Quiz-1 solutions

1) **Definition of K-L Divergence:**

→ Suppose there is a random variable 'x' for which we have 2 probability distributions P_x & q_x. Then the K-L Divergence is $D(P_x \| q_x) \triangleq -\sum_{x \in supp(P_x)} P(x) \log \frac{q(x)}{P(x)}$

→ A random variable can have only one true distribution, so K-L divergence or relative entropy talks about the deviations b/w the 2 distributions.

→ It is similar to the distance metric but fails to satisfy

a) $D(P \| q) = D(q \| P)$ (condition)

b) $D(P \| q) \leq D(P \| r) + D(r \| q)$ (Δle inequality condition)

$H(x) + H(y) \geq H(x, y)$:

$H(x) + H(y) - H(x, y)$

$$= -\sum_{x \in supp(P_x)} P(x) \log P(x) + -\sum_{y \in supp(P_y)} P(y) \log P(y) - \left[-\sum_{x,y \in supp(P_{x,y})} P(x,y) \log P(x,y) \right]$$

$$= -\sum_{\substack{\cancel{x \in supp(P_x)} \\ x,\, y \in supp(P_{x,y})}} P(x,y) \log P(x) - \sum_{x,y \in supp(P_{x,y})} P(x,y) \log P(y)$$

$$+ \sum_{x,y \in supp(P_{x,y})} P(x,y) \log P(x,y)$$

$$= \sum_{x,y \in supp(P_{x,y})} P(x,y) \log \frac{P(x,y)}{P(x) \cdot P(y)} = D\big(P(x,y) \,\|\, P(x) \cdot P(y) \big)$$

We know $D(P \| q) \geq 0$ for all prob distributions P, q.

$\therefore H(x) + H(y) \geq H(x,y)$

<u>If case:</u>

If X & Y are independent, then

$P(x,y) = P(x) \cdot P(y)$

So $D(P(x,y) \,||\, P(x) \cdot P(y)) = \sum_{\substack{x,y \in \\ supp(P_{x,y})}} P(x,y) \log \cancel{\frac{P(x) \cdot P(y)}{P(x) \cdot P(y)}}^{\,1}$

$= 0 \quad (\log(1) = 0)$

$\therefore H(x) + H(y) = H(x,y)$

<u>Only if case:</u>

Let $D(P(x,y) \,||\, P(x) \cdot P(y)) = 0$.

When is $D(p||q) = 0$?

Applying Jensens equality condition where

$\frac{p(x)}{q(x)} = \text{const } 'c' \quad \forall x \in supp(P_x)$

$1 = \sum_{x \in supp(P_x)} p(x) = \sum_{x \in supp(P_x)} c \cdot q(x) \quad \forall x \in supp(P_x)$

together will mean that $c = 1 \Rightarrow p(x) = q(x)$

$\Rightarrow \frac{p(x)}{q(x)} = 1$

So $\frac{P(x,y)}{P(x) \cdot P(y)} = 1 \Rightarrow P(x,y) = P(x) \cdot P(y)$

X & Y are independent R.V

2) X is a r.v taking values from $\mathcal{X}$

$|\mathcal{X}| = n$

$\Rightarrow \mathcal{X} = \{x_1, x_2, \dots, x_n\}$

Given A, we need to show there exists 2 probability distributions P_1 & P_2 for X, such that $D(P_1 \| P_2) = A$.

$P_1(x)$ is a probability distribution.

$$\Rightarrow \left.\begin{aligned} &\forall x_i \in \mathcal{X}, \quad P(X = x_i) \geq 0 \\ &\sum_{x_i} P(X = x_i) = 1 \end{aligned}\right\} \text{valid probability distribution conditions}$$

$$D(P_1 \| P_2) = A \Rightarrow A = -\sum_{x \in supp(P_1)} P_1(x) \log \frac{P_2(x)}{P_1(x)}$$

Let $P_1(x)$ be a prob distribution such that

$$P_1(X = x_j) = 1 \quad \cancel{\text{for any}}$$

$$\& \; P_1(X = X_i) = 0 \quad \forall i \in [n] - \{j\}$$

Then,

$$A = - \cancelto{1}{P_1(X = X_j)} \log \frac{P_2(X = X_j)}{\cancelto{1}{P_1(X = X_j)}}$$

$$= -\log P_2(X = X_j)$$

$$\therefore P_2(X = X_j) = 2^{-A}$$

We can have $\displaystyle \sum_{\substack{i \\ i \in [n] - \{j\}}} P_2(X = X_i) \leq 1 - 2^{-A}$ → one of the probability distribution condition is satisfied.

$$A = -\sum_{x \in supp(P_1)} P_1(x) \log \frac{P_2(x)}{P_1(x)}$$

$$= -H(x) - \sum_{x \in supp(P_1)} P_1(x) \log P_2(x)$$

$$-[A + H(x)] = \sum_{x \in supp(P_1)} P_1(x) \log P_2(x)$$

We know $A \geq 0$, $H(x) \geq 0$, $P_1(x) \geq 0$

$\therefore$ $\log P_2(x)$ $\forall x \in supp(P_1)$ must be -ve

$\Rightarrow P_2(x) \in [0, 1]$ → Another probability distribution condition is satisfied

$$\therefore \left.\begin{array}{l} P_2(x) \geq 0 \\ \& \sum_{x_i} P(X = x_i) = 1 \end{array}\right\} \forall x_i \in \mathcal{X}$$

$\therefore$ P_2 is a valid probability distribution.