CƠ SỞ LÝ THUYẾT THÔNG TIN THỐNG KÊ

5/2018

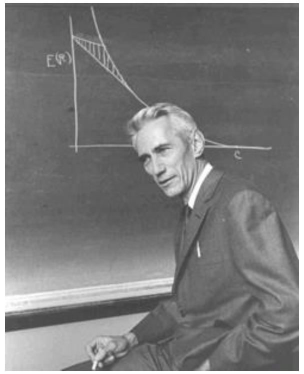
NỘI DUNG

1. Khái niệm lượng tin
2. Entropy của nguồn rời rạc
3. Entropy có điều kiện
4. Khả năng thông qua của kênh
5. Entropy của nguồn liên tục
6. Ôn tập

1. KHÁI NIỆM LƯỢNG TIN

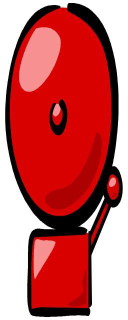
5/2018

THÔNG TIN LÀ GÌ (3)



If sự kiện a_0 chắc chắn xảy ra, $P(a_0) = 1$, thì sự kiện A không mang lại thông tin.

Nếu $A = a_0$ có xác suất xảy ra cao, ta có thể dự đoán một cách khá tự tin về sự kiện sẽ xảy ra, vì vậy lượng thông tin học được rất ít.



Nếu $A = a_0$ rất ít khi xảy ra, điều đó có thể thay đổi đột ngột các dự đoán của chúng ta.

NỘI DUNG

1. Thông tin và lượng thông tin
2. Quan hệ giữa độ bất định và xác suất
3. Tính lượng thông tin

THÔNG TIN LÀ GÌ (1)

- Thông tin có thể cân đong đo đếm được hay không?
- Nếu có thì có thể đo bằng gì?
- Hãy cho ví dụ về một tin chứa nhiều thông tin và một tin chứa ít thông tin.

THÔNG TIN LÀ GÌ (2)

Gỉa sử có sự kiện A với các đầu ra có thể có là $\{a_i\}$.

Mọi thứ chúng ta biết là xác suất của sự kiện đầu ra $P(a_i)$.

Đối với đồng xu, các sự kiện có thể xảy ra là mặt sấp hoặc ngửa:

P(mặt ngửa) = 1/2 & P(mặt sấp) = 1/2.

More generally:

$$\boxed{\begin{gathered} 0 \le P(a_i) \le 1 \\ \sum_i P(a_i) = 1 \end{gathered}}$$

VÍ DỤ: CẦN MẤY CÂU HỎI ĐỂ TÌM RA SỐ ĐƯỢC KHOANH TRÒN NẾU NGƯỜI HỎI CHỈ ĐƯỢC HỎI CÂU HỎI VỚI CÂU TRẢ LỜI DẠNG CÓ/KHÔNG?

1	2
3	4

Bao nhiêu câu hỏi?

2

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Bao nhiêu câu hỏi?

4

Tính ngẫu nhiên nằm ở chỗ sự không chắc chắn về vị trí hình tròn!

THÔNG TIN LÀ GÌ (2)

- Biết càng ít, thông tin chứa trong đó càng nhiều.

- Tin càng bất ngờ sẽ chứa càng nhiều thông tin.

- Nếu biết chắc chắn, sự kiện đó sẽ không chứa thông tin.

Nói cách khác:

- Tin nào có xác suất xảy ra càng nhỏ sẽ chứa nhiều thông tin và ngược lại.

- Tin nào có xác suất bằng 1 sẽ không chứa thông tin.

$$\boldsymbol{p(x_i)} \downarrow \quad \leftrightarrow \quad \boldsymbol{I(x_i)} \uparrow$$

$$\boldsymbol{p(x_i) = 1 \rightarrow I(x_i) = 0}$$

CÂU HỎI

Gỉa sử có 2 sự kiện:

- Tung một đồng xu

- Tung một viên xúc xắc

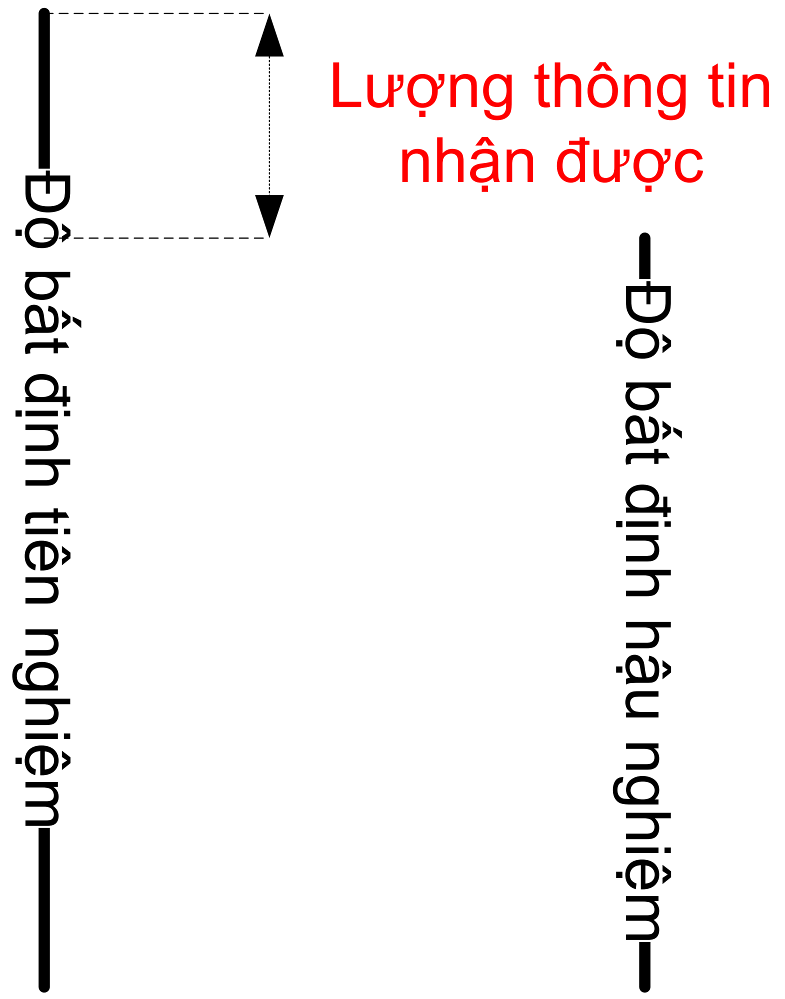
Hỏi sự kiện nào chứa nhiều thông tin hơn?

LƯỢNG THÔNG TIN NHẬN ĐƯỢC

- Lượng thông tin nhận được =

ĐBĐ tiên nghiệm – ĐBĐ hậu nghiệm

- $0 \leq \mathit{LTT\ nhận\ được}\ \leq \mathit{ĐBĐ\ tiên\ nghiệm}$
- Lượng thông tin có mối quan hệ với độ bất định (sự không chắc chắn).
- Vậy độ bất định được tính như thế nào?



QUAN HỆ GIỮA ĐBĐ VÀ XÁC SUẤT

- Giả sử có 2 sự kiện độc lập *A* và *B*, với xác suất tương ứng $P(A) = p_A$ và $P(B) = p_B$.
- $I(p_A)$ và $I(p_B)$ là độ bất định của 2 sự kiện A và B.
- *Tiên đề 1*:

$$I(p_A) \geq 0$$

- *Tiên đề 2*:

Nếu hai sự kiện đồng thời xảy ra, xác suất là $p_A . p_B$. Độ bất định khi đó sẽ là: $I(p_A . p_B) = I(p_A) + I(p_B)$

- *Tiên đề 3*:

Nếu $p_A = 1$ *thì* $I(p_A) = 0$

QUAN HỆ GIỮA ĐBĐ VÀ XÁC SUẤT

- Chỉ có duy nhất hàm Logarithm thỏa mãn tiên đề 3.

$$I(A) = \log(p_A)?$$

- Để thỏa mãn tiên đề 1:

$$I(A) = -\log(p_A) = \log\frac{1}{p_A}$$

- *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*.

