②

a) $D_{KL} = [p(x) \| q(x)] = D_{KL} = [p(x) \| q(x)]$

$$\Leftrightarrow \quad \mathbb{E}_p\left[\log \frac{p(x)}{q(x)}\right] = \mathbb{E}_q\left[\log \frac{q(x)}{p(x)}\right]$$

let $q(x) \begin{cases} 0.2 & \text{if } x=0 \\ 0.8 & \text{if } x=1 \end{cases}$ $\quad p(x) = \begin{cases} 0.5 & \text{if } x=0 \\ 0.5 & \text{if } x=1 \end{cases}$

$$\Rightarrow 0.5 \log \frac{0.5}{0.2} + 0.5 \log \frac{0.5}{0.8} = 0.8 \log \frac{0.8}{0.5} + 0.2 \log \frac{0.2}{0.5}$$

$$\approx\!> \quad 0{,}223 = 0{,}193 \quad ⚡$$

b)

$$\arg\min_\theta D_{KL}[p_{true}(x) \| p(x|\theta)] = \arg\min_\theta \mathbb{E}\left[\log \frac{p_{true}(x)}{p(x|\theta)}\right]$$

$$= \arg\min_\theta \mathbb{E}[\log p_{true}(x) - \log p(x|\theta)]$$

$$= \arg\max_\theta \mathbb{E}[\log p(x|\theta)]$$

$$\arg\max_\theta p(\{x_n\}_{n=1}^N|\theta) = \quad = \arg\max_\theta \frac{1}{N}\sum_{n=1}^N \log p(x_n|\theta) - \underbrace{\frac{1}{N}\sum_{n=1}^N \log p(x_n|\theta) + \mathbb{E}[\log p(x|\theta)]}_{\xrightarrow[N\to\infty]{} 0 \ (LLN)}$$

$$\overset{N\to\infty}{=} \arg\max_\theta \frac{1}{N}\sum_{n=1}^N \log p(x_n|\theta)$$

$$= \arg\max_\theta \frac{1}{N}\sum_{n=1}^N p(x_n|\theta)$$

$$= \theta_{MLE}$$

c) $L(\theta) = -\frac{1}{N}\sum_{n=1}^N p(x_n|\theta)$