

Review

- The KL distance from p_X to p_Y

$$D(p_X \| p_Y) \triangleq E \left\{ \log \frac{p_X(X)}{p_Y(X)} \right\}$$

$$= E \left\{ \log \frac{1}{p_Y(X)} \right\}$$

← mismatched,
not minimum...

$$- E \left\{ \log \frac{1}{p_X(X)} \right\}$$

← minimum
description
length

- "Shannon Code": a variable-length source code,
a PF code,
$$l_i = \lceil \log_2 \frac{1}{p_i} \rceil$$

- $D(p_X \| p_Y) \geq 0$ where the equality holds iff $p_X = p_Y$.

$$\begin{aligned} D(p_{X,Y} \| p_{Z,W}) &= E \left\{ \log \frac{p_{X,Y}(X,Y)}{p_{Z,W}(X,Y)} \right\} \\ &= D(p_X \| p_Z) + D(p_{Y|X} \| p_{W|Z}) \\ &= E \left\{ \log \frac{p_{X,Y}(X,Y)}{p_{W|Z}(Y|X)} \right\} \end{aligned}$$

$$\bullet \quad I(X; Y) = E \left[\log \frac{p_{X,Y}(X,Y)}{p_X(X) p_Y(Y)} \right]$$

$$= D(p_{X,Y} \parallel p_X p_Y)$$

≥ 0 where the equality holds iff X & Y are independent.

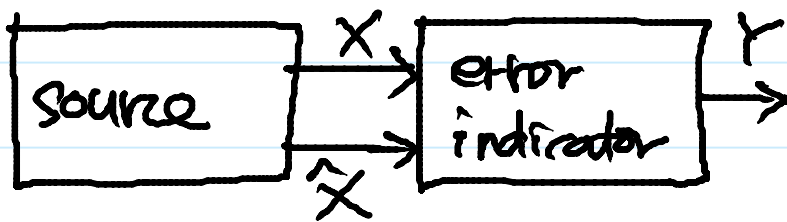
$$\bullet \quad I(X; Y | Z) \geq 0 \text{ where the equality holds if } X \text{ & } Y \text{ are conditionally independent given } Z$$

$$\left(p_{X,Y|Z}(x,y|z) = p_{X|Z}(x|z) \cdot p_{Y|Z}(y|z) \right)$$

$$\forall x \forall y \forall z$$

- Fano's inequality
Given two random variables X & $\hat{X}$
w/ the same sample space $\mathcal{X}$, the Fano's inequality mts
 $P_e \triangleq \Pr(X \neq \hat{X})$ & $H(X|\hat{X})$
in such a way that

$$H(X|\hat{X}) \leq h_b(P_e) + P_e \cdot \log(|\mathcal{X}| - 1) \\ \leq 1 + P_e \log |\mathcal{X}|$$



$$\begin{aligned}
 H(X|\hat{X}) &= H(X|\hat{X}) + H(Y|X, \hat{X}) \\
 &= H(X, Y|\hat{X}) \\
 &= H(Y|\hat{X}) + H(X|Y, \hat{X}) \\
 &\leq h_b(P_e) + \\
 &\quad \sum_{\hat{x}} \sum_{y \neq \hat{x}} H(X|Y=y, \hat{X}=\hat{x}) P_Y(Y=y, \hat{X}=\hat{x}) \\
 &= h_b(P_e) + \sum_{\hat{x}} H(X|Y=1, \hat{X}=\hat{x}) P_Y(Y=1, \hat{X}=\hat{x})
 \end{aligned}$$

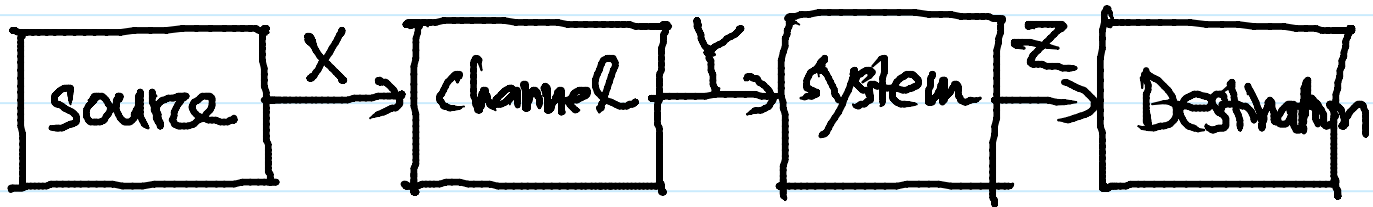
$$\begin{aligned}
 \sum_{\hat{x}} \sum_{y \neq \hat{x}} &= \sum_{\hat{x}} \sum_{y \neq \hat{x}} \\
 &= \sum_{\hat{x}} \sum_{y \neq \hat{x}}
 \end{aligned}$$

$$\begin{aligned}
&\leq h_b(P_e) + \sum_{\hat{x}} \log(|\mathcal{X}|-1) P_r(Y \neq 1, \hat{X} = \hat{x}) \\
&= h_b(P_e) + \boxed{\sum_{\hat{x}} P_r(Y \neq 1, \hat{X} = \hat{x})} \cdot \log(|\mathcal{X}|-1) \\
&\quad = P_r(Y \neq 1) \\
&= h_b(P_e) + P_e \cdot \log(|\mathcal{X}|-1)
\end{aligned}$$

