loss function: calculate the difference between output and real value:

1st Difference between r.v.

Set $Y_i \geq 0$, $Y_i' \geq 0$, $\sum_{i=1}^{n} Y_i = 1$, $\sum_{i=1}^{n} Y_i' = 1$

let $u_i = \log Y_i$, $v_i = \log Y_i'$, calculate $u_i - v_i = \log \frac{Y_i}{Y_i'}$

$E_{Y_i}(\log \frac{Y_i}{Y_i'}) = \sum_{i=1}^{n} Y_i * \log \frac{Y_i}{Y_i'}$ KL divergence

Since $\sum_{i=1}^{n} Y_i \cdot \log \frac{Y_i}{Y_i'} = -\sum_{i=1}^{n} Y_i \log \frac{Y_i'}{Y_i}$

log concave ∩

$-\log \sum_{i=1}^{n} Y_i \frac{Y_i'}{Y_i} \geq 0$

① KL-Divergence ≥ 0, can be loss function

$KL\text{-}D = \sum_{i=1}^{n} Y_i \log \frac{Y_i}{Y_i'} = \sum_{i=1}^{n} Y_i \log Y_i - \sum_{i=1}^{n} Y_i \log Y_i'$

$= E_{Y_i}(\log Y_i) - E_{Y_i}(\log Y_i')$

$E_{Y_i}(\log Y_i) \geq -\log \sum_{i=1}^{n} Y_i \cdot \frac{1}{Y_i} = 0$ for $KL\text{-}D \geq 0$

$-\sum_{i=1}^{n} Y_i \log Y_i' \geq 0 \Rightarrow$ cross entropy loss function related to w closely.

Application: 1) match difference between 2 distribution

2) with entropy: $H = -\sum_{i=1}^{n} P(x_i) \log P(x_i) = -E_{P_i}(\log P_i)$

$D_{KL}(P \| q) = \sum_{i=1}^{n} P(x_i)[\log P(x_i) - \log q(x_i)]$

$= E_{P_i}[\log P(x_i)] - E_{P_i}[\log q(x_i)]$