

data $\begin{cases} \text{compression} \\ \text{transmission} \end{cases}$

Review

• The source coding theorem

a sequence of fixed-length codes $\mathcal{C}^{(k)}$ ^{source}
 s.t. code rate $R^{(k)} \rightarrow R$ [code symbols / source symbol]
 error rate $P_e^{(k)} \rightarrow 0$,
 as $k \rightarrow \infty$

an achievable rate

• Direct Part of source coding theorem

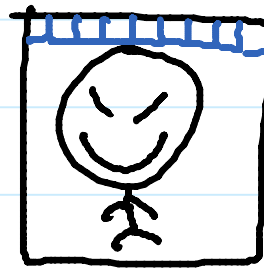
If $R \geq H(X)$, then R is achievable.

• Converse of source coding theorem

If R is achievable, then $R \geq H(X)$.

$\Leftrightarrow$ If $R < H(X)$, then R is not achievable.

$\begin{cases} \text{strong converse} \\ \text{weak converse} \end{cases}$



• Direct Part.

variable-length source coding

1st lesson : stack the outputs of the DMS and compress.

2nd lesson : consider the limit

fixed-length source coding

WLLN: $\frac{1}{n} \left(\sum_{i=1}^n -\log p_X(X_i) \right) \rightarrow H(X)$ in prob.

$$A_\varepsilon^{(n)} \triangleq \left\{ (x_1, \dots, x_n) : \left| -\frac{1}{n} \log p(x_1, x_2, \dots, x_n) - H(X) \right| < \varepsilon \right\}$$

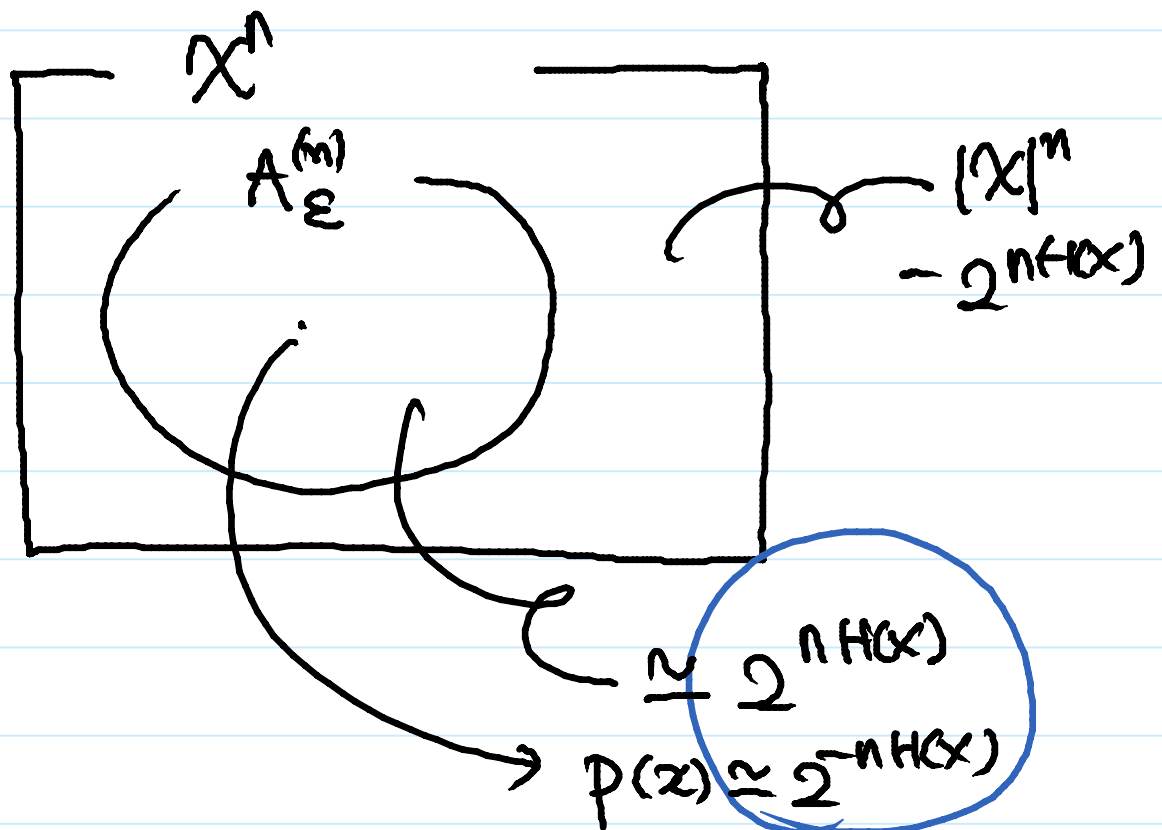
$P(A_\varepsilon^{(n)}) > 1 - \varepsilon$ for sufficiently large n

