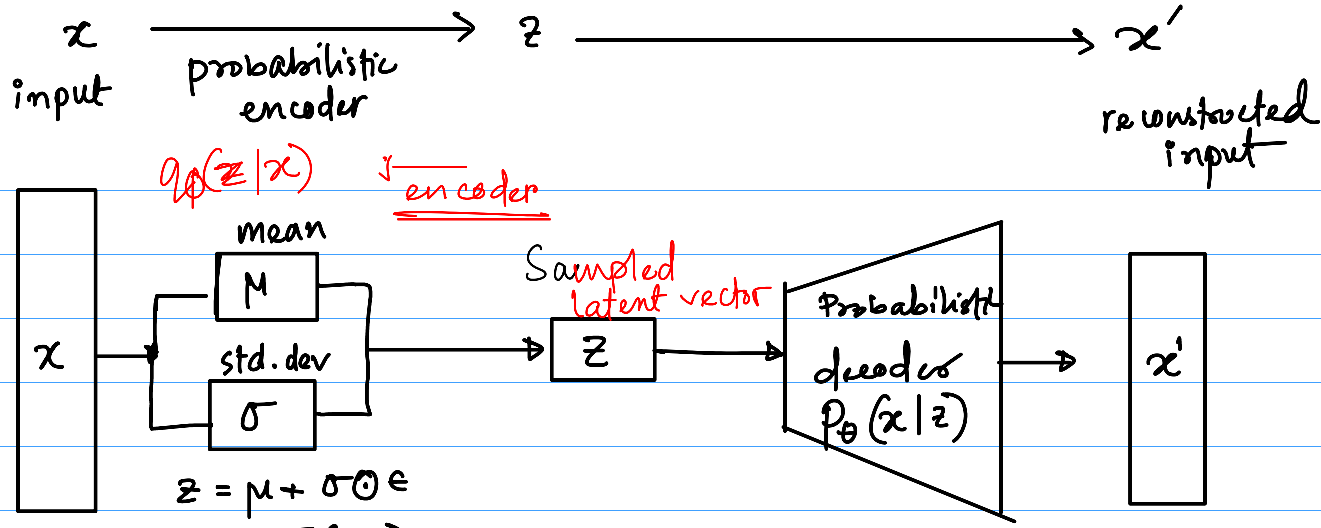


$$z = \mu + \sigma \odot \epsilon$$
$$\epsilon = \mathcal{N}(0, I)$$

Loss function : $L(\theta, \phi) =$

Probability Concepts Required:

$p(x)$: prob of a random var x

$p(x|y)$: prob of x given y has happened.

$$p(x|y) = \frac{p(y|x) \cdot P(x)}{P(y)}$$

$$= \frac{P(x|y)}{P(x)} P(y) \rightarrow \text{prior}$$

$\frac{P(x|y)}{P(x)} \rightarrow$ likelihood ratio

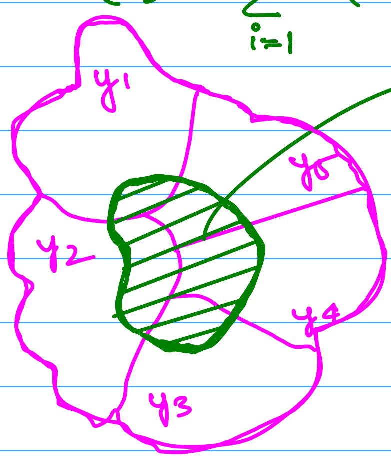
$$\text{posterior probability} = \frac{P(x,y)}{P(x)} \quad \text{joint prob dist}$$

∴ Total probability theorem: $y_1 \quad y_2 \quad y_3 \quad \cdots \quad y_n$ (no overlap
mutually exclusive $(y_i \cap y_j = 0)$

