



# CHƯƠNG 3. CƠ SỞ LÝ THUYẾT THÔNG TIN THỐNG KÊ

# Giới thiệu

- ▶ LTTT giải quyết ba khái niệm cơ bản:
- ▶ (a) đo lường thông tin nguồn (tốc độ mà tại đó nguồn tạo ra thông tin)
- ▶ (b) dung lượng của kênh (tốc độ tối đa mà tại đó có thể truyền thông tin tin cậy qua kênh với xác suất lỗi nhỏ tùy ý)
- ▶ (c) mã hóa (một hệ thống cho phép sử dụng một cách hiệu quả dung lượng kênh để truyền thông tin)
- ▶ Ba khái niệm này có quan hệ chặt chẽ với nhau thông qua một loạt các định lý, các định lý này tạo nên cơ sở cho lý thuyết thông tin và được tổng kết như sau:
- ▶ *Nếu tốc độ thông tin từ một nguồn tạo bản tin không vượt quá dung lượng của kênh truyền tin thì sẽ luôn tồn tại một kỹ thuật mã hóa sao cho thông tin được truyền qua kênh với xác suất sai lỗi nhỏ tùy ý, kể cả khi có mặt của nhiễu không mong muốn.*

# LƯỢNG TIN RIÊNG

# Đo lường thông tin

- ▶ Trong hệ thống truyền tin, thông tin là một “thứ gì đó” được truyền từ nguồn tới đích - thứ không sẵn có ở phía thu.
- ▶ Nguồn thông tin (số hoặc tương tự) tạo ra các bản tin (hay sự kiện), đầu ra của nó được lựa chọn ngẫu nhiên theo phân phối xác suất.
- ▶ Thông tin đo sự “lựa chọn ngẫu nhiên” trong việc lựa chọn một bản tin trong số nhiều bản tin và phía thu không có sự chắc chắn về việc bản tin nào được lựa chọn.
- ▶ Bản tin có xác suất xảy ra thấp luôn truyền tải nhiều thông tin hơn so với bản tin có xác suất xảy ra cao.

# Đo lường thông tin

- ▶ Ví dụ: “If a dog bites a man, it’s no news, but if a man bites a dog, it’s news”
- ▶ Bản tin có xác suất xảy ra càng thấp, nó càng chứa nhiều thông tin và ngược lại.
- ▶ Nếu một nguồn phát bản tin  $m$  với xác suất  $p$ , lượng tin được xác định một cách định lượng bởi Shannon như sau:

$$\triangleright I = \log_b \left( \frac{1}{p} \right) = -\log_b(p)$$

- ▶  $I$  gọi là lượng tin riêng (hay độ bất định) chứa trong bản tin  $m$ .
- ▶ **Logarithm cơ số nào?**
  - Nếu chọn cơ số  $e$ , đơn vị của độ bất định là nat
  - Nếu chọn cơ số 10, đơn vị của độ bất định là Hartley
  - Nếu chọn cơ số 2, đơn vị của độ bất định là bit.

# Đo lường thông tin

- ▶ Đối với nguồn rời rạc không nhớ (DMS), mỗi bản tin được phát độc lập với các bản tin trước đó.
- ▶ Khi nguồn gửi một số các bản tin độc lập  $m_1, m_2, \dots, m_M$  với xác suất tương ứng là  $p_1, p_2, \dots, p_M$  và nếu các bản tin nhận được là đúng thì lượng thông tin trong bản tin kết hợp ở phía thu là:
- ▶ 
$$I = \log_2 \left( \frac{1}{p_1} \right) + \log_2 \left( \frac{1}{p_2} \right) + \dots + \log_2 \left( \frac{1}{p_M} \right) = \sum_{j=1}^M \log_2 \left( \frac{1}{p_j} \right)$$

# ENTROPY

# Entropy

- ▶ Thông tin trung bình trong mỗi bản tin đơn lẻ được phát bởi nguồn được gọi là entropy của nguồn.
- ▶ Xét  $M$  bản tin khác nhau  $m_1, m_2, \dots, m_M$  với các xác suất tương ứng là  $p_1, p_2, \dots, p_M$ .
- ▶ Thông tin trung bình trong mỗi bản tin hay entropy là:
  - ▶  $H = \sum_{i=1}^M p_i \cdot I(m_i) = -\sum_{i=1}^M p_i \cdot \log p_i$  (bit/bản tin)
- ▶ Tính chất:
- ▶  $H_{min} = 0 \leftrightarrow$  chỉ có một bản tin hay  $p_i=1$  và  $p_j=0$  với mọi  $i \neq j$
- ▶  $H_{max} = \log M \leftrightarrow p_1 = p_2 = \dots = p_M = \frac{1}{M}$



## Ví dụ 2.1

- ▶ Chứng minh biểu thức:  $I(x_i x_j) = I(x_i) + I(x_j)$  nếu  $x_i, x_j$  độc lập với nhau.

# Giải

- ▶ Vì  $x_i, x_j$  độc lập nên  $P(x_i x_j) = P(x_i)P(x_j)$
- ▶ Ta có:  $I(x_i x_j) = \log \frac{1}{P(x_i x_j)}$
- ▶  $= \log \frac{1}{P(x_i)P(x_j)}$
- ▶  $= \log \frac{1}{P(x_i)} + \log \frac{1}{P(x_j)}$
- ▶  $= I(x_i) + I(x_j)$

## Ví dụ 2.2

- ▶ Nguồn không nhớ rời rạc  $X$  tạo ra bốn ký tự  $x_1, x_2, x_3, x_4$  với xác suất tương ứng là 0,5; 0,2; 0,2 và 0,1. Tính  $H(X)$  và lượng thông tin chứa trong bản tin  $x_4x_3x_1x_1$ .
- ▶ Nguồn không nhớ nhị phân phát hai bản tin với xác suất lần lượt là  $p$  và  $(1-p)$ . Chứng minh rằng entropy đạt giá trị cực đại khi hai bản tin có xác suất bằng nhau.

# Giải

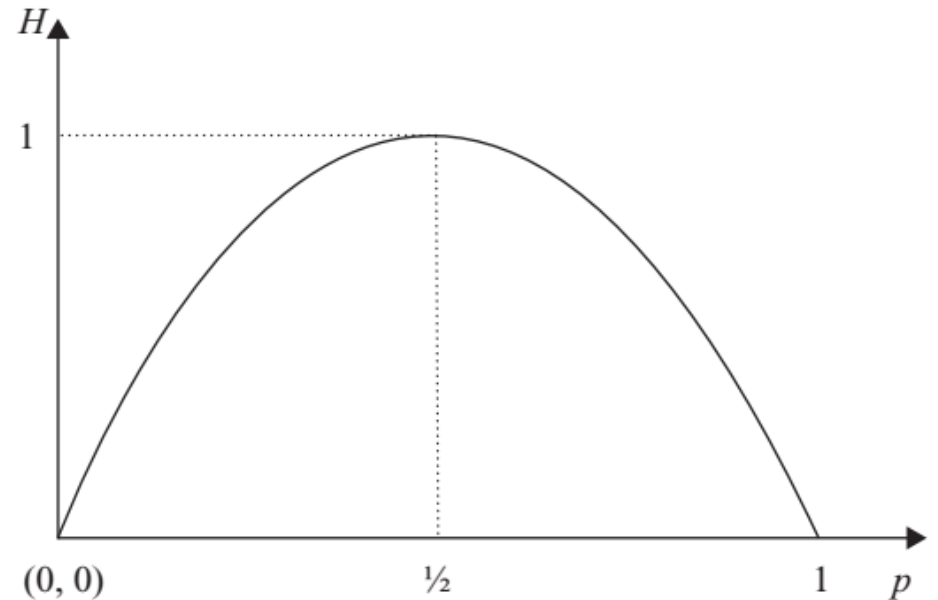
- ▷  $H(X) = -\sum_{i=1}^4 P(x_i) \log P(x_i)$
- ▷  $= -0,5 \cdot \log(0,5) - 0,2 \cdot \log(0,2) - 0,2 \cdot \log(0,2) - 0,1 \cdot \log(0,1)$
- ▷  $= 1,76 \text{ bit/ký tự}$
- ▷  $P(x_4 x_3 x_1 x_1) = (0,1)(0,2)(0,5)^2 = 5 \cdot 10^{-3}$
- ▷ Vì vậy:  $I(x_4 x_3 x_1 x_1) = -\log(5 \cdot 10^{-3}) = 7,64 \text{ bit}$
- ▷ Lượng thông tin trung bình chứa trong 1 chuỗi gồm 4 ký tự:
- ▷  $1,76 \times 4 = 7,04 \text{ (bit)}$

## Ví dụ 2.3

- ▶ Nguồn không nhớ nhị phân phát hai bản tin với xác suất lần lượt là  $p$  và  $(1-p)$ . Chứng minh rằng entropy đạt giá trị cực đại khi hai bản tin có xác suất bằng nhau.