$$|A_\varepsilon^{(n)}| > (1 - \varepsilon) 2^{n(H(X) - \varepsilon)} \quad \text{f.s.l. } n$$

$$|A_\varepsilon^{(n)}| < 2^{n(H(X) + \varepsilon)}$$

$$2^{-n(H(X) + \varepsilon)} \leq p(x_1, x_2, \dots, x_n) \leq 2^{-n(H(X) - \varepsilon)}$$

AEP (asymptotic equipartition property)



$$\begin{aligned}
 \bullet \quad P_C &\equiv \Pr(X^{(n)} \in C^{(n)}) \\
 &= \sum_{x^{(n)} \in C^{(n)}} \Pr(X^{(n)} = \underline{x}^{(n)})
 \end{aligned}$$



(Metrics)

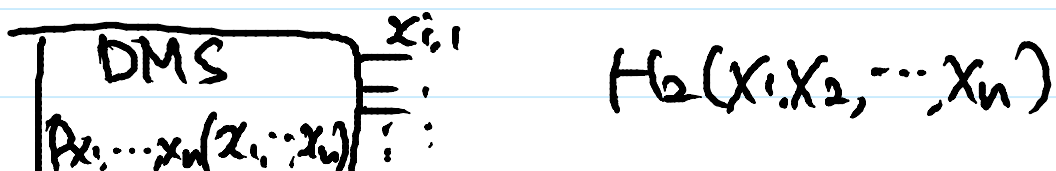
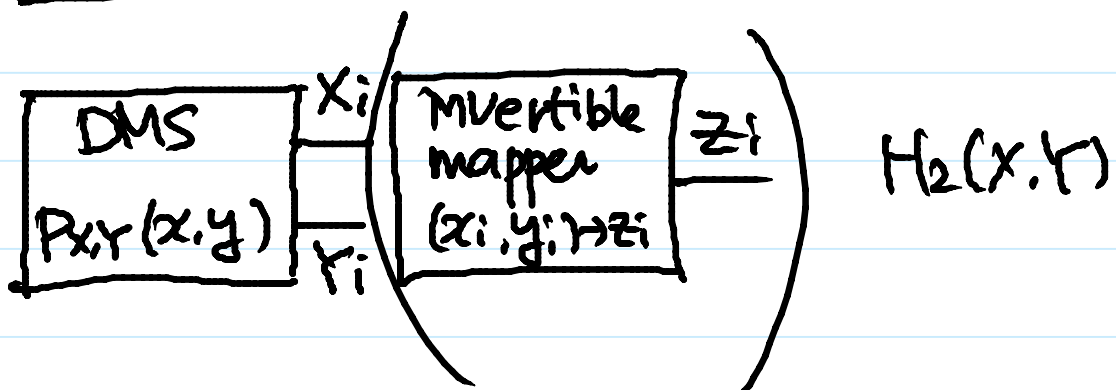
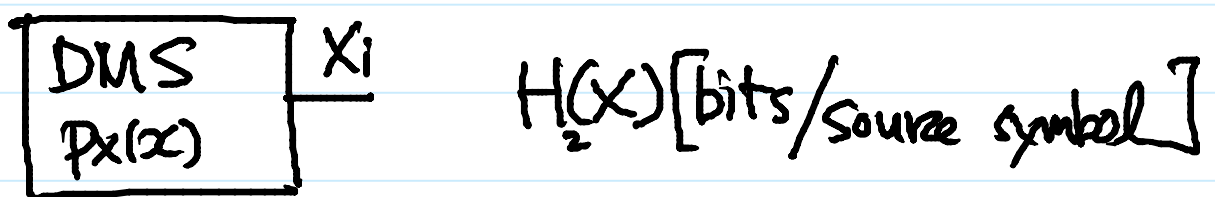
Shannon's Information Measures

