جلسه ۱۲ تئوری اطلاعات : ۱۵ آبان ۹۹

* بهینه بودن کدهای Huffman

* فصل 2 ← مجموعه های نوعی

← قانون اعداد بزرگ

← AEP

← مجموعه های نوعی typical set

← یک اثبات دیگر برای قضیه کدگذاری منبع

یادآوری لم جلسه قبل :

برای هر توزیع گسسته‌ای بر روی الفبای $\mathcal{X}$: $\quad p_1 \geq \dots \geq p_m$, $m = |\mathcal{X}|$

① در یک کد بهینه خاصیت زیر را دارد :

$$p_j > p_k \Rightarrow l_j \leq l_k$$

② در یک کد بهینه لحظه‌ای (prefix free) دو طولانی‌ترین کدواژه‌ها هم طول هستند.

③ یک کد لحظه‌ای بهینه وجود دارد که دو تا طولانی‌ترین کدواژه‌ها فقط در بیت آخر تفاوت دارند.

به کد بهینه‌ای که هر سه خواص لم بالا را داشته باشد ← Canonical

Canonical $\longleftarrow$ به کد بهینه‌ای که ~~...~~ خواص لم بالا را داشته باشد

$$\mathcal{X} \longrightarrow m=|\mathcal{X}| \longrightarrow \underline{P}=(P_1,\dots,P_m) \quad : \quad P_1 \geqslant \dots \geqslant P_m$$

Huffman reduction $\quad \underline{P}'=(P_1,\dots,P_{m-2},P_{m-1}+P_m) \quad : |\mathcal{X}'|=m-1$

$C^{(m)}_{can}(\underline{P})$ یک کد (بهینه) کانونی روی $\underline{P}$ در نظر می‌گیریم :

$C^{(m-1)}_{opt}(\underline{P}')$ یک کد بهینه برای توزیع $\underline{P}'$ داریم :

کار اول : کد بهینه $C_{opt}(\underline{P}')$

کار اول : کد بهینه $C_{opt}(p')$

p_1	$\longrightarrow$	w'_1	l'_1
p_2	$\longrightarrow$	w'_2	l'_2
$\vdots$		$\vdots$	
p_{m-2}		w'_{m-2}	l'_{m-2}
$p_{m-1}+p_{m-2}$		w'_{m-1}	l'_{m-1}

$\xRightarrow[\text{extend}]{\text{توسعه}}$

p_1		l'_1
p_2		
p_{m-2}		l'_{m-2}
p_{m-1}	$w'_{m-1} \lozenge 0$	$l'_{m-1}+1$
p_m	$w'_{m-1} \lozenge 1$	$l'_{m-1}+1$

$$\boxed{\bar{L}_{ext}(p) = \bar{L}_{opt}(p') + p_{m-1} + p_m}$$

کار دوم: کد $C_{can}(p)$ ← با merge کردن یک کد جدید می‌سازیم

$$\bar{L}_{mrg}(p') = \sum_{i=1}^{m-2} p_i l_i + p_{m-1}(l_{m-1}-1) + p_m(l_m - 1)$$

$$= \sum_{i=1}^{m} p_i \cdot l_i - p_{m-1} - p_m = \bar{L}_{can}(p) - p_{m-1} - p_m$$

$$\Rightarrow \bar{L}_{ext}(p) + \bar{L}_{mrg}(p') = \bar{L}_{opt}(p') + \bar{L}_{can}(p)$$

$$\Rightarrow \underbrace{\left[\bar{L}_{ext}(p) - \bar{L}_{can}(p)\right]}_{\geqslant 0} + \underbrace{\left[\bar{L}_{mrg}(p') - \bar{L}_{opt}(p')\right]}_{\geqslant 0} = 0$$

$\Rightarrow \bar{L}_{ext}(p)$: is an opimal code on $\mathcal{X}$

$\Rightarrow \bar{L}_{ext}(p)$: is an opimal code on $\mathcal{X}$

در نهایت یک قضیه زیر داریم

قضیه: کد هافمن بهینه است یعنی:

$$\bar{L}_{Huff} \leq \bar{L}_C$$

$\hookrightarrow$ any other uniqely decodable codes on $\mathcal{X}$

فصل 2 کتاب Cover:

AEP

مثال: $X_i \sim \text{Bern}(p)$

$$\underline{X} = (X_1, \dots, X_n) \overset{iid}{\sim} \text{Bern}(p)$$

k: تعداد یک ها در رشته $\underline{x}$

$$P(\underline{X}) = p^k (1-p)^{n-k}$$

$$\downarrow$$

$$\in \{0,1\}^n$$

$$\mathbb{P}[\text{یک رشته با تعداد } k \text{ یک}] = P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

تقریب استرلینگ:

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

$$\simeq \left(\frac{n}{e}\right)^n$$

$$\binom{n}{k} = \frac{(n/e)^n}{\left(\frac{k}{e}\right)^k \left(\frac{n-k}{e}\right)^{(n-k)}} = \frac{n^n}{k^k \cdot (n-k)^{(n-k)}}$$