# Giải

- ▶  $H = p \log \left( \frac{1}{p} \right) + (1 - p) \log \left( \frac{1}{1-p} \right)$
- ▶  $H' = \frac{1}{\ln 2} \left[ p \cdot p \cdot \left( -\frac{1}{p^2} \right) + \ln \left( \frac{1}{p} \right) + (1 - p)(1 - p) \cdot \left( \frac{1}{(1-p)^2} \right) - \ln \left( \frac{1}{1-p} \right) \right]$
- ▶  $= \frac{1}{\ln 2} \left[ \ln \left( \frac{1}{p} \right) - \ln \left( \frac{1}{1-p} \right) \right]$
- ▶  $H' = 0$  khi  $p = 1 - p$  hay  $p = \frac{1}{2}$
- ▶ Vậy entropy đạt cực đại khi hai bản tin có xác suất xảy ra như nhau.
- ▶ Khi đó:  $H_{max} = 2 \cdot \left( \frac{1}{2} \right) \log(2) = 1$  bit
- ▶  $H_{min} = 0$  khi  $p = 1$  hoặc  $p = 0$



# Tốc độ thông tin

- Nếu một nguồn tạo ra  $r$  bản tin/s, *tốc độ thông tin* hay lượng thông tin trung bình trong 1s được định nghĩa như sau:

$$R = \frac{H}{T} = rH \text{ (bit/s)}$$

ở đó  $T$  là thời gian cần để gửi một bản tin ( $T = 1/r$ )

Giả sử có 2 nguồn có cùng entropy  $H$ , phát  $r_1$  và  $r_2$  bản tin/s.

- Nguồn thứ nhất phát thông tin ở tốc độ  $R_1 = r_1 H$
- Nguồn thứ hai phát thông tin ở tốc độ  $R_2 = r_2 H$
- Nếu  $r_1 > r_2$ , nguồn thứ nhất sẽ phát nhiều thông tin hơn nguồn thứ hai mặc dù entropy là giống nhau.

*Vì vậy, một nguồn được đặc trưng bởi entropy và tốc độ thông tin của nó.*

## Ví dụ 2.5

- ▶ Một nguồn có 4 đầu ra với xác suất lần lượt là  $1/4$ ;  $1/2$ ;  $1/8$  và  $1/8$ .
- ▶ a. Tìm entropy của nguồn.
- ▶ b. Tìm tốc độ thông tin nếu có 8 đầu ra/s.
- ▶ Một tín hiệu tương tự có độ rộng băng thông  $B(\text{Hz})$  được lấy mẫu ở tốc độ Nyquist và các mẫu được lượng tử hóa thành 4 mức  $Q_1, Q_2, Q_3, Q_4$  với xác suất xảy ra tương ứng là  $p_1 = p_2 = 3/10$  và  $p_3 = p_4 = 2/10$ . Các mức lượng tử được giả định là độc lập với nhau. Tìm tốc độ thông tin của nguồn.



## Ví dụ 2.6

- ▶ Một tín hiệu tương tự có độ rộng băng thông  $B(\text{Hz})$  được lấy mẫu ở tốc độ Nyquist và các mẫu được lượng tử hóa thành 4 mức  $Q_1, Q_2, Q_3, Q_4$  với xác suất xảy ra tương ứng là  $p_1 = p_2 = 3/10$  và  $p_3 = p_4 = 2/10$ . Các mức lượng tử được giả định là độc lập với nhau. Tìm tốc độ thông tin của nguồn.

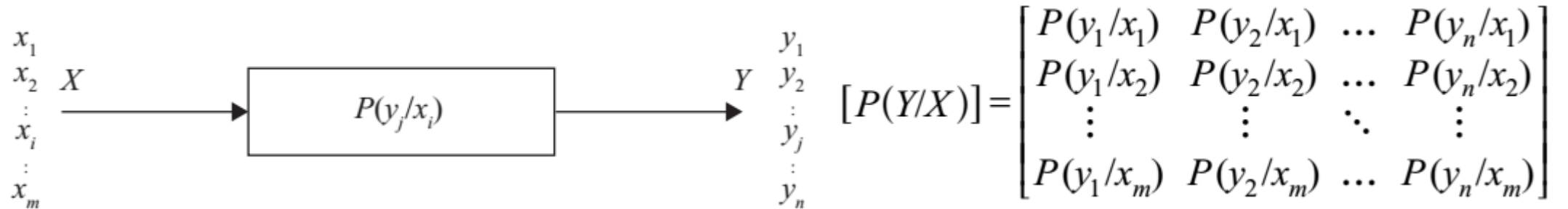
# MÔ HÌNH KÊNH

# Giới thiệu

- ▶ Kênh truyền tin là đường truyền hay phương tiện truyền mà thông qua đó các ký tự được truyền đến phía thu.
- ▶ Các loại kênh:
  - Kênh rời rạc không nhớ
  - Các kênh đặc biệt
  - Kênh không nhiễu
  - Kênh nhị phân đối xứng

# Kênh rời rạc không nhớ (DMC)

- ▶ Xét kênh DMC với đầu vào  $X$  và đầu ra  $Y$ .
  - **Rời rạc**: vì  $X, Y$  là các nguồn rời rạc.
  - **Không nhớ**: đầu ra của kênh chỉ phụ thuộc vào đầu vào tại thời điểm đó mà không phụ thuộc vào đầu vào tại các thời điểm trước đó.
  - $X = \{x_1, x_2, \dots, x_m\}; Y = \{y_1, y_2, \dots, y_n\}$
  - $P(y_j|x_i)$ : xác suất có điều kiện nhận được đầu ra  $y_j$  khi đầu vào là  $x_i$ .  
 $P(y_j|x_i)$  được gọi là **xác suất chuyển trên kênh**



- Ma trận  $[P(Y/X)]$  được gọi là **ma trận kênh** hay **ma trận chuyển đổi**.

# Kênh rời rạc không nhớ (DMC)

- ▷  $\sum_{j=1}^n P(y_j|x_i) = 1$  với mọi  $i$
- ▷  $[P(X)] = [P(x_1) \ P(x_2) \ \dots \ P(x_m)]$
- ▷  $[P(Y)] = [P(y_1) \ P(y_2) \ \dots \ P(y_n)]$
- ▷  $[P(Y)] = [P(X)] [P(Y|X)]$
- ▷ Nếu biểu diễn  $X$  dưới dạng ma trận đường chéo:
 
$$[P(X)]_d = \begin{bmatrix} P(x_1) & 0 & \dots & 0 \\ 0 & P(x_2) & \dots & 0 \\ 0 & 0 & \dots & P(x_m) \end{bmatrix}$$
- ▷  $[P(X, Y)] = [P(X)]_d [P(Y|X)]$
- ▷  $P(X, Y)$  là ma trận xác suất kết hợp và  $P(x_i, y_j)$  là xác suất kết hợp của việc phát  $x_i$  và nhận được  $y_j$ .

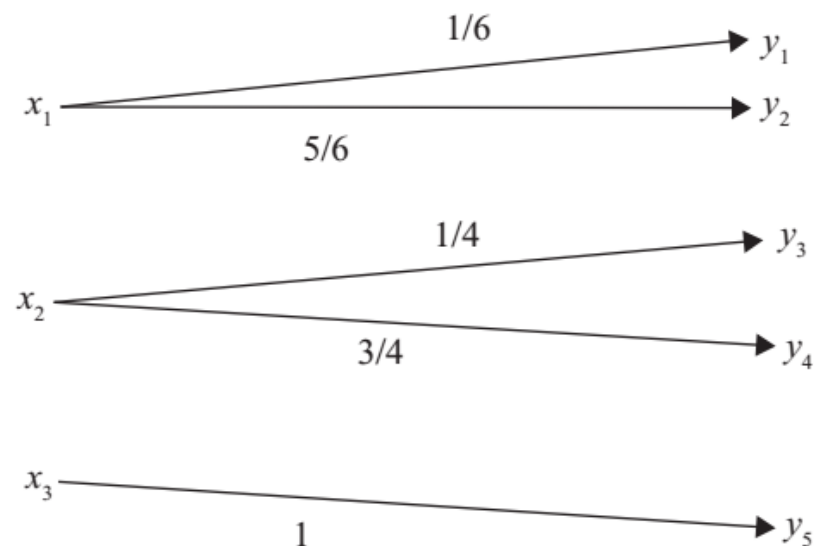
$$[P(Y|X)] = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) & \dots & P(y_n/x_1) \\ P(y_1/x_2) & P(y_2/x_2) & \dots & P(y_n/x_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(y_1/x_m) & P(y_2/x_m) & \dots & P(y_n/x_m) \end{bmatrix}$$

# Các kênh đặc biệt

- ▶ **Kênh không tổn thất:** là kênh chỉ chứa một thành phần khác 0 trong mỗi cột của ma trận kênh.
- ▶ Trong kênh này, không có thông tin nguồn nào bị mất trong quá trình truyền.

