Recap

KL-divergence: $\int p(x) \log \frac{p(x)}{q(x)} = \mathbb{E}_{x \sim p} \log \frac{p(x)}{q(x)}$

Gaussian entropy: $h(x) = \frac{1}{2} \log(2\pi e \cdot \sigma^2)$ (1-dim)

$h(x) = \frac{n}{2} \log(2\pi e) + \frac{1}{2} \log(\det(\Sigma))$

Maximum entropy

- Let X be a continuous random variable with dist. P s.t. $\mathbb{E}_P X = 0$, $\mathbb{E}_P XX^T = \Sigma$. (Say $X \in \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{n \times n}$)

  Let Y be an r.v. with distribution $Q = N(0, \Sigma)$.

  Then $h(X) \leq h(Y)$

$$D(P \| Q) \geq 0 \Rightarrow \underbrace{\mathbb{E}_{x\sim P} \log \frac{1}{q(x)}}_{\downarrow} \geq \mathbb{E}_{x\sim P} \log \frac{1}{p(x)} = h(X)$$

$$\mathbb{E}_{x\sim P} \frac{1}{\ln 2}\left[x^T \Sigma^{-1} x + \frac{1}{2} \ln\left((2\pi)^{n/2} \det(\Sigma)\right)\right]$$

$$= \mathbb{E}_{x\sim Q} \frac{1}{\ln 2}\left[x^T \Sigma^{-1} x + \frac{1}{2} \ln\left((2\pi)^{n/2} \det(\Sigma)\right)\right]$$

- For any r.v. X on $\mathbb{R}$, $h(X) \leq \frac{1}{2} \log\left(2\pi e \cdot \operatorname{Var}(X)\right)$

Continuous Fano

- Let X be a random variable and let $\hat{X}$ be an estimator

Then $\mathbb{E}[(X-\hat{X})^2] \geq \frac{1}{2\pi e} \cdot 2^{2h(X)}$

$$\forall a \quad \mathbb{E}[(X-a)^2] \geq \mathbb{E}[(X-\mu)^2] = \mathrm{Var}(X)$$

$$\mathbb{E}[(X-\hat{X})^2] \geq \mathrm{Var}(X)$$

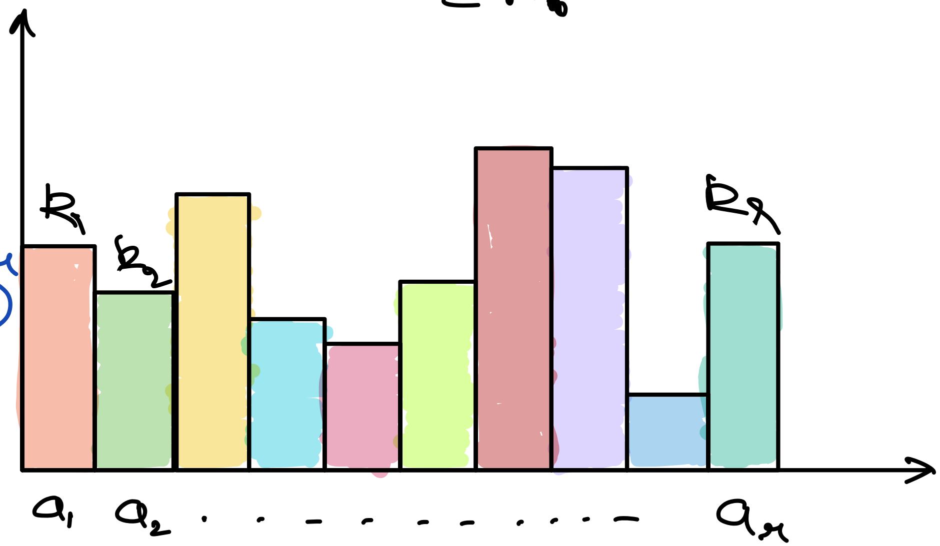
Ex: For $X \to Y \to \hat{X}$, $\mathbb{E}[(X-\hat{X}(Y))^2] \geq \frac{1}{2\pi e} 2^{2h(X|Y)}$

Concentration bounds via "Method of Types"

$|\mathcal{X}| = r$ $\quad$ n i.i.d. samples $\bar{x} = (x_1, \ldots, x_n)$

How likely is this sample?

$\sum k_i = n$



T_n : set of all n-types

$\le (n+1)^r$ $\qquad$ $\binom{n+r-1}{r-1} \le (n+1)^r$

$r-1$ dividers

$C_P : \{\bar{x} \mid P_{\bar{x}} = P\}$

$$P_{\bar{x}}(a_i) = \frac{|\{j \mid x_j = a_i\}|}{n}$$

Concentration inequalities

Chernoff/Hoeffding bounds

$$\mathbb{P}_{x_1 \dots x_n \in \{0,1\}} \left[\sum x_i \geq \left(\tfrac{1}{2} + \epsilon\right) \cdot n \right] \lessapprox 2^{-C\epsilon^2 n}$$

w.p. ½ each

$$\mathbb{P}_{\bar{x} \sim Q^n} [\text{Type of } \bar{x} \text{ is in the set } \Pi]$$

It's all about types

- Let Q be any distribution on $\mathcal{X}$.

