Class 7

2018年3月27日 星期二　10:04

Entropy

* H_p :

集合 $S=\{x_1,\cdots,x_r\}$ 上有分布 $P=P_s$ ，令 $P(x_i)=p_i$

$\Rightarrow H_p \triangleq -\sum_{i=1}^{r} p_i \cdot \log(p_i)$ ← $E(-\log P)$

* MacMillan Theorem :

$\Omega = \{ w=(w_1,w_2,\cdots) \}$ ↙确定了w_1后，w就已有一个存在概率
当只考虑w_1,w_2时等价于考虑 $S\times S$空间

MacMillan

~> $\Omega_n = \{(w_1,w_2,\cdots,w_n)\}$ (Ω_n,P_n) ← 有限集合

For every $\varepsilon>0$ and all sufficiently large n, one can find a subset $\Omega_n' \subseteq \Omega_n$

such that : 1. $e^{n(H-\varepsilon)} \le |\Omega_n'| \le e^{n(H+\varepsilon)}$

由2,3可推1 { 2. $\lim_{n\to\infty} P(\Omega_n')=1$

3. For every $w\in\Omega_n'$, we have $e^{-n(H+\varepsilon)} \le P(w) \le e^{-n(H-\varepsilon)}$

pf: Let $\Omega_n' = \{ w\in\Omega_n : \left|\frac{\nu_j(w)}{n}-p_j\right| \le \delta \}$ $(1\le j\le r)$

其中 $\nu_j(w)=\{i\in[n]: w_i=j\}$, $p_1,\cdots,p_r>0$

由 Chebyshev 不等式及弱大数率可知，对于 $\forall j$ $P(\left|\frac{\nu_j(w)}{n}-p_j\right|>\delta)=0$

2✓ 对其取补集可知 2 成立

下证由该方法构造的Ω_n'同时满足 3

$\because P(w)=P(w_1)\cdot P(w_2)\cdots P(w_n)$

$= p_1^{\nu_1(w)}\cdots p_r^{\nu_r(w)}$

$= e^{\sum \nu_i(w)\cdot\log p_i} = e^{n\cdot\sum_i \frac{\nu_i(w)}{n}\cdot\log p_i}$

$= e^{n\sum p_i\cdot\log p_i}\cdot e^{n\sum(\underbrace{\frac{\nu_i(w)}{n}-p_i}_{<\varepsilon\text{ uses 2}})\cdot\log p_i}$

3✓ $= e^{nH}\cdot e^{n\cdot O(\delta)}$

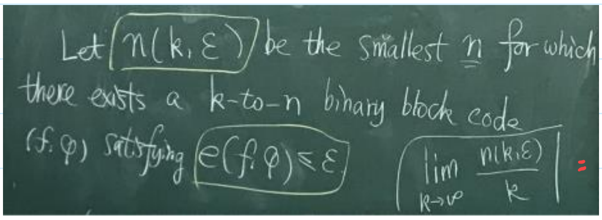
* Discrete Memoryless Source (DMS) with generic distribution P:

字 s 在信中有概率分布 (S,P)

$S^k \underset{\varphi}{\overset{f}{\rightleftarrows}} \{0,1\}^n$

定义 $\mathrm{Error}(f,\varphi)=P_k(\varphi\circ f(w)\ne w)$ $\begin{cases}(f,\varphi): \text{code}\\ w\in\Omega_k\end{cases}$

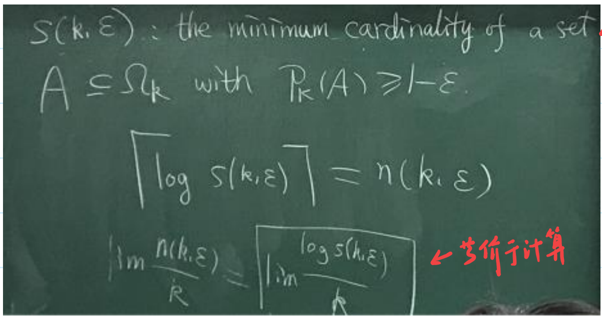
目标： $\frac{n}{k}\downarrow$, $\mathrm{Error}(f,\varphi)\downarrow$



= H(P)
极限即为Entropy

pf: $\exists (f,\varphi)$ s.t. $e(f,\varphi)\le\varepsilon$

$\Leftrightarrow \exists A \subseteq \Omega_k$, $P_k(A) \geq 1-\varepsilon$ 且 $|A| \leq 2^n$ 在空间 Ω_k 中找一个大概率事件占比尽可能小



← 等价于计算

Let $B(k,\delta)$ be the set of those sequences $w \in \Omega_k$ which have probability:

$$\because \ e^{-k(H+\delta)} \leq P(w) \leq e^{-k(H-\delta)}$$

由 MacMillion:

$$B(k,\delta) \supseteq \Omega_k'$$

$$\rightsquigarrow \begin{cases} \lim_{k\to\infty} P(B(k,\delta)) \to \lim_{k\to\infty} P(\Omega_k') = 1 & (*) \\ |B(k,\delta)| \leq e^{k(H+\delta)} & (**) \end{cases}$$