▶ **Ví dụ:**

$$[P(Y|X)] = \begin{bmatrix} 1/6 & 5/6 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 3/4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

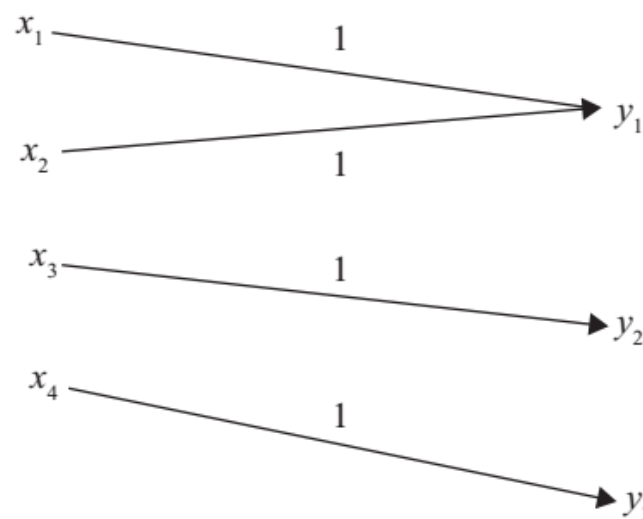


# Các kênh đặc biệt

- ▶ **Kênh xác định:** là kênh mà ma trận kênh chỉ chứa một thành phần khác 0 trong mỗi hàng.
- ▶ Bởi vì mỗi hàng chỉ có một thành phần khác 0 nên thành phần này phải bằng 1. Điều này nghĩa là khi một ký tự đầu vào được truyền qua kênh, ta luôn biết sẽ nhận được ký tự đầu ra nào.

▶ Ví dụ:

▶  $[P(Y|X)] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

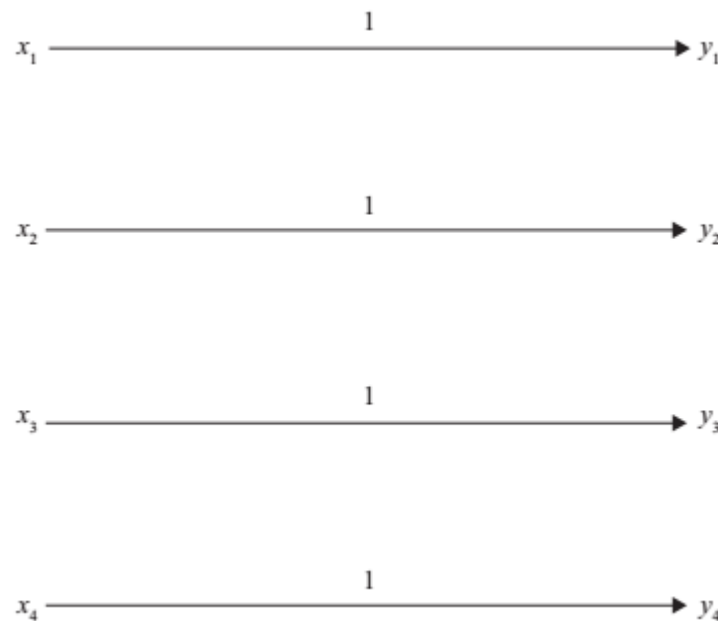


# Các kênh đặc biệt

- ▶ **Kênh không nhiễu:** là kênh thỏa mãn điều kiện là kênh không tổn hao và là kênh xác định.
- ▶ Ma trận kênh chỉ có một thành phần ở mỗi hàng và mỗi cột, và thành phần này phải bằng 1.

▶ Ví dụ:

$$[P(Y|X)] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

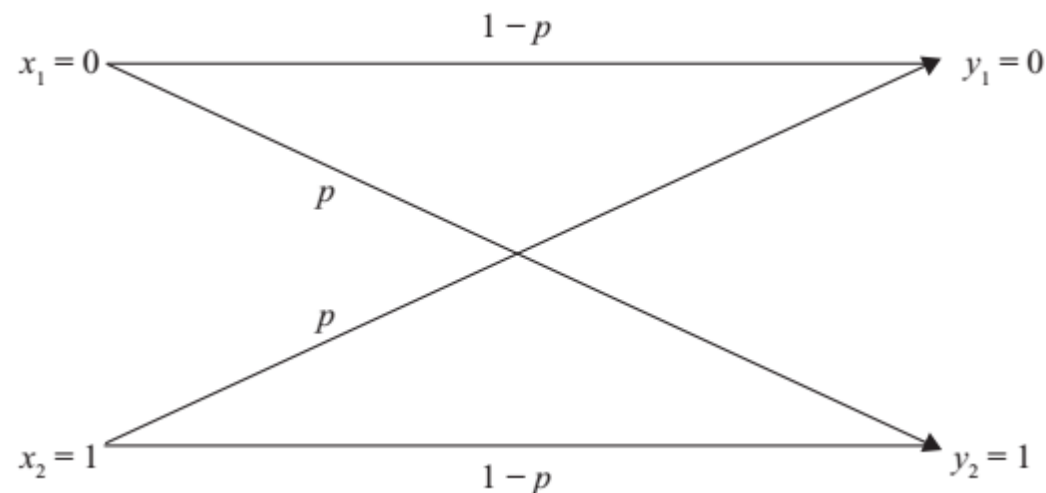




# Các kênh đặc biệt

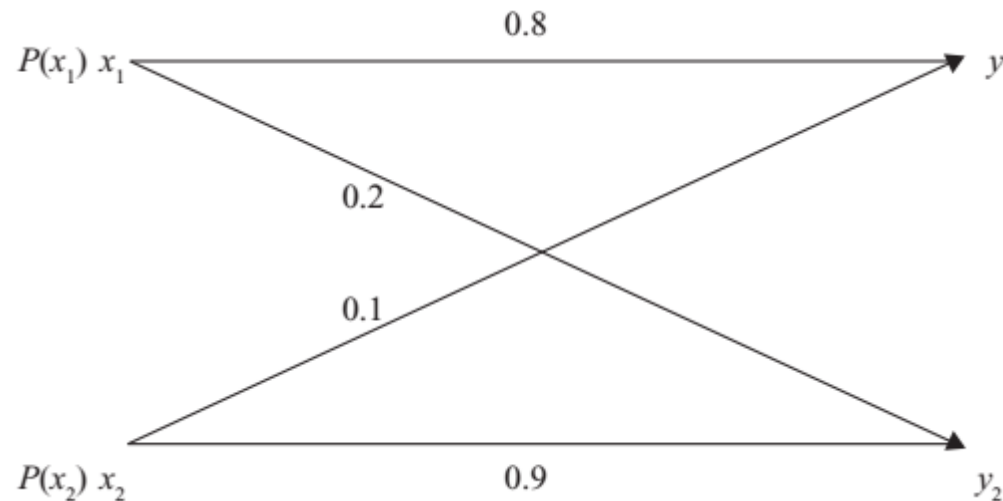
- ▶ **Kênh nhị phân đối xứng (BSC):** là kênh được mô tả bởi hình dưới đây.
- ▶ Kênh có hai đầu vào và hai đầu ra. Kênh được gọi là *đối xứng* vì xác suất chuyển sai từ 0 đầu vào thành 1 đầu ra bằng xác suất chuyển sai từ 1 đầu vào thành 0 đầu ra.

- ▶ 
$$[P(Y|X)] = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

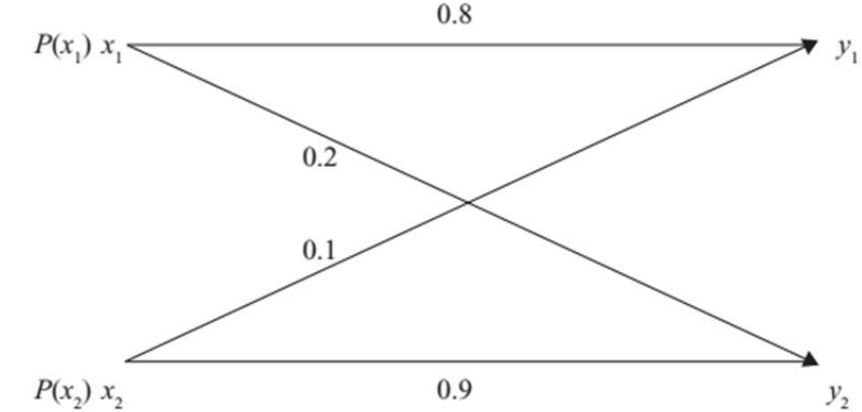


## Ví dụ 2.7

- ▶ Cho kênh nhị phân như mô tả trong hình dưới đây, hãy tìm:
- ▶ (a) Ma trận kênh
- ▶ (b)  $P(y_1)$  và  $P(y_2)$  khi  $P(x_1) = P(x_2) = 0,5$ .
- ▶ (c) Các xác suất kết hợp  $P(x_1, y_2)$  và  $P(x_2, y_1)$  khi  $P(x_1) = P(x_2) = 0,5$ .



