1) (a) $H(X) = \sum_x p(x) \log_2 \left(\frac{1}{p(x)}\right)$

$0 \le p(x) \le 1 \Rightarrow \frac{1}{P(x)} \ge 1 \Rightarrow \log_2 \frac{1}{P(x)} \ge \cancel{\log_2} \nearrow 0$

$\downarrow$ non-negative ①

$\Rightarrow \log_2 \frac{1}{P(x)} \ge 0$ ② (positive and zero)

① and ② $\Rightarrow \boxed{\sum_x p(x) \log_2 \frac{1}{P(x)} \ge 0}$

---

1) (b) Take $Y(x) = \frac{q(x)}{p(x)}$

Take $\Phi(x) = \log_2(X) \rightarrow$ concave

$E[\Phi(Y)] = \sum_x p(x) \log_2 Y(x) = \sum_x p(x) \log_2 \frac{q(x)}{P(x)} = -\sum_x p(x) \log_2 \frac{p(x)}{q(x)}$

$\Phi[E[Y]) = \log_2 \sum_x p(x)\, Y(x) = \log_2 \sum_x \cancel{p(x)} \frac{q(x)}{\cancel{p(x)}} = \log_2 1 = 0$

by Jensen's inequality (since $\log_2(x)$ is ~~convex~~ concave):

$\Phi(E[x]) \le E[\Phi(x)] \Rightarrow \boxed{0 \le -\left(-\sum_x p(x) \log_2 \frac{p(x)}{q(x)}\right) = KL(p\|q)}$

---

1) (c) $KL(p(x,y)\,||\,p(x)p(y)) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$

$= \sum_{x,y} p(x,y)\left(\log \frac{p(x,y)}{p(x)} - \log p(y)\right)$

$= \underbrace{\sum_{x,y} p(x) \underbrace{p(Y|X=x) \log(y|x)}_{H(Y|X=x)}}_{-H(Y|X)} - \underbrace{\sum_y \log p(y) \overbrace{\sum_x p(x,y)}^{P(y)}}_{-H(Y)}$

$= -(-H(Y)) - H(Y|X) = \boxed{H(Y) - H(Y|X) = I(Y;X)}$

Mahsa Naserifar
CSC411 - Fall 18

Scanned by CamScanner

2) ~~a~~ Hypothesis $\frac{1}{2}\left(\frac{1}{m}\sum_{i=1}^{m}(h_i(x)-t)\right)^2 \leq \frac{1}{m}\sum_{i=1}^{m}\frac{1}{2}(h_i(x)-t)^2$

consider: $Y(x) = h_i(x) - t$

$$\Rightarrow E[Y(x)] = \frac{1}{m}\sum_{i=1}^{m}(h_i - t) \Rightarrow \Phi(E[Y]) = \frac{1}{2}\left(\frac{1}{m}\sum_{i=1}^{m}(h_i-t)\right)^2$$

consider: $\Phi(Y) = \frac{(\overbrace{h_i - t}^{Y})^2}{2} \Rightarrow E[\Phi(Y)] = \frac{1}{m}\sum_{i=1}^{m}\frac{(h_i-t)^2}{2}$

$\frac{Y^2}{2}$ is convex

by the Jensen's inequality: $\Phi(E[Y]) \leq E[\Phi(Y)]$

$$\Rightarrow \boxed{\underbrace{\frac{1}{2}\left(\frac{1}{m}\sum_{i=1}^{m}(h_i-t)\right)^2}_{L(\bar{h},t)} \leq \frac{1}{m}\sum_{i=1}^{m}\underbrace{\frac{(h_i-t)^2}{2}}_{L(h_i,t)}}$$

---

3) consider $w_b = \sum_{i=1}^{N} w_i \mathbb{I}\{h_t \neq t\}$ and consider $w = \sum_{i=1}^{N} w_i$

and $w_g = \sum_{i=1}^{N} w_i \mathbb{I}\{h_t = t\}$

$$err' = \frac{w_b'}{w'} = \frac{w_b e^{(-\alpha)\times(-1)}}{w_b e^{(-\alpha)(-1)} w_g^{-\alpha}} = \frac{w_b e^{\alpha}}{w_b e^{\alpha} + w_g e^{-\alpha}}$$

$$= \frac{w_b\sqrt{\frac{1-err}{err}}}{w_b\sqrt{\frac{1-err}{err}} + w_g\sqrt{\frac{err}{1-err}}}$$

$$\because \sqrt{\frac{1-err}{err}} = \sqrt{\frac{1-\frac{w_b}{w}}{\frac{w_b}{w}}} = \sqrt{\frac{\frac{w-w_b}{w}}{\frac{w_b}{w}}} = \sqrt{\frac{w_g}{w_b}}$$

$$\because \sqrt{\frac{err}{1-err}} = \sqrt{\frac{\frac{w_b}{w}}{\frac{w_g}{w}}} = \sqrt{\frac{w_b}{w_g}}$$

$$\because err' = \frac{w_b\sqrt{\frac{w_g}{w_b}}}{w_b\sqrt{\frac{w_g}{w_b}} + w_g\sqrt{\frac{w_b}{w_g}}} = \frac{\cancel{\sqrt{w_b w_g}}}{2\cancel{\sqrt{w_b w_g}}} \boxed{= \frac{1}{2}}$$

Mahsa Naserifar
CSC411 — Fall 18