• Implications

- 문제 (i) If $P_e \rightarrow 0$, then $H(X|\hat{X}) \rightarrow 0$
 ⇐ (ii) If $H(X|\hat{X}) \rightarrow 0$, then $P_e \rightarrow 0$
 ↳ contrapositive

Data Processing Theorem



~~...~~

$$I(X; Y) \geq I(X; Z)$$

Markov chain

Def. Given 3 jointly distributed r.v.'s X, Y , and Z , we say X, Y , and Z form a Markov chain if

$$p_{X,Y,Z}(x,y,z) = p_X(x) p_{Y|X}(y|x) p_{Z|Y}(z|y)$$

$\forall x, \forall y, \forall z.$

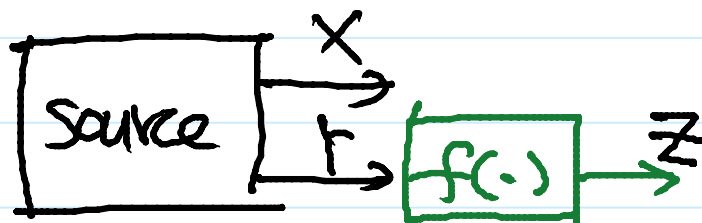
$\uparrow$
 $p_{Z|X,Y}(z|x,y)$

($\Leftrightarrow X$ & Z are conditionally independent given Y).

Notation $X \rightarrow Y \rightarrow Z$
 $\equiv Z \rightarrow Y \rightarrow X$

• Properties

(iii) Given two jointly distributed discrete r.v.'s X and Y , let $Z = f(Y)$ w/ f a deterministic function.



$$X \rightarrow Y \rightarrow Z$$

$$(iv) \quad X \rightarrow Y \rightarrow Z \Rightarrow I(X; Y) \geq I(X; Z)$$

$$\Downarrow \Rightarrow I(Z; Y) \geq I(Z; X)$$

$$I(X; Z | Y) = 0$$

Proof)

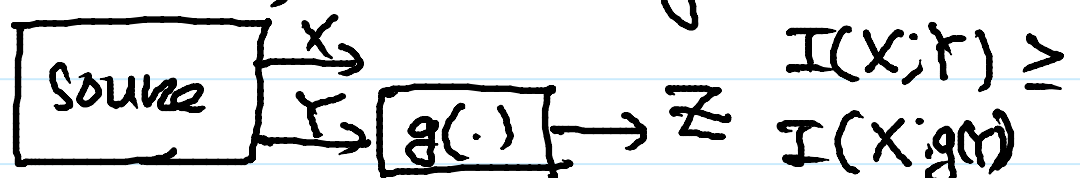
$$I(X; Y, Z) = I(X; Y) + I(X; Z | Y)$$

$$= I(X; Z) + \underbrace{I(X; Y | Z)}$$

$$5 = x + y \quad y \geq 0 \quad \rightarrow 0$$

$$5 \geq x$$

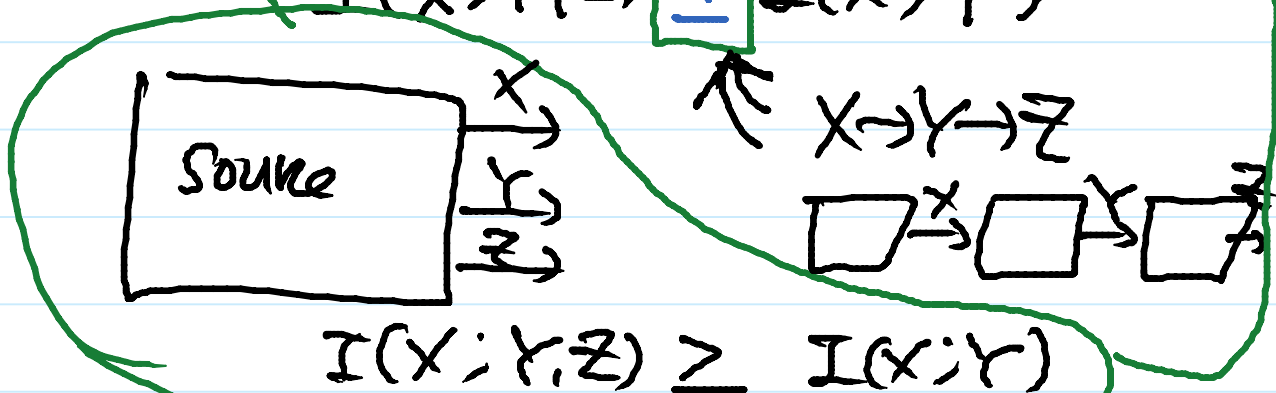
(v) Given X & Y , let $Z = g(Y)$. Then,



$$I(X; Y) \geq I(X; g(Y))$$

(vi) Q. Given X, Y, Z , what is the correct inequality relation b/w

$$I(X; Y|Z) \leq I(X; Y)$$



(proof)

$$\begin{aligned} I(X; Z) + I(X; Y|Z) &= I(X; Y, Z) \\ &\geq 0 \quad = I(X; Y) \end{aligned}$$

$$I(X; Y|Z) \leq I(X; Y)$$

where equality holds iff X & Z are indep

Data compression

- zero-error data compression
- fixed-length source coding
- Shannon's information metrics
- channel coding theorem
 - discrete channel
 - continuous channel (Gaussian)
 - continuous-time channel (")
- rate distortion theory

data
TF

