

$$D_{KL}(P \parallel q) = E_P[\log P/q]$$

$$= \sum p(x) \log \frac{p(x)}{q(x)}$$

$$H(p, q) = H(p) + D_{KL}(P \parallel q)$$

$$-\sum p(x) \log q(x) = -\sum p(x) \log p(x) + \sum p(x) \log \frac{p(x)}{q(x)}$$

$$= \cancel{-\sum p(x) \log p(x)} + \cancel{\sum p(x) \log p(x)} - \sum p(x) \log q(x)$$