# Giải



▷ a. Ma trận kênh là:

▷  $[P(Y|X)] = \begin{bmatrix} P(y_1|x_1) & P(y_2|x_1) \\ P(y_1|x_2) & P(y_2|x_2) \end{bmatrix} = \begin{bmatrix} 0,8 & 0,2 \\ 0,1 & 0,9 \end{bmatrix}$

▷ b.  $[P(Y)] = [P(X)] [P(Y|X)] = \begin{bmatrix} 0,5 & 0,5 \end{bmatrix} \begin{bmatrix} 0,8 & 0,2 \\ 0,1 & 0,9 \end{bmatrix}$

▷  $[P(Y)] = \begin{bmatrix} 0,45 & 0,55 \end{bmatrix} = [P(y_1) \ P(y_2)]$

▷ c.  $[P(X, Y)] = [P(X)]_d [P(Y|X)] = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,5 \end{bmatrix} \begin{bmatrix} 0,8 & 0,2 \\ 0,1 & 0,9 \end{bmatrix}$

▷  $[P(X, Y)] = \begin{bmatrix} 0,4 & 0,1 \\ 0,05 & 0,45 \end{bmatrix} = \begin{bmatrix} P(x_1, y_1) & P(x_1, y_2) \\ P(x_2, y_1) & P(x_2, y_2) \end{bmatrix}$

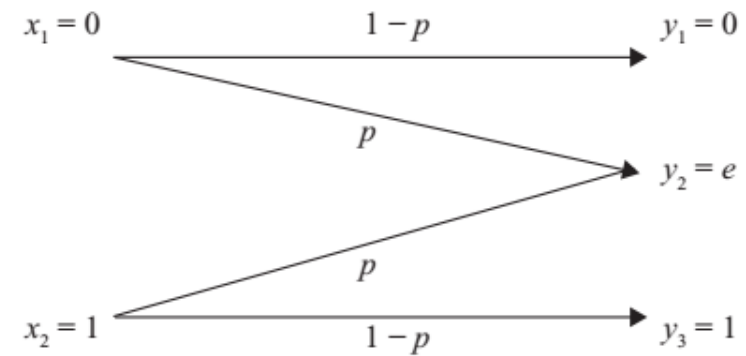
▷ Vậy  $P(x_1, y_2) = 0,1$  và  $P(x_2, y_1) = 0,05$

## Ví dụ 2.8

- ▶ Cho kênh được mô tả bởi ma trận kênh:
- ▶  $[P(Y|X)] = \begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$
- ▶ (a) Vẽ sơ đồ kênh tương ứng
- ▶ (b) Nếu nguồn tạo ra các đầu vào có xác suất như nhau, tìm các xác suất liên quan đến đầu ra của kênh với  $p=0,2$ .

# Giải

(a). Sơ đồ kênh tương ứng là:



- Kênh này gọi là kênh nhị phân có xóa.
- Nó có hai đầu vào 0 và 1 và ba đầu ra là 0, 1 và e ở đó e chỉ khái niệm “có xóa” (erasure), nghĩa là đầu ra có không chắc chắn và có thể bị xóa

(b).  $[P(Y)] = [P(X)] [P(Y|X)] = \begin{bmatrix} 0,5 & 0,5 \end{bmatrix} \begin{bmatrix} 0,8 & 0,2 & 0 \\ 0 & 0,2 & 0,8 \end{bmatrix} = \begin{bmatrix} 0,4 & 0,2 & 0,4 \end{bmatrix}$

Vì vậy  $P(y_1) = 0,4$  và  $P(y_2) = 0,2$ ;  $P(y_3) = 0,4$

# ENTROPY KẾT HỢP VÀ ENTROPY CÓ ĐIỀU KIỆN

# Tập xác suất hoàn chỉnh

- Xét một kênh có đầu vào và đầu ra tạo thành hai tập sự kiện  $[X] = [x_1, x_2, \dots, x_m]$  và  $[Y] = [y_1, y_2, \dots, y_n]$  với không gian mẫu  $S_1$  và  $S_2$ .
- Mỗi sự kiện  $x_i$  của  $S_1$  tạo nên một sự kiện  $y_j$  của  $S_2$ .
- Tập sự kiện hoàn chỉnh trong không gian mẫu tích  $S = (S_1 S_2)$  là:
- $[XY] = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_n \\ \dots & \dots & \dots & \dots \\ x_m y_1 & x_m y_2 & \dots & x_m y_n \end{bmatrix}$

# Tập xác suất hoàn chỉnh

- Tập xác suất được gọi là hoàn chỉnh nếu tổng các xác suất bằng 1.
- Ta có 3 sơ đồ xác suất hoàn chỉnh:
- $P(X) = [P(x_j)]; P(Y) = [P(y_k)]; P(X, Y) = [P(x_j, y_k)]$

<b>Entropy biên của X</b>	$H(X) = - \sum_{j=1}^m P(x_j) \cdot \log P(x_j)$	$P(x_j) = \sum_{k=1}^n P(x_j, y_k)$
<b>Entropy biên của Y</b>	$H(Y) = - \sum_{k=1}^n P(y_k) \cdot \log P(y_k)$	$P(y_k) = \sum_{j=1}^m P(x_j, y_k)$
<b>Entropy kết hợp của X và Y</b>	$H(X, Y) = - \sum_{j=1}^m \sum_{k=1}^n P(x_j, y_k) \cdot \log P(x_j, y_k)$	



# Entropy có điều kiện

- ▷  $H(X)$ : sự không chắc chắn trung bình của đầu vào kênh
- ▷  $H(Y)$ : sự không chắc chắn trung bình của đầu ra kênh
- ▷  $H(X,Y)$ : sự không chắc chắn trung bình của kênh truyền tổng thể
- ▷  $[X|y_k] = [x_1|y_k \quad x_2|y_k \quad \dots \quad x_m|y_k]$
- ▷ Khi đó:  $P(X|y_k) = [P(x_1|y_k) \quad P(x_2|y_k) \quad \dots \quad P(x_m|y_k)]$
- ▷ 
$$= \left[ \frac{P(x_1, y_k)}{P(y_k)} \quad \frac{P(x_2, y_k)}{P(y_k)} \quad \dots \quad \frac{P(x_m, y_k)}{P(y_k)} \right]$$
- ▷ Do  $P(x_1, y_k) + P(x_2, y_k) + \dots + P(x_m, y_k) = P(y_k)$
- ▷ Nên:  $\sum_{j=1}^m P(x_j|y_k) = 1$
- ▷ Vậy  $[X|y_k]$  được coi là một tập xác suất hoàn chỉnh.

# Entropy có điều kiện

$$H(X|y_k) = - \sum_{j=1}^m P(x_j|y_k) \cdot \log P(x_j|y_k)$$

- Lấy trung bình của entropy có điều kiện với tất cả các giá trị có thể có của  $y_k$ , ta có giá trị entropy có điều kiện trung bình của hệ thống:
- $H(X|Y) = \overline{H(X|y_k)} = \sum_{k=1}^n P(y_k) H(X|y_k)$
- $= - \sum_{k=1}^n P(y_k) \sum_{j=1}^m P(x_j|y_k) \cdot \log P(x_j|y_k)$

$$H(X|Y) = - \sum_{j=1}^m \sum_{k=1}^n P(x_j, y_k) \cdot \log P(x_j|y_k)$$

Tương tự:

$$H(Y|X) = - \sum_{j=1}^m \sum_{k=1}^n P(x_j, y_k) \cdot \log P(y_k|x_j)$$

# Ý nghĩa của entropy có điều kiện

- ▶  $H(X|Y)$  và  $H(Y|X)$  là entropy có điều kiện trung bình hay gọi đơn giản là entropy có điều kiện.
- ▶  $H(X|Y)$  đo sự không chắc chắn trung bình còn lại về đầu vào kênh sau khi đã quan sát đầu ra kênh.
- ▶  $H(Y|X)$  là sự không chắc chắn trung bình của đầu ra kênh khi  $X$  đã được truyền; nó cung cấp phép đo sai lỗi hoặc nhiễu.
- ▶ Vậy có 5 đại lượng entropy:  
 $H(X), H(Y), H(X,Y); H(X|Y)$  và  $H(Y|X)$