$\rightsquigarrow$ 由 $s(k,\varepsilon)$ minimum 性质，$s(k,\varepsilon) \leq B(k,\delta)$

$\leq$ ✓ 由 $(*)$ $\quad \overline{\lim_{k\to\infty}} \frac{1}{k} \log s(k,\varepsilon) \leq \overline{\lim_{k\to\infty}} \frac{1}{k} \log |B(k,\delta)|$

For every $A \subseteq \Omega_k$ with $P_k(A) \geq 1-\varepsilon$

由于 $P_k(A)$ 实际上可无限趋于 1，有:

$$P_k(A \cap B(k,\delta)) \geq \frac{1-\varepsilon}{2} \quad (k\to\infty)$$

$$\because |A| \geq |A \cap B(k,\delta)| \geq \sum_{w \in A\cap B(k,\delta)} P_k(w) \cdot e^{k(H-\delta)} \geq \frac{1-\varepsilon}{2} e^{k(H-\delta)}$$

$\geq$ ✓ $\quad \therefore \underline{\lim_{k\to\infty}} \frac{1}{k} \log s(k,\varepsilon) \geq \underline{\lim_{k\to\infty}} \frac{1}{k} \log e^{k(H-\delta)} \cdot \frac{1-\varepsilon}{2}$

由 ε-k 语言对上下极限说明: $\lim_{k\to\infty} \frac{\log s(k,\varepsilon)}{k} = \text{Entropy}$

※ 推广:

给 $S \to R^+$: mass function 赋上权重 M_i ← 给信中每一个字再赋上权重

$M(w) = M_1(w_1) \cdot M_2(w_2) \cdots M_k(w_k)$, $w \in \Omega_k$

$$s(k,\varepsilon) \triangleq \text{minimum of } \underset{\sum_{w\in A} M(w)}{\overset{\shortparallel}{M(A)}} \text{ for } A \subseteq \Omega_k \text{ with } P_k(A) \geq 1-\varepsilon$$

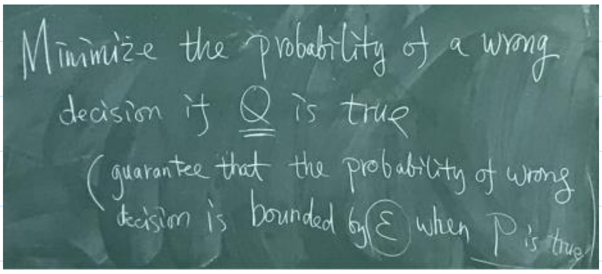
$$\Rightarrow \lim_{k\to\infty} \left(\frac{\log s(k,\varepsilon)}{k} - E_k \right) = 0 \qquad E_k \triangleq \frac{1}{k} \sum_{i=1}^{k} \sum_{x\in S} P_i(x) \log \frac{M_i(x)}{P_i(x)}$$ ← 结论

Coding Problem

※ 统计问题:

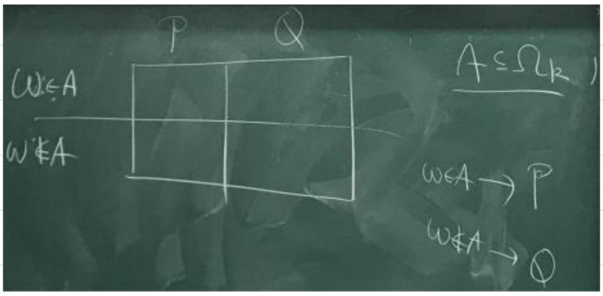
$$\text{现有概率分布} \begin{cases} P = \{ P(x) : x \in X \} \\ Q = \{ Q(x) : x \in X \} \end{cases} \to D_{P\|Q}$$

样本空间中抽 k 次，$w = w_1 \cdots w_k$, $A \subseteq \Omega_k$ $\quad (A = \{w\})$ $\begin{matrix} \nearrow P \\ \searrow Q \end{matrix}$



$$\beta(k,\varepsilon) = \min_{\substack{A\subseteq\Omega_k \\ P_k(A)\ge 1-\varepsilon}} Q_k(A)$$

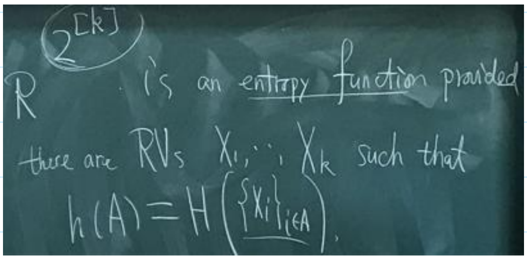
找到A使 $(w\notin A, P)\le\varepsilon$



$$D_{P\|Q} = \lim_{k\to\infty}\frac{\log\beta(k,\varepsilon)}{k} = -\sum_{x\in S} P(x)\log\frac{P(x)}{Q(x)}$$

Predict on Sampling

HW:



$$\text{polymatroid:}\begin{cases} h(\emptyset)=0 \\ h(A)\le h(B),\ A\subseteq B \\ h(A\cup B)+h(A\cap B)\le h(A)+h(B) \end{cases}$$

Homework