VÍ DỤ

Ví dụ 1:

$$p_i = 1 \Rightarrow I(1) = \log\frac{1}{1} = 0$$

Ví dụ 2:

$$p_i = 0.5 \Rightarrow I(0.5) = log_2\frac{1}{0.5} = 1[bit]$$

TÍNH LƯỢNG THÔNG TIN

- Lượng thông tin nhận được = ĐBĐ tiên nghiệm – ĐBĐ hậu nghiệm

$$I(A) = -\log(p_A) = \log\frac{1}{p_A}$$

$\boldsymbol{a_i}$ $\boldsymbol{b_j}$

$p(a_i)$ $p(a_i|b_j)$

$I(a_i) = -logp(a_i)$ $I(a_i|b_j) = -logp(a_i|b_j)$

- Lượng thông tin nhận được:

$$I(a_i; b_j) = I(a_i) - I(a_i|b_j) = -logp(a_i) - (-logp(a_i|b_j))$$

$$\boldsymbol{I(a_i; b_j) = log\frac{p(a_i|b_j)}{p(a_i)}} \quad \textbf{(bit)}$$

BÀI TẬP

Một thành phố nọ có 1% dân số là sinh viên. Trong số sinh viên có 50% là nam thanh niên. Số nam thanh niên trong thành phố là 32% dân số. Giả sử ta gặp một nam thanh niên. Hãy tính lượng thông tin chứa trong tin khi biết rằng đó là một nam sinh viên

2. ENTROPY CỦA NGUỒN RỜI RẠC

5/2018

NGUỒN RỜI RẠC

- Xét nguồn rời rạc A:

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_s \\ p(a_1) & p(a_2) & \dots & p(a_s) \end{pmatrix}$$

- Lượng tin chứa trong a_1:

$$I(a_1) = -logp(a_1)$$

- Lượng tin chứa trong a_2:

$$I(a_2) = -logp(a_2)$$

- .....
- Lượng tin chứa trong a_s

$$I(a_s) = -logp(a_s)$$

$$0 \le P(a_i) \le 1$$

$$\sum_i P(a_i) = 1$$

Lượng thông tin trung bình chứa trong tin a_i bằng bao nhiêu?

ENTROPY CỦA NGUỒN RỜI RẠC

- *Lượng thông tin trung bình chứa trong một tin của nguồn A* hay còn gọi là entropy của nguồn A được tính như sau:

$$H(A) = M\left[I(\mathrm{p}(a_i))\right] = -\sum_i \mathrm{p}(a_i)\log \mathrm{p}(a_i) \quad \text{bits}$$

- Tính chất:

$$0 \le H(A) \le \log s$$

- $H(A) = 0 \leftrightarrow p(a_i) = 1 \text{ và } p(a_j) = 0 \text{ với } \forall j \neq i$
- $H(A) = logs \leftrightarrow p(a_i) = p(a_j) = \frac{1}{s}$

ENTROPY CỦA NGUỒN RỜI RẠC NHỊ PHÂN

Ví dụ: Nguồn nhị phân không nhớ BMS (Binary Memoryless Source)

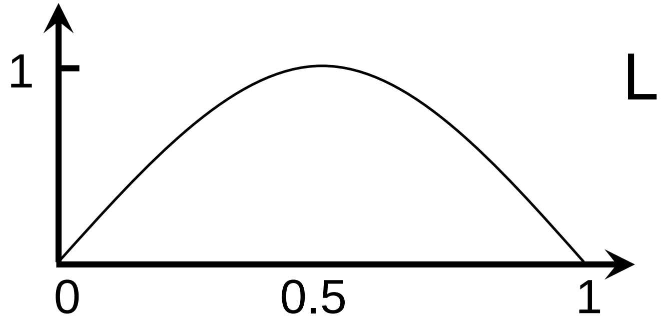
$$\boxed{\text{BMS}} \longrightarrow 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ \ldots$$

Cho $p = P(X_k = 1)$

$q = P(X_k = 0) = 1 - p$

Khi đó: $H = p\log\frac{1}{p} + (1-p)\log\frac{1}{1-p}$

Thường ký hiệu $h(p)$



Lượng thông tin (sự không chắc chắn) đạt cực đại khi:

$$p = q = \frac{1}{2}$$

BÀI TẬP

1. Có 2 hộp đựng bút chì, mỗi hộp đựng 20 bút chì. Hộp thứ nhất có 10 bút chì trắng, 5 bút chì đen và 5 bút chì đỏ. Hộp thứ 2 có 8 bút chì trắng, 8 bút chì đen, 4 bút chì đỏ. Thực hiện các 2 phép thử lấy hú hoạ một bút chì từ mỗi hộp. Hỏi rằng phép thử nào trong hai phép thử nói trên có độ bất định lớn hơn?