## Ví dụ 2.9

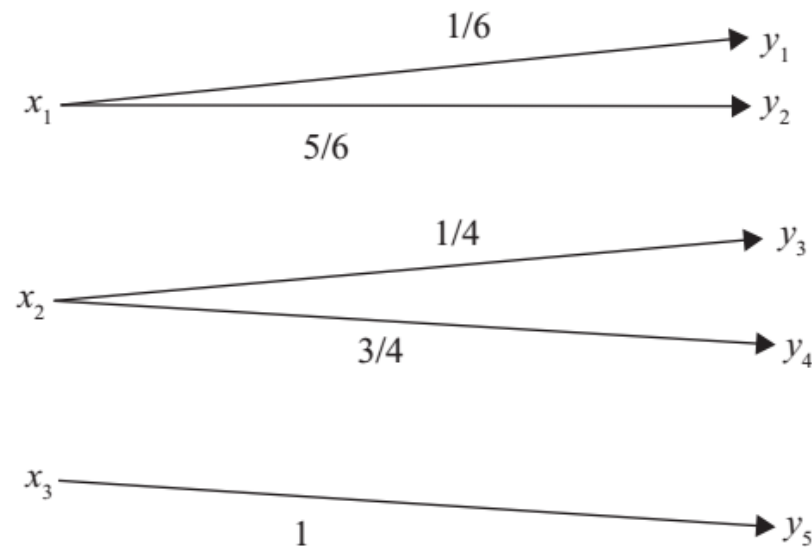
- ▷ Chứng minh rằng
- ▷ (a)  $H(X,Y)=H(X|Y) + H(Y)$
- ▷ (b)  $H(X,Y)=H(Y|X) + H(X)$

# Giải

- ▷ (a)  $H(X, Y) = - \sum_{j=1}^m \sum_{k=1}^n P(x_j, y_k) \cdot \log P(x_j, y_k)$
- ▷  $= - \sum_{j=1}^m \sum_{k=1}^n P(x_j, y_k) \cdot \log [P(x_j|y_k)P(y_k)]$
- ▷  $= - \sum_{j=1}^m \sum_{k=1}^n P(x_j, y_k) \cdot [\log P(x_j|y_k) + \log P(y_k)]$
- ▷  $= H(X|Y) - \sum_{k=1}^n \sum_{j=1}^m P(x_j, y_k) \cdot \log P(y_k)$
- ▷  $= H(X|Y) - \sum_{k=1}^n [\sum_{j=1}^m P(x_j, y_k)] \log P(y_k)$
- ▷  $= H(X|Y) - \sum_{k=1}^n [P(y_k)] \log P(y_k)$
- ▷  $= H(X|Y) + H(Y)$
- ▷ (b) Tương tự ta có:  $H(X, Y) = H(Y|X) + H(X)$

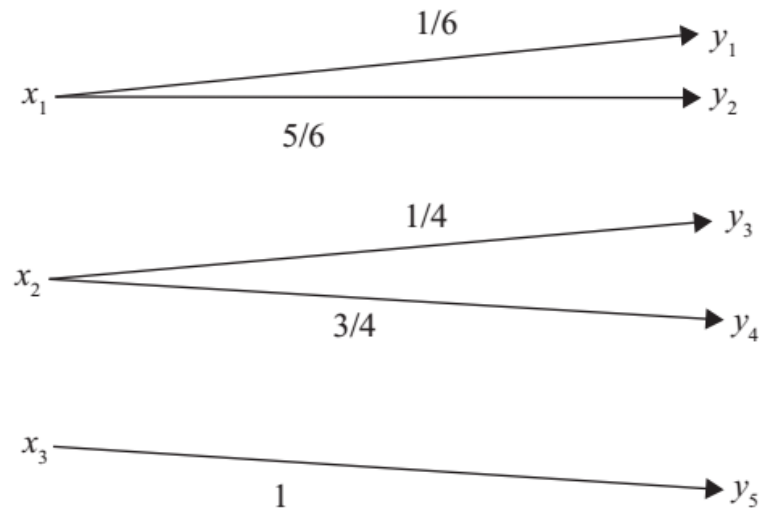
## Ví dụ 2.10

- ▷ Đối với kênh không tổn thất (ví dụ như hình dưới đây), chứng minh rằng  $H(X|Y)=0$



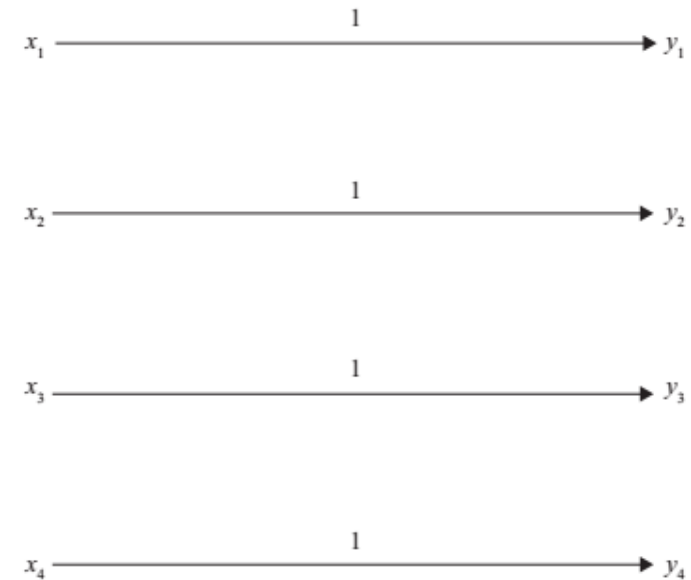
# Giải

- ▷ Đối với một kênh không tổn thất, khi nhận được đầu ra  $y_j$ , rõ ràng sẽ biết được  $x_i$  nào được truyền.
- ▷  $P(x_i|y_j) = 1$  hoặc 0
- ▷ Do đó:
- ▷  $H(X|Y) = -\sum_{j=1}^m \sum_{k=1}^n P(x_j, y_k) \cdot \log P(x_j|y_k) = 0$



## Ví dụ 2.11

- ▷ Đối với kênh không nhiễu (ví dụ như hình dưới đây) với  $m$  ký tự đầu vào và  $m$  ký tự đầu ra, chứng minh rằng:
- ▷ (a)  $H(X) = H(Y)$
- ▷ (b)  $H(Y|X)=0$





# Giải

▷ (a) Đối với kênh không nhiễu, xác suất chuyển được cho bởi:

$$▷ P(y_j|x_i) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$▷ \text{Vì vậy: } P(x_i, y_j) = P(y_j|x_i)P(x_i) = \begin{cases} P(x_i) & i = j \\ 0 & i \neq j \end{cases}$$

$$▷ \text{Và } P(y_j) = \sum_{i=1}^m P(x_i, y_j) = P(x_j)$$

$$▷ \text{Do đó: } H(Y) = -\sum_{j=1}^m P(y_j) \cdot \log P(y_j) = -\sum_{j=1}^m P(x_j) \cdot \log P(x_j) = H(X)$$

$$▷ \text{(b) } H(Y|X) = -\sum_{j=1}^m \sum_{k=1}^m P(x_j, y_k) \cdot \log P(y_k|x_j)$$

$$▷ = -\sum_{j=1}^m P(x_j) \sum_{k=1}^m \log P(y_k|x_j)$$

$$▷ = -\sum_{j=1}^m P(x_j) \log 1 = 0$$

# THÔNG TIN TƯỞNG HỒ

# Thông tin tương hỗ

$x_j$

$y_k$

**Xác suất tiên nghiệm  $P(x_j)$**

Độ bất định tiên nghiệm  $-\log P(x_j)$

**Xác suất hậu nghiệm  $P(x_j|y_k)$**

Độ bất định hậu nghiệm  $-\log P(x_j|y_k)$

Lượng thông tin  $I(x_j; y_k)$  thu được về  $x_j$  sau khi nhận  $y_k$  chính là mức giảm độ bất định:

