Information-theoretic inequalities.

1. Lemma [Log sum inequalities].

Let P, Q be two discrete distributions. Then

$$D_{KL}(P, Q) \geqslant p \cdot \log \frac{p}{q}, \quad (1)$$

where $p = \|P\|_1$, $q = \|Q\|_1$.

Proof.

To prove Eq (1), we only need to show that

$$\sum_{i=1}^{n} P_i \cdot \log \frac{P_i}{q_i} \geqslant p \cdot \frac{\log p}{\log q},$$

which is equivalent to prove that

$$\sum_{i=1}^{n} P_i \cdot \log \frac{q \cdot P_i}{p \cdot q_i} \geqslant 0.$$

To see this, observe that $\log \frac{1}{x} \geqslant 1 - x$ for any $x > 0$, and let $\lambda = \frac{q}{p}$. Then

$$\begin{aligned}\sum_{i=1}^{n} P_i \cdot \log \frac{q \cdot P_i}{p \cdot q_i} &= \sum_{i=1}^{n} P_i \cdot \log \lambda \cdot \frac{P_i}{q_i} \geqslant \sum_{i=1}^{n} P_i \cdot \left(1 - \frac{1}{\lambda} \cdot \frac{q_i}{P_i}\right) = \sum_{i=1}^{n} P_i \cdot \left(1 - \frac{p}{q} \cdot \frac{q_i}{P_i}\right) \\ &= \sum_{i=1}^{n} P_i - \frac{p}{q} \cdot q_i \\ &= 0.\end{aligned}$$

$\blacksquare$

2. Information Processing Inequality.

Let X, Y be two random variables, which take values on $\mathcal{X}$. Let $f: \mathcal{X} \to f(\mathcal{X})$. Then

$$D_{KL}(P_{f(X)}, P_{f(Y)}) \leq D_{KL}(X, Y).$$

Proof. $$D_{KL}(X,Y) = \sum_{w \in \Sigma} P_X(w) \cdot \log \frac{P_X(w)}{P_Y(w)}$$

$$= \sum_{i \in f(\Sigma)} \sum_{w \in f^{-1}(i)} P_X(w) \cdot \log \frac{P_X(w)}{P_Y(w)}$$

$$\geq \sum_{i \in f(\Sigma)} P_{f(X)}(i) \cdot \log \frac{P_{f(X)}(i)}{P_{f(Y)}(i)}$$

$$= D_{KL}(f(X), f(Y)).$$

$\blacksquare$