2. Một thiết bị điện tử gồm 16 khối có giá trị như nhau về độ tin cậy và được mắc nối tiếp. Giả sử có một khối hỏng. Hãy sử dụng một thiết bị đo tín hiệu ra để xác định khối hỏng. Tính số lần đo trung bình tối thiểu cần thực hiện bằng thiết bị đo này để có thể xác định được khối hỏng. Nêu thuật toán đo? Giả sử khối hỏng là khối thứ 12 hãy chỉ ra các lần đo cần thiết và kết quả đo tương ứng, các phán đoán đưa ra sau mỗi lần đo?

3. ENTROPY CÓ ĐIỀU KIỆN

5/2018

TRƯỜNG SỰ KIỆN ĐỒNG THỜI

- Cho hai nguồn rời rạc A và B

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_s \\ p(a_1) & p(a_2) & \dots & p(a_s) \end{pmatrix}$$

$$B = \begin{pmatrix} b_1 & b_2 & \dots & b_t \\ p(b_1) & p(b_2) & \dots & p(b_t) \end{pmatrix}$$

$$0 \le p(a_i), p(\mathrm{b}_i) \le 1$$

$$\sum_i \mathrm{p}(a_i) = \sum_i \mathrm{p}(\mathrm{b}_i) = 1$$

- Trường sự kiện đồng thời của A và B:

$$AB = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_t & \dots & a_sb_t \\ p(a_1b_1) & p(a_1b_2) & \dots & p(a_1b_t) & \dots & p(a_sb_t) \end{pmatrix}$$

$$0 \le p(a_i \mathrm{b}_j) \le 1$$

$$\sum_i \sum_j p(a_i \mathrm{b}_j) = 1$$

ENTROPY CÓ ĐIỀU KIỆN

$\boldsymbol{a_i}$ —— $\boldsymbol{b_j}$

$p(a_i)$ $p(a_i|b_j)$

$I(a_i) = -logp(a_i)$ $I(a_i|b_j) = -logp(a_i|b_j)$

A ***B***

$H(A) = M[I(p(a_i)]$ $H(A/\mathrm{B}) = M\left[I(p(a_i|b_j)\right]$?

- Entropy có điều kiện về trường tin A khi đã biết trường tin B:

$$H(A/\mathrm{B}) = \sum_i \sum_j I(p(a_i|b_j).p(a_i, b_j)$$

$$\boxed{H(A/\mathrm{B}) = -\sum_i \sum_j p(a_i, b_j).\ logp(a_i|b_j)}$$

H(A|B): Lượng thông tin tổn hao trung bình của một tin của nguồn A khi bên thu nhận được một tin của nguồn B

ENTROPY CÓ ĐIỀU KIỆN

- Tương tự ta có:

$$H(\mathrm{B}|\mathrm{A}) = -\sum_i \sum_j p(a_i, b_j).\, logp(b_j|a_i)$$

- Entropy của trường sự kiện đồng thời:

$$H(\mathrm{A}, \mathrm{B}) = -\sum_i \sum_j p(a_i, b_j).\, logp(a_i, b_j)$$

- Tính chất của entropy có điều kiện:

1) $0 \leq H(A|B) \leq H(A)$
2) $H(A, B) = H(A) + H(B|A) = H(B) + H(A|B)$
3) $H(A, B) \leq H(A) + H(B)$