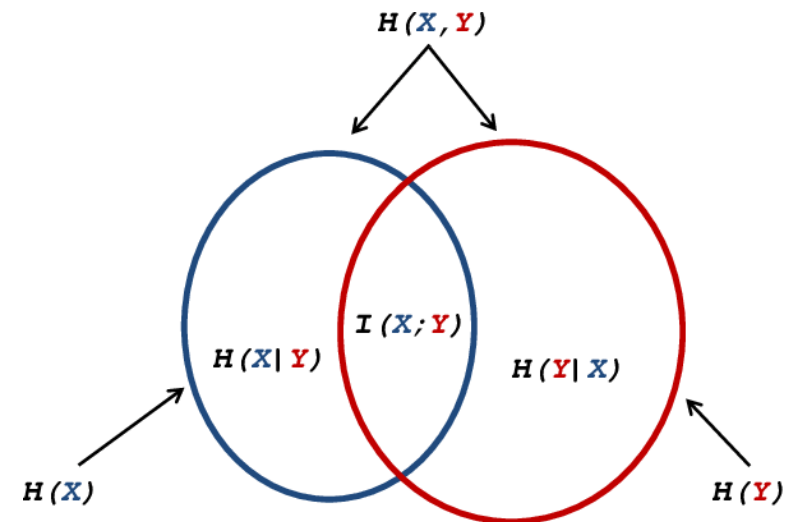
$$\begin{aligned} I(x_j; y_k) &= \text{Độ bất định tiên nghiệm} - \text{Độ bất định hậu nghiệm} \\ &= -\log P(x_j) - [-\log P(x_j|y_k)] \\ &= \log \frac{P(x_j|y_k)}{P(x_j)} = \log \frac{P(x_j|y_k) \cdot P(y_k)}{P(x_j)P(y_k)} = \log \frac{P(x_j, y_k)}{P(x_j)P(y_k)} \\ &= \log \frac{P(y_k|x_j)}{P(y_k)} = I(y_k; x_j) \end{aligned}$$

# Thông tin tương hỗ

- ▶ Trung bình của lượng thông tin  $I(x_j; y_k)$  được gọi là thông tin tương hỗ:
- ▶  $I(X; Y) = E[I(x_j; y_k)] = \sum_{j=1}^m \sum_{k=1}^n P(x_j; y_k) \cdot I(x_j; y_k)$
- ▶  $= \sum_{j=1}^m \sum_{k=1}^n P(x_j; y_k) \cdot \log \frac{P(x_j|y_k)}{P(x_j)}$
- ▶  $= \sum_{j=1}^m \sum_{k=1}^n P(x_j; y_k) \cdot [\log P(x_j|y_k) - \log P(x_j)]$
- ▶  $= H(X) - H(X|Y)$
- ▶  $I(Y; X) = H(Y) - H(Y|X)$

# Tính chất của lượng thông tin tương hỗ

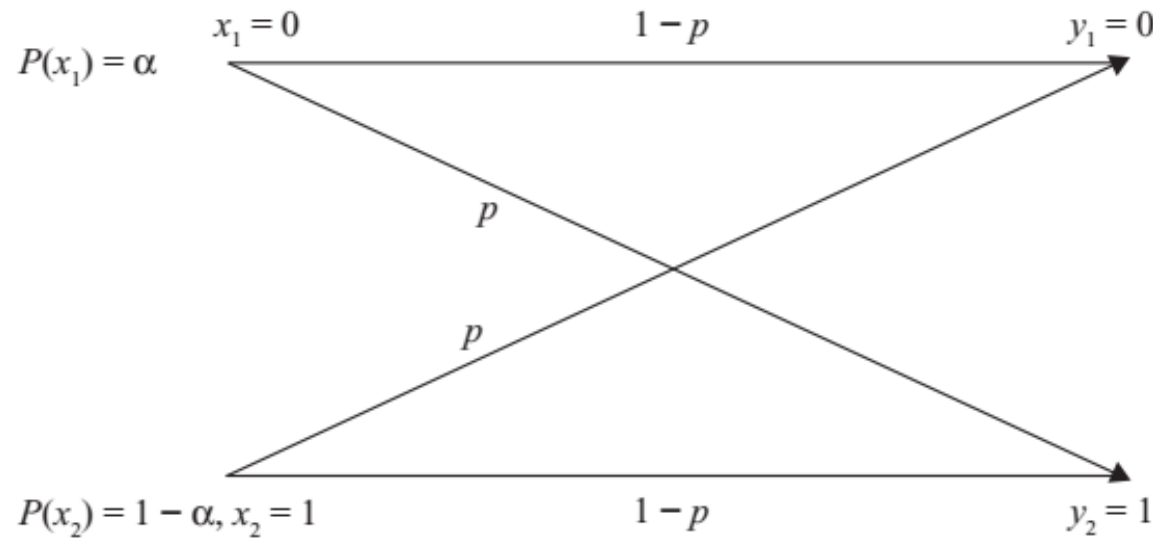
1.  $I(X;Y) \geq 0$
2. Thông tin tương hỗ có tính chất đối xứng;  $I(X;Y) = I(Y;X)$
3.  $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$
4.  $I(X;Y) = H(X) + H(Y) - H(X,Y)$



Lược đồ Venn

## Ví dụ 2.12

- ▷ Xét một kênh rời rạc đối xứng với  $P(x_1) = \alpha$ .
- ▷ (a) Chứng minh rằng:  $I(X; Y) = H(Y) + p \log p + (1 - p) \log(1 - p)$
- ▷ (b) Tính  $I(X; Y)$  với  $\alpha = 0,5$  và  $p = 0,1$
- ▷ (c) Tính  $I(X; Y)$  với  $\alpha = 0,5$  và  $p = 0,5$  và thảo luận về kết quả.



# DUNG LƯỢNG KÊNH

# Định nghĩa

- ▶ **Dung lượng kênh** của một kênh không nhớ rời rạc được định nghĩa là maximum của thông tin tương hỗ. Vì vậy, dung lượng kênh  $C$  được xác định bởi biểu thức:

$$C = \max_{\{P(x_j)\}} I(X; Y) \text{ bit/bản tin}$$

- ▶ Nếu có  $r$  bản tin được truyền trong 1s, tốc độ truyền tin cực đại trong 1s là  $rC$  (bit/s). Đại lượng này được gọi là dung lượng kênh trong 1s **hay khả năng thông qua của kênh**:

$$C' = rC \text{ (bit/s)}$$

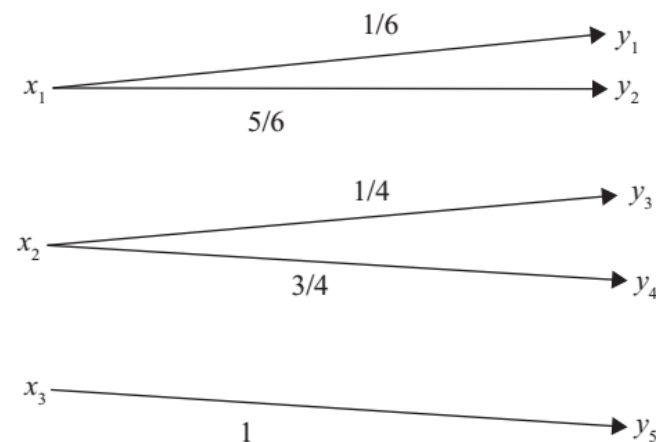


# Các kênh đặc biệt

## ▷ Kênh không tổn thất:

Đối với kênh không tổn thất:  $H(X|Y)=0$

Vì vậy:  $I(X;Y)= H(X) - H(X|Y) = H(X)$



Do đó, thông tin tương hỗ bằng với entropy của nguồn phát và không có thông tin vào bị mất trong quá trình truyền.

Khi đó, dung lượng kênh của kênh không tổn thất là:

$$C = \max_{\{P(x_j)\}} I(X;Y) = \log_2 M \text{ (bit/bản tin)}$$

M là số lượng bản tin chứa trong nguồn X

# Các kênh đặc biệt

## ▷ Kênh xác định:

Đối với kênh xác định:  $H(Y|X)=0$  đối với tất cả các phân phối đầu vào của  $P(x_j)$ .

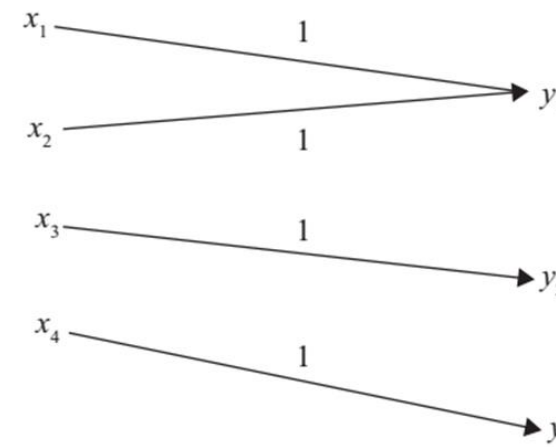
Vì vậy:  $I(X;Y) = H(Y) - H(Y|X) = H(Y)$

Do đó, thông tin tương hỗ bằng với entropy của đầu ra.

