# تئوری اطلاعات: جلسه هفتم، ۹۹/۷/۲۱

Cover ← Sufficient Statistics, Fano's Inequality

مباحث علوم کامپیوتر ← Shearer's Lemma, Counting Primes using entropy, Sums of Binomial Coefficients

---

* مرور نامساوی پردازش:

$$X - Y - Z$$

$$X \perp Z \mid Y \longleftrightarrow p(x,z|y) = p(x|y)\cdot p(z|y)$$

Data processing Ineq. $I(X;Y) \geq I(X;Z)$

شرط تساوی: $X - Z - Y$

* آمار T کافی : Sufficient Statistics

فرض : A family of PMF $\{f_\theta(x)\}$

$$\underbrace{\{f_\theta(x)\}}_{\text{unknown}} \xrightarrow{\text{sample}} X \longrightarrow \underbrace{T(X)}_{\text{آمار یا Statistics}}$$

تابعی از مشاهدات X

با توجه به ساختار مسئله داریم :

$$\theta - X - T(X)$$

نامساوی پردازش $\Longrightarrow$

$$I(\theta; T(X)) \leq I(\theta; X)$$

* تعریف آمار کافی : یک تابع $T(X)$ را آمار کافی برای خانواده $\{f_\theta\}$ می‌گوییم اگر که $X \perp \theta \mid T(X)$ برای هر توزیع روی $\theta$. به بیان دیگر

$$\theta - T(X) - X$$

در نتیجه برای آمار کافی $T(X)$ داریم :

$$I(\theta; T(X)) = I(\theta; X)$$

مثال : مرحق کنیم

iid $\quad \boxed{X_1, \dots, X_n} \quad , \quad X_i \in \{0, 1\}$

$X \longleftrightarrow \boxed{X_1, \dots, X_n}$

$$\mathbb{P}[X_i = 1] = \theta$$

$$T(X_1, \dots, X_n) = \sum_{i=1}^{n} X_i$$

$$\mathbb{P}\Big[(X_1, \dots, X_n) = (x_1, \dots, x_n) \;\Big|\; \overbrace{\sum X_i}^{T(X)} = k \;, \theta\Big] = \underbrace{\begin{cases} \dfrac{1}{\binom{n}{k}} & : \sum x_i = k \\ 0 & : \text{o.w.} \end{cases}}_{\text{مستقل از } \theta}$$

$$\Longleftrightarrow \quad X - T(X) - \theta$$

$$= \mathbb{P}\big[(X_1, \dots, X_n) = (x_1, \dots, x_n) \;\big|\; T(X) = k\big]$$

* تعریف : آمار کافی کمینه (minimal sufficient statistics) :

$T(X)$ یک آمار کافی کمینه است اگر که تابعی از هر آماره‌های کافی دیگر مثل $U(X)$ باشد. گویا زنجیره مارکف زیر را داریم :

$$X - U(X) - T(X) - \theta$$

* نامساوی فانو :

فرض کنید $(X,Y) \sim P_{XY} = P_X(x)\cdot P_{Y|X}(y|x)$

X را مستقیما مشاهده نمی‌کنیم. می‌خواهیم از روی Y تخمینی برای X به دست بیاوریم.

$\hat{X} = g(Y)$ به عنوان تخمین X و می‌خواهیم درمورد خطای این تخمین صحبت کنیم.

$$X - Y - \hat{X}$$

$$P_e = \mathbb{P}[\hat{X} \neq X]$$

$\uparrow$

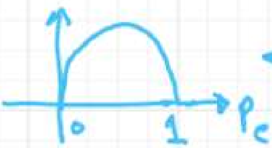
احتمال خطای تخمین

حالت خاص : $P_e = 0 \xleftarrow{\text{شهودی}} H(X|Y) = 0$

قضیه (نامساوی فانو):

برای تخمینگر $\hat{X}$ که داشته باشیم $X - Y - \hat{X}$ و $P_e = \mathbb{P}[X \neq \hat{X}]$

داریم:




$$\longleftarrow h_2(P_e) + P_e \log_2 |\mathcal{X}| \geq H(X|\hat{X}) \geq H(X|Y)$$

نتایج: ① اگر $P_e$ کوچک باشد $\longleftarrow$ $\underbrace{H(X|Y)}$ هم کوچک است.