LƯỢNG THÔNG TIN CHÉO TRUNG BÌNH

$\boldsymbol{a_i}$ —— $\boldsymbol{b_j}$

$p(a_i)$ $\quad$ $p(a_i|b_j)$

$I(a_i) = -logp(a_i)$ $\quad$ $I(a_i|b_j) = -logp(a_i|b_j)$

$\boldsymbol{A}$ $\quad$ $\boldsymbol{B}$

$H(A) = M[I(p(a_i)]$ $\quad$ $H(A/\mathrm{B}) = M[I(p(a_i|b_j)]$

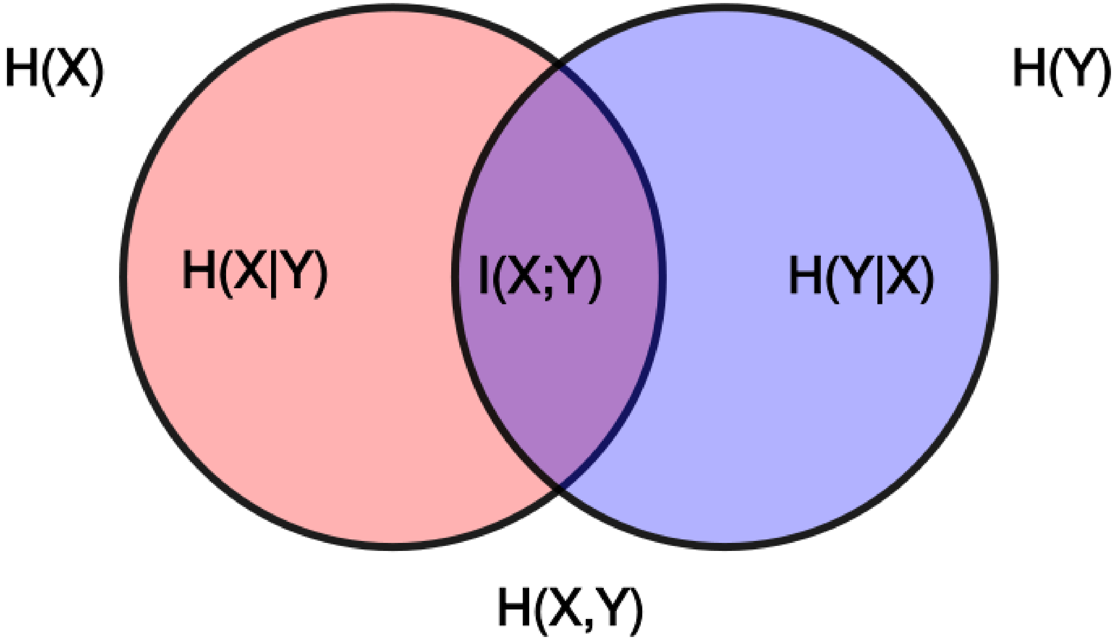
$$I(a_i; b_j) = I(a_i) - I(a_i|b_j)$$

$$I(A, B) = H(A) - H(A|B) = M[I(a_i; b_j)]$$

$$\boxed{I(A, B) = \sum_i \sum_j p(a_i, b_j). log \frac{p(a_i|b_j)}{p(a_i)}}$$

I(A,B): Lượng thông tin trung bình truyền được qua kênh khi thực hiện phát và thu một tin bất kỳ.

LƯỢC ĐỒ VENN



$$
\begin{aligned}
I(X;Y) &= H(X) - H(X \mid Y) \\
&= H(Y) - H(Y \mid X) \\
&= H(X) + H(Y) - H(X,Y)
\end{aligned}
$$

TÍNH CHẤT CỦA LƯỢNG THÔNG TIN CHÉO TRUNG BÌNH

$$I(X;Y) = I(Y;X)$$

$$I(X;Y) \geq 0$$

$$I(X;X) = H(X)$$

$$I(X;Y) \leq \min\{H(X), H(Y)\}$$

$$0 \leq H(X) \leq \log|\mathrm{X}| \longrightarrow$$ Entropy đạt cực đại khi các xác suất bằng nhau

$$Y = g(X), \text{ then } H(Y) \leq H(X)$$

BÀI TẬP

1. Cho sơ đồ một kênh rời rạc không nhớ (DMC) trong đó nguồn phát X gồm hai tin x_1 và x_2; nguồn Y gồm hai tin y_1 và y_2 . Biết $p(x_1)=1/2$, $p(y_1/x_1)=1$, $p(y_1/x_2)=\alpha$, $p(y_2/x_1)=0$, $p(y_2/x_2)=1-\alpha$.

   a. Hãy tính H(X), H(Y), H(X,Y).

   b. Tìm điều kiện của α để H(Y) đạt giá trị cực đại. Khi đó, giá trị của I(X,Y) bằng bao nhiêu (dẫn giải một cách chi tiết nhất có thể để có được kết quả đó).