Khi đó, dung lượng kênh của kênh xác định là:

$$C = \max_{\{P(x_j)\}} I(X;Y) = \max_{\{P(x_j)\}} H(Y) = \log_2 n \text{ (bit/bản tin)}$$

$n$  là tổng số bản tin nhận được.



# Các kênh đặc biệt

- ▷ Kênh không nhiễu:
- ▷ Kênh không nhiễu vừa là kênh không tổn thất vừa là kênh xác định.

Vì vậy:  $I(X;Y) = H(X) = H(Y)$

Khi đó, dung lượng kênh của kênh không nhiễu là:

$$C = \max_{\{P(x_j)\}} I(X; Y) = \log_2 M = \log_2 n \text{ (bit/bản tin)}$$



# Các kênh đặc biệt

▷ Kênh nhị phân đối xứng (BSC):

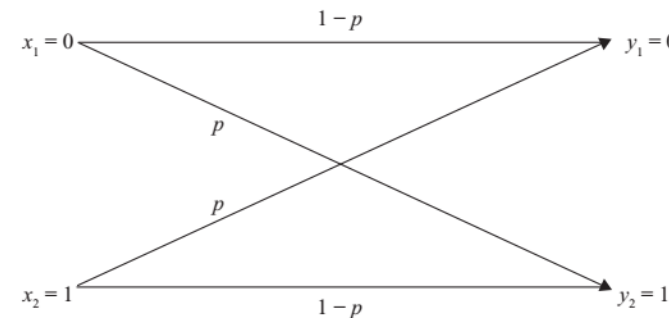
▷ Trong ví dụ 2.12 ta có:

▷  $I(X; Y) = H(Y) + p \log p + (1 - p) \log(1 - p)$

▷ Vì đầu ra là nhị phân nên  $H(Y)_{\max} = \log 2 = 1$  khi  $p(y_1) = p(y_2) = 0,5$ .

Khi đó, dung lượng kênh của kênh BSC là:

$$C = \max_{\{P(x_j)\}} I(X; Y) = 1 + p \log p + (1 - p) \log(1 - p) \text{ (bit/bản tin)}$$



# ĐỊNH LÝ MÃ HÓA KÊNH CỦA SHANNON

# Định lý mã hóa kênh

- ▶ Hiệu năng của hệ thống truyền tin phụ thuộc vào nhiều tham số: độ rộng băng tần kênh, tỉ số tín hiệu/nhiều (SNR), xác suất lỗi bit...
- ▶ Định lý mã hóa kênh của Shannon quan tâm đến tốc độ truyền tin qua kênh truyền:
- ▶ (1) Cho một nguồn không nhớ rời rạc  $X$  với entropy  $H(X)$  và tốc độ thông tin  $r_s = \frac{1}{T_s}$  (bản tin/s). Với một kênh không nhớ rời rạc có dung lượng kênh  $C$  và tốc độ truyền  $r_c = \frac{1}{T_c}$  (bản tin/s) thì nếu:

$$\frac{H(X)}{T_s} \leq \frac{C}{T_c} \text{ hay } r_s \cdot H(X) \leq r_c C$$

thì luôn tồn tại một hệ thống mã hóa sao cho các bản tin được truyền qua kênh có nhiễu và có thể khôi phục lại với xác suất lỗi nhỏ tùy ý.

- ▶ (2) Ngược lại nếu  $\frac{H(X)}{T_s} \geq \frac{C}{T_c}$ , không thể truyền tin qua kênh và khôi phục lại nó với xác suất lỗi nhỏ tùy ý.

# KÊNH LIÊN TỤC

# Kênh liên tục

- ▶ Có nhiều hệ thống truyền tin (điều chế biên độ, điều chế tần số...) sử dụng nguồn liên tục và kênh liên tục.
- ▶ Trong phần này sẽ mở rộng khái niệm lý thuyết thông tin đã phát triển cho kênh rời rạc sang kênh liên tục.



# Entropy vi phân

- ▶ Xét một nguồn thông tin tạo ra tín hiệu liên tục  $x(t)$  được truyền qua một kênh liên tục.
- ▶ Tập các tín hiệu có thể có được coi là một tập hợp các dạng sóng được tạo ra bởi một số quá trình ngẫu nhiên dừng.
- ▶  $x(t)$  được giả định có độ rộng băng tần hữu hạn sao cho nó có thể được biểu diễn một cách hoàn chỉnh bằng một số giá trị mẫu theo chu kỳ của nó.
- ▶ Tại một thời điểm lấy mẫu, tập hợp các giá trị mẫu có thể có tạo thành một biến ngẫu nhiên liên tục  $X$ .

# Entropy vi phân

- ▶ *Lưu ý:* định nghĩa cho entropy có thể được chuyển từ biến ngẫu nhiên rời rạc sang biến ngẫu nhiên liên tục, nhưng khái niệm và ý nghĩa vật lý thì không thể. Lý do là vì lượng thông tin trong biến ngẫu nhiên liên tục là vô hạn, và để biểu diễn chính xác biến ngẫu nhiên liên tục, ta phải yêu cầu một số lượng bit vô hạn để biểu diễn. Lượng thông tin riêng, và vì vậy entropy, là vô hạn.
- ▶ Nếu  $f(x)$  là hàm mật độ xác suất thì lượng thông tin trung bình trên mỗi giá trị mẫu của  $x(t)$  được cho bởi:

$$h(X) = - \int_{-\infty}^{\infty} f(x) \log_2 f(x) dx \text{ bit/mẫu}$$

- ▶ Entropy này được gọi là **entropy vi phân** của  $X$ .

## Ví dụ 2.12

- ▶ Tìm entropy vi phân  $h(X)$  của biến ngẫu nhiên có phân bố đồng nhất:

- ▶ 
$$f(x) = \begin{cases} 1/b & 0 \leq x \leq b \\ 0 & x \text{ khác} \end{cases}$$

# Entropy vi phân có điều kiện

- ▶  $X$  và  $Y$  là các biến ngẫu nhiên với hàm mật độ xác suất kết hợp là  $f(x, y)$  và hàm mật độ xác suất biên  $f(x), f(y)$ .
- ▶ Entropy có điều kiện trung bình của biến ngẫu nhiên liên tục  $X$  với điều kiện  $Y$  được định nghĩa là:
  - ▶ 
$$h(X|Y) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log f(x|y) dx dy$$
  - ▶ Tương tự:
  - ▶ 
$$h(Y|X) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log f(y|x) dx dy$$
  - ▶ Thông tin tương hỗ:
  - ▶ 
$$I(X; Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log \frac{f(x, y)}{f(x)f(y)} dx dy$$

# Nhiều Gauss

- ▷ Nhiều Gauss là nhiễu có hàm mật độ xác suất là phân bố Gauss với trung bình 0 và phương sai  $\sigma^2$  là:

- ▷  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)}$

- ▷ Khi đó entropy vi phân của nhiễu Gauss là:

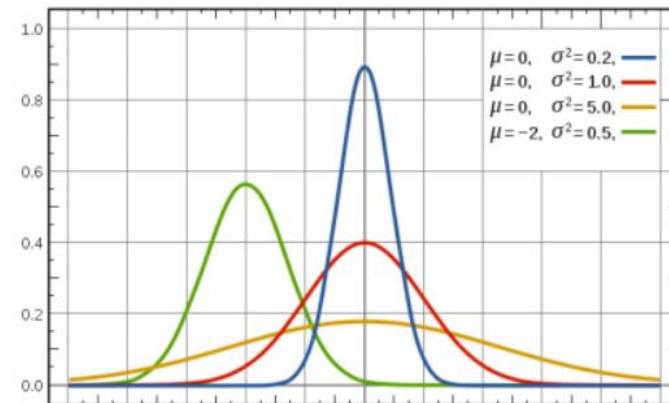
- ▷  $h(X) = - \int_{-\infty}^{\infty} f(x) \log f(x) dx = \int_{-\infty}^{\infty} f(x) \log\left(\frac{1}{f(x)}\right) dx$

- ▷  $= \int_{-\infty}^{\infty} f(x) \log(\sqrt{2\pi\sigma^2}) dx + \int_{-\infty}^{\infty} f(x) \log\left(e^{x^2/(2\sigma^2)}\right) dx$

- ▷  $= \log(\sqrt{2\pi\sigma^2}) \int_{-\infty}^{\infty} f(x) dx + \frac{\log e}{2\sigma^2} \int_{-\infty}^{\infty} f(x) x^2 dx$