$$\begin{aligned}
 H(X) &= -\sum (\log p_i) p_i = E\left\{\log \frac{1}{p_X(x)}\right\} \\
 H(X, Y) &= -\sum_x \sum_y (\log p_{xy}) p_{xy} = E\left\{\log \frac{1}{p_{X,Y}(x,y)}\right\} \\
 H(X_1, X_2, \dots, X_n) &= E\left\{\log \frac{1}{p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)}\right\}
 \end{aligned}$$

x	$P_X(X=x)$
1	p_1
2	p_2
$\vdots$	$\vdots$
$ X $	$p_{ X }$

$x \backslash y$	y_1	y_2
x_1	p_{11}	p_{12}
x_2	p_{21}	p_{22}

Physical Meaning of Entropy



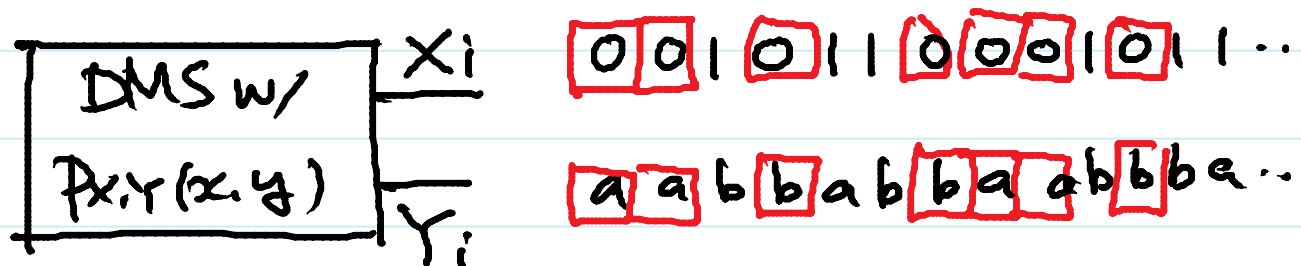
$$\boxed{p_{x_1, \dots, x_n}(x_1, \dots, x_n)} \begin{matrix} \vdots \\ \vdots \\ \vdots \end{matrix} \\ \hline x_{1:n}$$

$$1 \leq (x_1, x_2, \dots, x_n)$$

• Properties of Entropy

- $H(X) \geq 0$ equality holds iff X is deterministic.
- $H(X) \leq \log |X|$ equality holds iff X is uniform.
- $H_a(X) < H_b(X)$ if $a > b$
- $H_a(X) = H_b(X) \frac{\log b}{\log a}$
- $Y = f(X)$ where f is one-to-one correspondence
 $\Rightarrow H(X) = H(Y)$

• Conditional Entropy



Q. How many bits/ Y_i symbol when $X_i = 0$?

A ① $H(Y)$ ② $\sum_{y \in \mathcal{Y}} P_{Y|X}(y|x=0)$

$$x \log 1/p_{YX}(y|x=0)$$

Entropy of Y given $\{X=0\}$
 $H(Y|X=0) = H(Y|X=0)$
;