2. Các tín hiệu x_1 và x_2 có xác suất xuất hiện tiên nghiệm tương ứng là $p(x_1)=3/4$ và $p(x_2)=1/4$ được truyền theo kênh nhị phân rời rạc đối xứng không nhớ có nhiễu có xác suất chuyển sai $p_e=1/8$. Tính:

   a. Lượng tin tức riêng có điều kiện $I(x_2/y_2)$

   b. Lượng tin tức chéo $I(\mathrm{x}_2;y_2)$

   c. Các đại lượng $H(X/y_1)$, $H(\mathrm{X}), H(X,Y), H(X/Y), I(X;Y)$

4. KHẢ NĂNG THÔNG QUA CỦA KÊNH (DUNG LƯỢNG KÊNH)

5/2018

NỘI DUNG

- Vận tốc phát của nguồn rời rạc
- Khả năng phát của nguồn rời rạc
- Khả năng thông qua của kênh
- Định lý về khả năng thông qua của kênh (Định lý mã hóa thứ hai của Shannon)

VẬN TỐC PHÁT CỦA NGUỒN RỜI RẠC

- Xét nguồn rời rạc A:

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_s \\ p(a_1) & p(a_2) & \dots & p(a_s) \end{pmatrix}$$

- Giả sử thời hạn tín hiệu là T_n [s].
- Số tin mà nguồn phát ra trong một đơn vị thời gian là:

$$v_n = \frac{1}{T_n} \text{ [số tin/s]}$$

- v_n: vận tốc phát của nguồn rời rạc

KHẢ NĂNG PHÁT CỦA NGUỒN RỜI RẠC

- Lượng thông tin chứa trong một tin của nguồn: H(A) (bit/tin)
- Số tin mà nguồn phát ra trong một đơn vị thời gian là v_n (tin/s)
- Lượng thông tin mà nguồn phát ra trong một đơn vị thời gian là:

$$H'(A) = H(A).v_n \text{ (bit/s)}$$

- H’(A): Khả năng phát của nguồn rời rạc

KHẢ NĂNG THÔNG QUA CỦA KÊNH (1)

- v_k: vận tốc kênh hay số tin mà kênh truyền được trong một đơn vị thời gian (tin/s).
- H(A) (bit/tin): Lượng thông tin chứa trong một tin của nguồn.
- H(A|B): Lượng thông tin tổn hao trung bình của một tin của nguồn A khi bên thu nhận được một tin của nguồn B
- I(A,B)=H(A)-H(A|B) (bit/tin): Lượng thông tin trung bình truyền được qua kênh khi thực hiện phát và thu một tin bất kỳ.
- Lượng thông tin trung bình truyền qua kênh trong một đơn vị thời gian:

$$v_k.\mathrm{I}(\mathrm{A},\mathrm{B}) \quad \text{(bit/s)}$$

KHẢ NĂNG THÔNG QUA CỦA KÊNH (2)

- Lượng thông tin trung bình *cực đại* truyền qua kênh trong một đơn vị thời gian:

$$C' = \max_{A} v_k . I(A, B) \text{ (bit/s)}$$

- C' cho biết khả năng tải tin tối đa của kênh truyền.
- C' được gọi là khả năng thông qua của kênh.
- $C = \max_{A} I(A, B)$ (bit/một lần sử dụng kênh) gọi là dung lượng kênh.

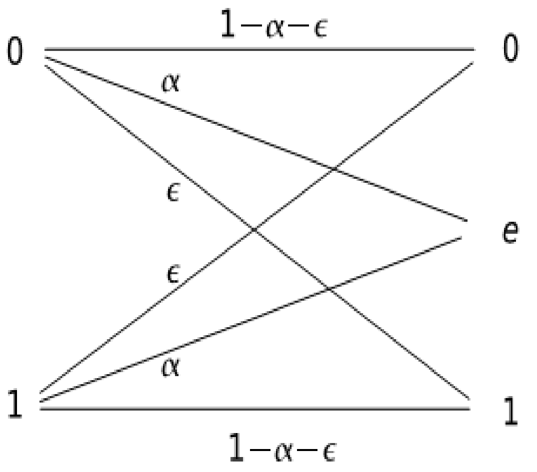
ĐỊNH LÝ VỀ KHẢ NĂNG THÔNG QUA CỦA KÊNH

Luôn tồn tại các bộ mã hóa kênh để có thể đạt được truyền thông tin cậy với xác suất lỗi nhỏ tùy ý, nếu khả năng phát của nguồn nhỏ hơn khả năng thông qua của kênh $H'(A) < C'$.