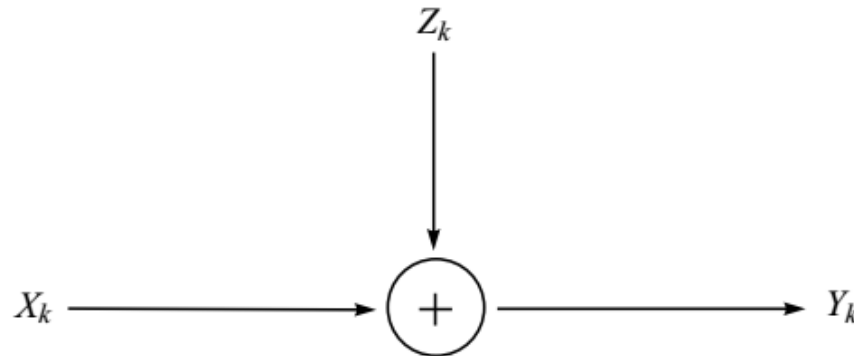
- ▷  $= \log(\sqrt{2\pi\sigma^2}) \times 1 + \frac{\log e}{2\sigma^2} \sigma^2 = \log(\sqrt{2\pi\sigma^2}) + \log\sqrt{e}$

- ▷  $= \log \sqrt{2\pi\sigma^2 e} \text{ bit/bản tin}$



# Kênh Gauss

- ▶ Kênh Gauss là kênh rời rạc về thời gian với đầu ra  $Y_k$  tại thời điểm  $k$  là kết quả của tổng giữa đầu vào  $X_k$  và nhiễu Gauss  $Z_k$ . Nhiễu này có phân bố Gauss với trung bình 0 và phương sai  $\sigma^2$ .
- ▶ Do đó:  $Y_k = X_k + Z_k$ .
- ▶ Nhiễu  $Z_k$  độc lập với đầu vào  $X_k$



# Dung lượng kênh của kênh Gauss

- ▶ Xét một quá trình ngẫu nhiên dừng  $X(t)$  có trung bình 0, được hạn chế băng tần đến  $W(Hz)$ .  $X_k$  ( $k = 1, 2, \dots, K$ ) là các biến ngẫu nhiên liên tục nhận được từ việc lấy mẫu quá trình  $X(t)$  tại tần số Nyquist  $2W$ .
- ▶ Các ký tự này được truyền qua kênh nhiễu cũng được hạn chế băng tần  $W(Hz)$ .
- ▶ Đầu ra kênh chịu tác động của nhiễu Gauss trắng cộng có trung bình 0 và mật độ phổ công suất  $N_0/2$ .
- ▶ Nếu gọi  $Y_k$  là các mẫu của tín hiệu nhận được thì  $Y_k = X_k + N_k$
- ▶  $N_k$  là mẫu nhiễu với trung bình 0 và phương sai  $\sigma^2 = N_0 \cdot W$ .
- ▶ Do máy phát thường bị hạn chế công suất, ta có ràng buộc về công suất trung bình của  $X_k$ :  $E[X_k^2] = P$ .
- ▶ Ta có dung lượng kênh của kênh Gauss:

$$C = \max_{f_{X_k}(x)} \{I(X; Y) | E[X_k^2] = P\}$$

$f_{X_k}(x)$  là hàm phân bố xác suất của  $X_k$

# Dung lượng kênh của kênh Gauss

- ▷ Ta có:  $I(X_k; Y_k) = h(Y_k) - h(Y_k|X_k)$
- ▷ Với  $X_k$  và  $N_k$  là các BNN độc lập. Do đó:  $h(Y_k|X_k) = h(N_k)$
- ▷ Vì  $h(N_k)$  độc lập với  $X_k$  nên để tính  $\max[I(X_k; Y_k)]$  ta tính  $\max[h(Y_k)]$ .
- ▷ **Định lý:** Trong các tín hiệu ngẫu nhiên có cùng công suất trung bình, tín hiệu nào có phân bố Gauss sẽ cho entropy vi phân lớn nhất.
- ▷ Do đó:  $h(Y_k)$  đạt max khi  $Y_k$  là BNN Gauss.
- ▷  $Y_k$  và  $N_k$  đều có phân bố Gauss nên  $X_k$  cũng có phân bố Gauss.
- ▷ Mặt khác, nếu cộng hai BNN Gauss độc lập với nhau, phương sai của BNN tổng sẽ bằng tổng của hai phương sai. Do vậy phương sai của  $Y_k$  bằng:  $P+N_0$ .  $W$ .



# Dung lượng kênh của kênh Gauss

▷ Khi đó:  $\max h(Y_k) = \frac{1}{2} \log[2\pi e(P + N_0 \cdot W)]$

▷  $h(N_k) = \frac{1}{2} \log[2\pi e(N_0 \cdot W)]$

▷ Vậy:  $C = \max[h(Y_k)] - h(N_k) = \frac{1}{2} \log\left(1 + \frac{P}{N_0 \cdot W}\right)$  bit/mẫu

▷ Nếu truyền  $2W$  mẫu/s, ta có:

$$C = W \log\left(1 + \frac{P}{N_0 \cdot W}\right) = W \log(1 + SNR)(\text{bit/s})$$

▷ Đây chính là công thức cơ bản về dung lượng kênh nhiễu Gauss trắng cộng có băng tần hạn chế với đầu vào có công suất trung bình hạn chế và băng tần hạn chế.

▷ Nó còn được gọi là định lý thứ 3 của Shannon, **định lý dung lượng kênh**.

## Ví dụ 2.13

- ▶ Tìm dung lượng của kênh điện thoại có độ rộng băng tần  $B = 3$  kHz và  $\text{SNR} = 40$  dB.

## Ví dụ 2.14

- ▶ Chứng minh rằng dung lượng kênh của một kênh Gauss trắng cộng lý tưởng với độ rộng băng tần vô hạn được cho bởi:
- ▶  $C_{max} = \frac{1}{\ln 2} \frac{S}{\eta} = 1,44 \cdot \frac{S}{\eta}$
- ▶ Ở đó  $S$  và  $\frac{S}{\eta}$  là công suất tín hiệu trung bình và mật độ phổ công suất của nhiễu Gauss trắng cộng tương ứng.

# Giải

▷  $C = B \log \left( 1 + \frac{S}{\eta B} \right)$

▷ Đặt  $\frac{S}{\eta B} = x$

▷ Khi đó:  $C = \frac{S}{\eta x} \log(1 + x) = \frac{1}{\ln 2} \frac{S}{\eta} \frac{\ln(1+x)}{x}$

▷ Khi  $B \rightarrow \infty, x \rightarrow 0$ :

▷  $C_{max} = \frac{1}{\ln 2} \frac{S}{\eta} \lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = \frac{1}{\ln 2} \frac{S}{\eta} = 1,44 \cdot \frac{S}{\eta} \text{ (bit/s)}$

# Bài tập

1. Một hình ảnh truyền hình đen trắng độ phân giải cao có  $3 \times 10^6$  phần tử ảnh và 16 mức sáng khác nhau. Các ảnh được truyền tại tốc độ 24 ảnh/s. Tất cả các mức sáng có tần suất xảy ra như nhau và tất cả các phần tử ảnh độc lập với nhau. Tìm tốc độ thông tin trung bình của nguồn hình ảnh truyền hình này.
2. Một nguồn điện báo tạo ra hai ký hiệu, gạch ngang và dấu chấm. Khoảng thời gian của dấu chấm là 0,2s và khoảng thời gian của dấu gạch ngang gấp 2 lần của dấu chấm. Xác suất của dấu chấm gấp hai lần dấu gạch ngang và khoảng thời gian giữa các ký tự là 0,1s. Tìm tốc độ thông tin của nguồn điện báo.
3. Một kênh Gauss có độ rộng băng tần 1 MHz. Tính dung lượng kênh nếu tỷ lệ công suất tín hiệu trên mật độ phổ nhiễu ( $S/N_0$ ) là  $10^5$ . Tính tốc độ thông tin tối đa.

# Bài tập

4. Một tín hiệu tương tự có độ rộng băng tần 3 kHz được lấy mẫu tại 1,5 lần tốc độ Nyquist. Các mẫu liên tiếp độc lập thống kê với nhau. Mỗi mẫu được lượng tử hóa thành một trong số 256 mức có xác suất như nhau.

- (a) Tìm tốc độ thông tin của nguồn
- (b) Liệu có thể truyền tin không lỗi cho nguồn qua một kênh Gauss trắng cộng với độ rộng băng tần 10 kHz và  $\text{SNR} = 20 \text{ dB}$  hay không?
- (c) Tìm SNR tối thiểu cho việc truyền tin không lỗi ở câu b).
- (d) Xác định độ rộng băng tần yêu cầu cho kênh Gauss trắng cộng đối với việc truyền tin không lỗi khi  $\text{SNR} = 20 \text{ dB}$ .