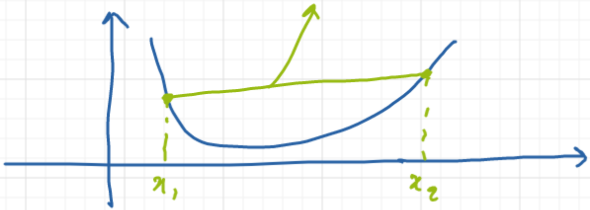
تئوری اطلاعات : جلسه پنجم ، 99/7/14 :

یادآوری : توابع محدب (convex)

$f(x)$ بر روی بازه (a,b) محدب است:

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \underline{\lambda f(x_1) + (1-\lambda) f(x_2)} \qquad \forall x_1, x_2 \in (a,b), \quad \forall \lambda \in [0,1]$$



نامساوی ینسن   Jensen's Ineq.

$f(x)$: یک تابع محدب و X یک متغیر تصادفی دلخواه

$f(x)$: یک تابع محدب و X یک متغیر تصادفی دلخواه

$$f\left(\mathbb{E}[X]\right) \leq \mathbb{E}\left(f(X)\right)$$

اگر f اکیداً محدب باشد

$$\boxed{\text{حالت تساوی} \iff \text{با احتمال } 1 \quad X=\mathbb{E}[X] \quad \text{یعنی } X \text{ یک متغیر تصادفی ثابت باشد.}}$$

اثبات برای متغیر های تصادفی گسسته :

استقراء روی اندازه الفبای $\mathcal{X}$ می‌زنیم:

ابتدا با فرض $|\mathcal{X}|=2$:

$$\mathcal{X}=\{x_1, x_2\}, \quad x_1 \to p_1, \quad x_2 \to p_2 \longrightarrow p_1+p_2=1$$

$$\Rightarrow f\left(\mathbb{E}[X]\right) = f\left(p_1x_1+\overset{(1-p_1)}{p_2}x_2\right) \leq p_1 f(x_1) + p_2 f(x_2) = \mathbb{E}\left[f(X)\right]$$

Whiteboard 2162120-5114-5080

برای $|\mathcal{X}|=n$

$$\mathcal{X}=\{x_1,\dots,x_n\}$$
$$\begin{array}{ccc} x_1 & & x_n \\ \downarrow & & \downarrow \\ p_1 & & p_n \end{array}$$

$$f(\mathbb{E}[X]) = f(p_1x_1+\dots+p_nx_n) = f\left(\frac{\overbrace{p_1}^{q_1}x_1+\dots+\overbrace{p_{n-1}}^{q_{n-1}}x_{n-1}}{1-p_n}(1-p_n)+p_nx_n\right)$$

$$\leq (1-p_n)\,\underline{f(q_1x_1+\dots+q_{n-1}x_{n-1})}+p_nf(x_n)$$

$$\leq (1-p_n)\left[q_1f(x_1)+\dots+q_{n-1}f(x_{n-1})\right]+p_nf(x_1)$$

$$=\sum_{i=1}^{n}p_if(x_i) = \mathbb{E}[f(X)]$$

* نامنفی بودن آنتروپی نسبی :

$$D(p\,\|\,q)\geq 0$$
$$\forall\ p,q \text{ over } \mathcal{X}$$

Whiteboard 2162120-5114-5080

$$-D(p\|q) = -\sum_{x\in\mathcal{X}} p(x)\log\frac{p(x)}{q(x)}$$

$$= -\mathbb{E}\left[\log\frac{p(X)}{q(X)}\right] \leq \log \mathbb{E}\,\frac{p(X)}{q(X)}$$

$$= \mathbb{E}\left[\log\frac{q(X)}{p(X)}\right]$$

$$\frac{p(X)}{q(X)} = c \longleftrightarrow$$

$$\leq \log \mathbb{E}\left[\frac{q(X)}{p(X)}\right] = \log \sum_{x\in\mathcal{X}:\ p(x)\neq 0} p(x)\cdot\frac{q(x)}{p(x)}$$

$$p(x) = c\cdot q(x)$$

$$= \log \sum_{x\in\mathcal{X}:\ p(x)\neq 0} q(x) \leq \log 1 = 0$$

$$\sum_{x:\ p(x)\neq 0}$$

$$\Rightarrow \boxed{D(p\|q)\geq 0}$$

$$1 = c\cdot 1$$

$$\Rightarrow \boxed{D(p\|q) \geq 0}$$

$$1 = C \cdot 1$$

$$\Leftrightarrow C = 1 \Leftrightarrow p(x) = q(x)$$

شرط تساوی :

$$p = q \Leftrightarrow D(p\|q) = 0$$

اطلاعات متقابل نامنفی است :

$$I(X;Y) = D\left(p_{XY}(x,y) \,\|\, p_X(x) \cdot p_Y(x)\right) \geq 0$$

$$I(X;Y) = 0 \Leftrightarrow X \perp Y$$

$$I(X;Y|Z) = D\left(p(x,y|z) \,\|\, p(x|z) \cdot p(y|z)\right) \geq 0$$

کران بالا برای $H(X)$:

کران بالا برای $H(X)$:

اگر X یک متغیر تصادفی با الفبای $\mathcal{X}$ باشد آنگاه داریم:

$$H(X) \leq \log_2 |\mathcal{X}|$$

• ابتدا دقت کنید توزیع یکنواخت $u(x) = \frac{1}{|\mathcal{X}|}$: $x \in \mathcal{X}$

$$H(U) = -\sum_{x} \frac{1}{|\mathcal{X}|} \log_2 \frac{1}{|\mathcal{X}|} = \log_2 |\mathcal{X}|$$

• برای هر توزیع $p(x)$ $\leftarrow$ $H(X) \leq \log_2 |\mathcal{X}|$

$$0 \leq D(p \,\|\, u) = \sum_{x \in \mathcal{X}} p(x) \log_2 \frac{p(x)}{u(x)}$$

$$= \sum_{x} p(x) \log_2 p(x) - \sum_{x} p(x) \log_2 \frac{1}{|\mathcal{X}|}$$

Whiteboard 2162120-5114-5080

$$= -H(X) + \log_2 |\mathcal{X}| \Rightarrow H(X) \leq \log_2 |\mathcal{X}|$$

• تعبیر :

(Information can't hurt)

(Conditioning Reduces Entropy)

$$0 \leq I(X;Y) = \begin{cases} H(X) - H(X|Y) \rightarrow \\ H(Y) - H(Y|X) \rightarrow \end{cases}$$

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

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

تساوی

Side Information

$$X \perp Y \iff H(X|\underset{\uparrow}{Y}) = H(X)$$

$$\iff H(Y|X) = H(Y)$$

لزوما رابطه زیر برقرار نیست

$$H(X|Y=y) \cancel{\leq} H(X)$$

---

* کران استقلال :

$$(X_1, \dots, X_n) \sim p(x_1, \dots, x_n)$$

Whiteboard 2162120-5114-5080

۴. کران استقلال: $(X_1,\dots,X_n) \sim p(x_1,\dots,x_n)$

$$0 \leq \underbrace{H(X_1,\dots,X_n)}_{} \leq \sum_{i=1}^{n} H(X_i)$$

$$= H(X_1) + \underline{H(X_2|X_1)} + \cdots + \underline{H(X_n|X_{n-1},\dots,X_1)}$$

chain rule

$$\leq H(X_1) + H(X_2) + \cdots + H(X_n)$$

تساوی

$$p(x_1 \dots x_n) = p(x_1)\cdots p(x_n)$$

مثال: $P_X \sim X = X_1 = \dots = X_n$

$$H(X_1,\dots,X_n) = H(X)$$

$$\leq \sum_{i=1}^{n} H(X_i) = n\,H(X)$$

تمرین : log sum : log sum Inequality

Whiteboard 2162120-5114-5080

فرض کنید $\left\{\begin{matrix} a_1 \cdots a_n \\ b_1 \cdots b_n \end{matrix}\right.$ اعداد نامنفی هستند.

$$\sum_{i=1}^{n} a_i \lg \frac{a_i}{b_i} \geq \left(\sum_{i=1}^{n} a_i\right) \lg \frac{\left(\sum_{i=1}^{n} a_i\right)}{\left(\sum_{i=1}^{n} b_i\right)}$$

تساوی $\iff \frac{a_i}{b_i} = \text{Const}$

$$-\sum_{i=1}^{n} a_i \ln \frac{a_i}{b_i} = \sum_{i=1}^{n} a_i \ln \frac{b_i}{a_i}$$

$$= \frac{\left(\sum_{j=1}^{n} a_j\right)}{\left(\sum_{j=1}^{n} a_j\right)} \sum_{i=1}^{n} a_i \lg \frac{b_i}{a_i}$$

$$= \left(\sum_{j=1}^{n} a_j\right) \sum_{i=1}^{n} \boxed{\frac{a_i}{\left(\sum_{j=1}^{n} a_j\right)}} \lg \frac{b_i}{a_i}$$

احتمال $= q_i = \frac{a_i}{\sum_j a_j}$

Whiteboard 2162120-5114-5080

$$= \frac{\;}{\left(\sum_{j=1}^{n} a_j\right)} \sum_{i=1} a_i \lg \frac{b_i}{a_i}$$

$$q(i) = q_i = \frac{a_i}{\sum_j a_j}$$

$$= \left(\sum_{j=1}^{n} a_j\right) \underbrace{\sum_{i=1}^{n} \boxed{\frac{a_i}{\left(\sum_{j=1}^{n} a_j\right)}} \lg \frac{b_i}{a_i}}_{\mathbb{E}_q \left[\lg \frac{b_i}{a_i}\right] \to \mathbb{E}_{I \sim q}\left[\lg \frac{b_I}{a_I}\right]}$$

Jensen's Ineq

$$\le \left(\sum_{j=1}^{n} a_j\right) \lg \mathbb{E}_{I \sim q}\left[\frac{b_I}{a_I}\right]$$

$$= \left(\sum_{j=1}^{n} a_j\right) \lg \left(\sum_{i=1}^{n} \frac{\cancel{a_i}}{\left(\sum_{j=1}^{n} a_j\right)} \cdot \frac{b_i}{\cancel{a_i}}\right)$$

$$= \left(\sum_{i=1}^{n} a_i\right) \lg \left(\frac{\sum_{i=1}^{n} b_i}{\sum_{i=1}^{n} a_i}\right) \quad \checkmark$$