Nếu $H'(A) \geq C'$ thì không thể tồn tại bộ mã hóa kênh nào để đạt được xác suất lỗi nhỏ tùy ý.

BÀI TẬP

1. Cho sơ đồ kênh rời rạc không nhớ (DMC) như hình vẽ. Biết thời hạn các ký hiệu phát 0 và 1 đều là T_p.
a. Hãy tính dung lượng của kênh.
b. Trong trường hợp kênh nhị phân đối xứng $(\alpha = 0)$ dung lượng kênh bằng bao nhiêu?



2. Xét một kênh rời rạc nhị phân đối xứng không nhớ có ma trận kênh cho như sau: $\begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}$. Biết đầu vào kênh là một nguồn rời rạc nhị phân không nhớ $X = \{0,1\}$ với $p(0) = 1/2$, đầu ra kênh là một nguồn rời rạc nhị phân không nhớ $Y = \{0,1\}$.

a. Hãy tính H(X), H(Y), H(X,Y) và I(X,Y).

b. Xác định các giá trị của ε để dung lượng của kênh đạt cực đại, và cực tiểu.

5. ENTROPY CỦA NGUỒN LIÊN TỤC

5/2018

NỘI DUNG

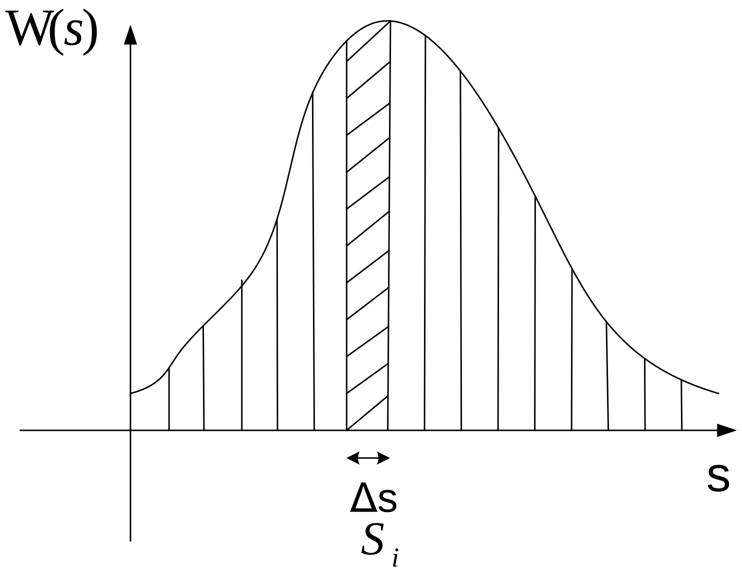
- Entropy vi phân của nguồn liên tục
- Entropy kết hợp và entropy có điều kiện của nguồn liên tục
- Tính chất của entropy vi phân
- Lượng thông tin chéo trung bình
- Khả năng thông qua của kênh Gausse
- Định lý về khả năng thông qua của kênh liên tục

ENTROPY VI PHÂN

- Xét nguồn liên tục S có hàm mật độ xác suất (pdf) là W(s).
- Nhắc lại:
  - $\boxed{\int_{-\infty}^{\infty} W(s)ds = 1}$
  - $\boxed{m = \int_{-\infty}^{\infty} s.W(s)ds}$ - Trung bình của nguồn S
  - $\boxed{\sigma_s^2 = \int_{-\infty}^{\infty} (s-m)^2 W(s)ds}$ - Phương sai của nguồn S
  - Khi nguồn S có trung bình m = 0 thì: $\sigma_s^2 = \int_{-\infty}^{\infty} s^2 W(s)ds$