  Let $\bar{x}, \bar{y} \in \mathcal{X}^n$ have type ($P_{\bar{x}} = P_{\bar{y}}$). Then

$$Q^n(\bar{x}) = Q^n(\bar{y})$$

$P(a_1) = \frac{k_1}{n} \quad \cdots \quad P(a_r) = \frac{k_r}{n}$

Proof:

$$Q^n(\bar{x}) = \prod_{j=1}^{n} Q(x_j)$$

$$= \prod_{i=1}^{r} (Q(a_i))^{k_i}$$

$$= Q^n(\bar{y})$$

<u>Ex:</u> $V = \begin{cases} 0 & \text{wp } 1/2 \\ 1 & \text{wp } 1/2 \end{cases}$ $\quad \bar{X} = \begin{cases} (x_1 \ldots x_n) \sim P_0^n & \text{if } V = 0 \\ (x_1 \ldots x_n) \sim P_1^n & \text{if } V = 1 \end{cases}$

Show that $g(\bar{x}) = P_{\bar{x}}$ is a sufficient statistic

Sequences of type P

- $\dfrac{2^{n \cdot H(P)}}{(n+1)^{r}} \leq \underbrace{|C_P|}_{} \leq 2^{n \cdot H(P)}$

$\hookrightarrow = |\{ \bar{x} \mid P_{\bar{x}} = P \}| = \#$ sequences of type P

Proof: Let $P(a_1) = \dfrac{k_1}{n}, \ldots\ldots, P(a_r) = \dfrac{k_r}{n}$ $\quad \sum k_i = n$

$$|C_P| \overset{?}{=} \frac{n!}{k_1! \cdots k_r!}$$

$$n^n = (k_1 + \cdots + k_r)^n = \sum_{j_1 + \cdots + j_r = n} \frac{n!}{j_1! \cdots j_r!} \cdot k_1^{j_1} \cdots k_r^{j_r}$$

Ex: Largest term has $j_1 = k_1, \ldots, j_r = k_r$

$$n^n = (k_1 + \cdots + k_r)^n \leq (n+1)^r \cdot \underbrace{\frac{n!}{k_1! \cdots k_r!}}_{|\mathcal{C}_p|} \cdot k_1^{k_1} \cdots k_r^{k_r}$$

$$1 \leq (n+1)^r \cdot |\mathcal{C}_p| \cdot \left(\frac{k_1}{n}\right)^{k_1} \cdots \left(\frac{k_r}{n}\right)^{k_r}$$

$$= (n+1)^r \cdot |\mathcal{C}_p| \cdot (P(a_1))^{n \cdot P(a_1)} \cdots (P(a_r))^{n \cdot P(a_r)}$$

$$= (n+1)^r \cdot |\mathcal{C}_p| \cdot 2^{n \cdot P(a_1) \log P(a_1) + \cdots + n \cdot P(a_r) \log P(a_r)}$$

$$\therefore 1 \leq (n+1)^r \cdot |\mathcal{C}_p| \cdot 2^{-n \cdot H(P)} \Rightarrow |\mathcal{C}_p| \geq \frac{2^{n \cdot H(P)}}{(n+1)^r}$$

Ex: Prove $|\mathcal{C}_p| \leq 2^{n \cdot H(P)}$

Probability of a type

- For any product distribution Q^n and any type $P \in \mathcal{T}_n$

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

Proof:

$$\Pr_{\bar{x}\sim Q^n}\left[P_{\bar{x}} = P \right] = \sum_{\bar{x}\in \mathcal{C}_P} Q^n(\bar{x})$$

$$= |\mathcal{C}_P| \cdot Q^n(\bar{x}) \qquad \bar{x} \text{ of type } P$$

$$= |\mathcal{C}_P| \prod \frac{Q(x_i)}{P(x_i)} \cdot P(x_i)$$

$$= \underbrace{|\mathcal{C}_P|}_{2^{nH(P)}} \cdot \underbrace{\frac{Q^n(\bar{x})}{P^n(\bar{x})}}_{2^{-nD(P\|Q)}} \cdot \underbrace{P^n(\bar{x})}_{2^{-nH(P)}}$$

$\bar{x}$ of type P, $\# a_1 = n \cdot P(a_1), \ldots\ldots, \# a_r = n \cdot P(a_r)$

$$P^n(\bar{x}) = P(a_1)^{n \cdot P(a_1)} \cdots P(a_r)^{n \cdot P(a_r)}$$

$$= 2^{-n \cdot H(P)}$$

$$\frac{Q^n(\bar{x})}{P^n(\bar{x})} = \left(\frac{Q(a_1)}{P(a_1)}\right)^{n \cdot P(a_1)} \cdots \left(\frac{Q(a_r)}{P(a_r)}\right)^{n \cdot P(a_r)}$$

$$= 2^{n \cdot P(a_1) \log \frac{Q(a_1)}{P(a_1)} + \cdots + n \cdot P(a_r) \cdot \log \frac{Q(a_r)}{P(a_r)}}$$

$$= 2^{-n \cdot D(P \| Q)}$$

Chernoff / Hoeffding bounds again

$\mathcal{X} = \{0, 1\}$ $\quad Q = \begin{cases} 0 & \text{wp } 1/2 \\ 1 & \text{wp } 1/2 \end{cases}$ $\quad P = \begin{cases} 0 & \text{wp } 1/2 - \epsilon \\ 1 & \text{wp } 1/2 + \epsilon \end{cases}$

$$\mathbb{P}_{\bar{x} \sim Q^n}[\bar{x} \text{ has type } P] = \mathbb{P}_{\bar{x} \sim Q^n}\left[x_1 + \cdots + x_n = \left(\tfrac{1}{2} + \epsilon\right) n\right]$$

$$\leq 2^{-n D(P \| Q)}$$

$$\mathbb{P}_{\bar{x} \sim Q^n}\left[x_1 + \cdots + x_n \geq \left(\tfrac{1}{2} + \epsilon\right) n\right] \leq (n+1) \cdot 2^{-n D(P \| Q)}$$