نتایج: ① اگر $P_e$ کوچک باشد ← $H(X|Y)$ هم کوچک است.

یعنی $Y$ حرف های زیادی درمورد $X$ می زند.

② اگر $H(X|Y)$ بزرگ باشد (یعنی $Y$ حرف زیادی درمورد $X$ نزند)

← $P_e$ بزرگ

اثبات:

$$E = \begin{cases} 1 & : \hat{X} \neq X \\ 0 & : \hat{X} = X \end{cases} \to P_e = \mathbb{P}[E=1]$$

متغیر نشانگر خطا ← $E$

$$H(E, X | \hat{X}) = H(X|\hat{X}) + \overbrace{H(E|X, \hat{X})}^{=0}$$

از طرف دیگر

$$= \underbrace{H(E|\hat{X})}_{\substack{\leq H(E) \\ = h_2(P_e)}} + \underbrace{H(X|E, \hat{X})}_{?}$$

$$H(X|E, \hat{X}) = \underbrace{H(X|E=0, \hat{X})}_{=0} \underbrace{\mathbb{P}[E=0]}_{(1-P_e)} + \underbrace{H(X|E=1, \hat{X})}_{\substack{\leq H(X) \\ \leq \log_2 |\mathcal{X}|}} \cdot \overbrace{\mathbb{P}[E=1]}^{P_e}$$

$$\leq P_e \cdot \log_2 |\mathcal{X}|$$

$$\Longrightarrow H(X|Y) \leq \overbrace{H(X|\hat{X}) \leq h_2(P_e) + P_e \log_2 |\mathcal{X}|}$$

$$\leq P_e \cdot \log_2|\mathcal{X}| \qquad\qquad \leq \log_2|\mathcal{X}|$$

$$\Rightarrow \quad H(X|Y) \leq H(X|\hat{X}) \leq h_2(P_e) + P_e \log_2|\mathcal{X}|$$

$$I(X;Y) \geq I(X;\hat{X}) \quad \Longleftarrow \quad X - Y - \hat{X}$$ از صورت مسئله

$$H(X|Y) \leq H(X|\hat{X})$$

---

* شمارش اعداد اول با استفاده از مفاهیم آنتروپی:

$\pi(n)$ : تعداد اعداد اول کمتر مساوی $n$.

$$\pi(n) \xrightarrow{n \to \infty} \infty$$

Prime number theorem (PNT):

$$\pi(n) \sim \frac{n}{\ln(n)} \quad : n \longrightarrow \infty$$

اثبات خیلی راحت برای نامتناهی بودن اعداد اول و یک کران پایین برای $\pi(n)$.

اعداد $\{1,2,\dots,n\}$ را در نظر بگیرید و سپس $N$ را به صورت یکنواخت از این مجموعه

انتخاب کنید. متغیر تصادفی ←

متغیر تصادفی → $N = P_1^{X_1} \cdot P_2^{X_2} \cdots P_{\pi(n)}^{X_{\pi(n)}}$

$$2^{X_i} \leqslant P_i^{X_i} \leqslant N \leqslant n$$

$$\Rightarrow \quad X_i \leqslant \log_2(n) \qquad \forall\, i \in [1 : \pi(n)]$$

$$\log_2(n) = H(N) = H(X_1, \dots, X_{\pi(n)})$$

$$\leqslant H(X_1) + \dots + H(X_{\pi(n)})$$

$$\leqslant \pi(n) \log_2\left(1 + \log_2(n)\right) \quad \Rightarrow \quad \pi(n) \geqslant \frac{\log_2(n)}{\log_2\left(1 + \log_2(n)\right)}$$

$$\Rightarrow \quad \pi(n) \xrightarrow{n \to \infty} \infty$$