Event X, union of multiple mutually exclusive events:

$$P(X) = \sum_{i=1}^{N} P(x|y_i)\, P(y_i) \quad \text{(2)}$$

$$P(y|x) = \frac{P(x,y)}{P(x)} \quad \text{(1)}$$



$$P(x) = \sum_{i=1}^{5} P(x, y_i)$$

$$= \sum_{i=1}^{5} P(x|y_i)\, P(y_i) \quad \text{(2)}$$

$$p(y|x) = \frac{P(x|y)\, P(y)}{\boxed{\sum_{i=1}^{N} P(x|y)\, P(y_i)}} \qquad \boxed{\sum_{i=1}^{N} P(x|y)\, P(y_i)} \rightarrow P(x)$$

$$P(y|x) = \frac{P(x|y)\, P(y)}{P(x)}$$

$$= \frac{P(x|y)\, P(y)}{\sum_{i=1}^{N} P(x|y_i)\, P(y_i)}$$

Expectation of a Random Variable:

$$E(x) = \sum_{i=1}^{K} x_i \cdot P(x = x_i) = E(x)$$

③
$P(3) = \frac{1}{6}$
$P(3) = 1/6$

$p(x=3 \mid y \text{ is odd})$

$$P(x \mid y) = \frac{P(y \mid x)\, P(y)}{P(x)}$$

$=$

For fair die: $\{1, 2, 3, 4, 5, 6\}$ $P = 1/6$

$E(x) =$

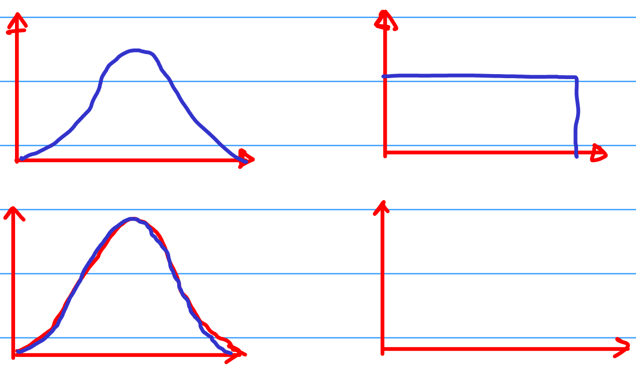
$$E(x) = \sum_{i=1}^{6} x_i .\, P(x_i)$$

$$= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6}$$

$$= \frac{21}{6} =$$

$$= \frac{21}{6} =$$

KL-Divergence: Kullback Divergence: 2020



high value

low value

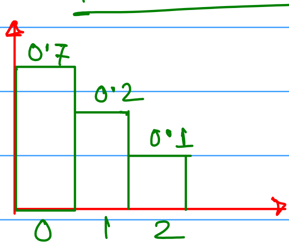
Measure of how one probability is different from the other: for two probability distribution P and Q

$$D_{KL} = (P \| Q) = \sum_{x} P(x_i = x) \log \frac{P(X = x)}{Q(X = x)}$$

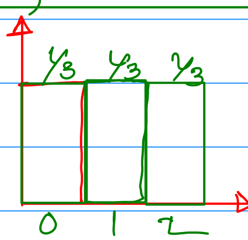
$$= \sum_{x} P(x) \log \frac{P(x)}{Q(x)} \qquad D_{KL}$$

Example:

P distribution



Q distribution:



uniform distribution

Here $D_{KL}(Q \| P) = \sum_{x=0}^{2} Q(x) \log \frac{Q(x)}{P(x)}$

$$= \frac{1}{3} \log\left(\frac{1/3}{0.7}\right) + \frac{1}{3} \log\left(\frac{1/3}{0.2}\right) + \frac{1}{3} \log\left(\frac{1/3}{0.1}\right)$$

$= 0.09673$ nats.

Properties: ① $KL(P \| Q)$ or $KL(Q \| P) \geq 0$
② $KL(P \| Q) \neq KL(Q \| P)$ Non symmetric

Consider

$$P(x) = \mathcal{N}(x; \mu_1, \Sigma_1)$$
$$Q(x) = \mathcal{N}(x; \mu_2, \Sigma_2)$$

Σ_1 and Σ_2 covariance matrix

$$\mathcal{N}(x, \mu, \sigma) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left(\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

Equation of normal distribution :

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

then

$$D_{KL} = (P(x) \| Q(x)) = \frac{1}{2}\left[\log \frac{|\Sigma_2|}{|\Sigma_1|} - d + tr(\Sigma_2^{-1}\Sigma_1) + (\mu_2 - \mu_1)\Sigma_2^{-1}(\mu_2 - \mu_1)\right]$$

$\boxed{\text{Proof:}}$

$$KL(P(x) \| Q(x)) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$
$$= \sum_x P(x) \cdot \left[\log P(x) - \log Q(x)\right]$$