1-4630

$$\binom{n}{k} = \frac{\left(\frac{n}{e}\right)^n}{\left(\frac{k}{e}\right)^k \left(\frac{n-k}{e}\right)^{(n-k)}} = \frac{n^n}{k^k \cdot (n-k)^{(n-k)}}$$

$$= \left(\frac{k}{n}\right)^{-k} \cdot \left(\frac{n-k}{n}\right)^{-(n-k)}$$

فرض کنیم که $\boxed{k = n.p}$

$$\boxed{= p^{-k} \cdot (1-p)^{-(n-k)}}$$

$$= p^{-np} \cdot (1-p)^{-n(1-p)} = 2^{-np \lg p} \cdot 2^{-n(1-p)\lg(1-p)}$$

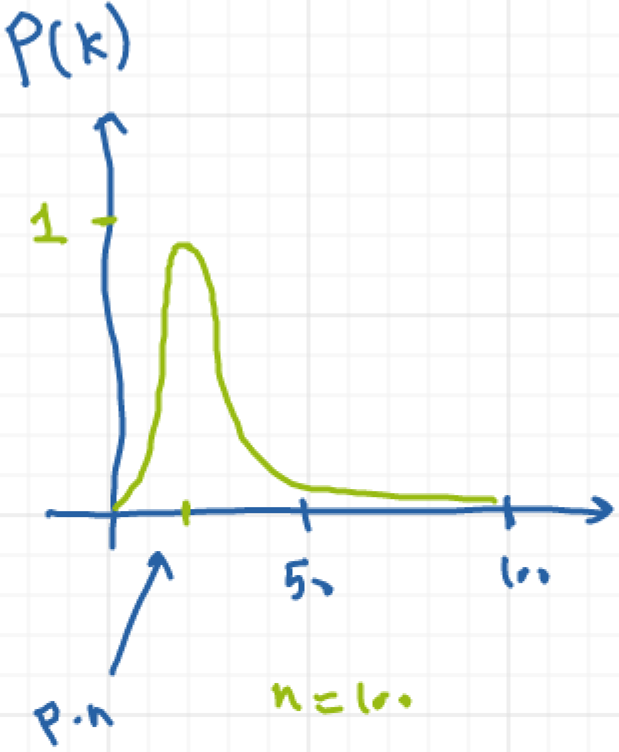
$$= \underbrace{2^{+n h_2(p)}}$$

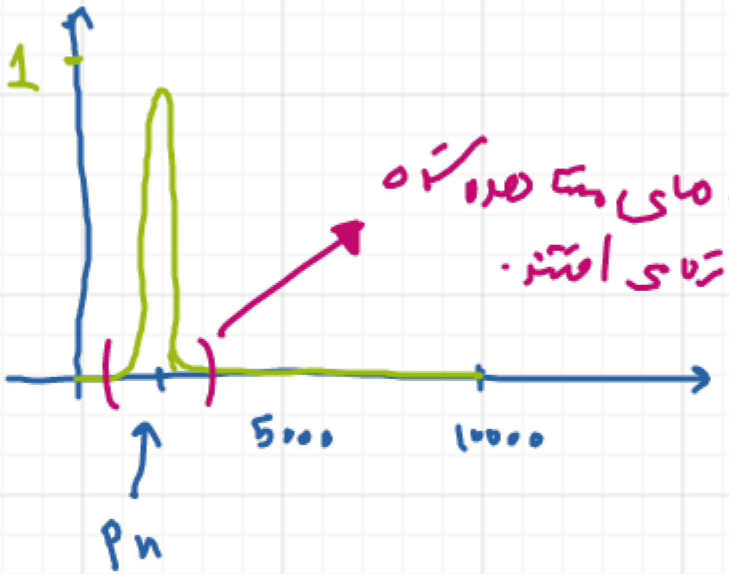
$h_2(p) \rightarrow$ آنتروپی X

تقریبا تعداد رشته های به طول np که همه احتمال را به خودشان نسبت داده اند

$$P(k)\Big|_{k=np} \simeq 1$$

مثال : ترسیم کنید $p = ۰.۲$





همه رشته های محتمل عملاً
در این بازه‌ای هستند.

یادآوری: قانون اعداد بزرگ (Weak Law of Large Number):

$X_1 \cdots X_n$ iid : $P(x)$

یک متغیر تصادفی $\leftarrow \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{n \to \infty} \mathbb{E}[X]$ in prob. ; $X \sim P(x)$

حد گرفتن در احتمال:

$$\lim_{n \to \infty} \mathbb{P}\left[\left| \bar{X}_n - \mathbb{E}[X] \right| > \varepsilon \right] \longrightarrow 0$$

قضیه: اگر $X_1, X_2 \dots X_n \overset{iid}{\sim} p(x)$ آنگاه:

$$-\frac{1}{n}\log \underbrace{P(X_1,\dots,X_n)}_{R.V.} \xrightarrow{n\to\infty} H(X) \quad \text{in prob.}$$

اثبات:

$$-\frac{1}{n}\log P(X_1\dots X_n) \overset{iid}{=} -\frac{1}{n}\log\big(P(X_1)\cdots P(X_n)\big)$$

$$= -\frac{1}{n}\sum_{i=1}^{n} \overbrace{\log P(X_i)}^{=Y_i} \xrightarrow[\text{wLLN}]{n\to\infty} -\mathbb{E}[Y_i]$$

$$\xrightarrow{n\to\infty} -\mathbb{E}[\log P(X_i)]$$

$$= \mathbb{E}[-\log P(X_i)] = H(X)$$

$$P(X_1,\dots,X_n) \simeq 2^{-nH(X)} \longleftarrow \text{نتیجه}$$