ENTROPY VI PHÂN

- Chia s thành các đoạn nhỏ đều nhau Δs
- Nguồn S được rời rạc hóa thành nguồn rời rạc S'
- $\boxed{S' = \{s_i, p(s_i)\}; p(s_i) = \Delta s. W\,(s_i)}$



$$H(S') = -\sum_{-\infty}^{\infty} p_i \log p_i = -\sum_{-\infty}^{\infty} \mathrm{W}(\mathrm{s}_i)\Delta s \log(\mathrm{W}(\mathrm{s}_i)\Delta s)$$

$$= -\sum_{-\infty}^{\infty} \Delta s.W(\mathrm{s}_i)\log \mathrm{W}(\mathrm{s}_i) - \sum_{-\infty}^{\infty} W(\mathrm{s}_i)\Delta s \log \Delta s = -\sum_{-\infty}^{\infty} \Delta s.W(\mathrm{s}_i)\log \mathrm{W}(\mathrm{s}_i) - \log \Delta s$$

$$\boxed{H(\mathrm{S}) = \lim_{\Delta s \to 0} H(S') = \int_{-\infty}^{\infty} W(s)\log \frac{1}{W(s)} ds + \lim_{\Delta s \to 0} \log \frac{1}{\Delta s}}$$

ENTROPY VI PHÂN

$$H(S) = \lim_{\Delta s \to 0} H(S') = \int_{-\infty}^{\infty} W(s) \log \frac{1}{W(s)} ds + \lim_{\Delta s \to 0} \log \frac{1}{\Delta s}$$

- $\lim\limits_{\Delta s \to 0} \log \frac{1}{\Delta s} = \infty$

- Số hạng thứ nhất phụ thuộc vào bản chất thống kê của nguồn. Do đó, chọn số hạng thứ nhất đặc trưng cho nguồn.

- Định nghĩa: Entropy vi phân (hay entropy tương đối) của nguồn liên tục S với hàm phân bố xác suất $W(s)$ là:

$$h(S) = -\int_{-\infty}^{\infty} \mathrm{W(s)} \log \mathrm{W(s)} ds = \int_{-\infty}^{\infty} \mathrm{W(s)} \log \frac{1}{\mathrm{W(s)}} ds$$

VÍ DỤ: PHÂN BỐ ĐỒNG NHẤT

Xét biến ngẫu nhiên có phân bố đồng nhất: $f(x) = \begin{cases} \frac{1}{a} & 0 \leq x \leq a \\ 0 & x\ khác \end{cases}$

Tính entropy vi phân của X.

Entropy vi phân của X là: $$h(X) = -\int_0^a \frac{1}{a} \log \frac{1}{a}\, dx = \log a$$

Lưu ý: Với $a < 1$, $\log a < 0$, và entropy vi phân mang giá trị âm. Như vậy, không giống entropy rời rạc, entropy vi phân có thể âm.

VÍ DỤ: PHÂN BỐ CHUẨN

Cho nguồn X có phân bố chuẩn, công suất trung bình (σ^2), trung bình 0:

$$X \sim \Phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2}{2\sigma^2}}$$

Tính toán entropy theo đơn vị bit, ta có:

$$h(\Phi) = -\int \Phi \log \Phi dx = -\int \Phi(x)\left[-\frac{x^2}{2\sigma^2}\log e - \log\sqrt{2\pi\sigma^2}\right]dx$$

$$= \frac{\log e}{2\sigma^2}\int \Phi(x)x^2 dx + \frac{1}{2}\log 2\pi\sigma^2 \int \Phi(x)dx = \frac{1}{2}\log e + \frac{1}{2}\log 2\pi\sigma^2$$

$$= \boxed{\frac{1}{2}\log 2\pi e\sigma^2}$$

ENTROPY KẾT HỢP VÀ ENTROPY CÓ ĐIỀU KIỆN

Định nghĩa:

Cho hai biến ngẫu nhiên liên tục (X, Y) có hàm phân bố xác suất kết hợp $f(x, y)$.

Entropy kết hợp được định nghĩa như sau:

$$\boxed{h(X, Y) = -\iint f(x, y) log f(x, y) dxdy}$$

Entropy có điều kiện được tính như sau:

$$\boxed{h(X|Y) = -\iint f(x, y) log f(x|y) dxdy}$$

TÍNH CHẤT CỦA ENTROPY VI PHÂN

- $h(X)$ có thể âm.
- $h(X,Y) = h(X) + h(Y|X) = h(Y) + h(X|Y)$
- $h(X|Y) \leq h(X)$
- Trong những quá trình có cùng công suất trung bình (σ^2), tín hiệu có phân bố chuẩn sẽ cho entropy vi phân lớn nhất.

$$\max h(X) = \frac{1}{2}\log(2\pi e\sigma^2)$$

LƯỢNG THÔNG TIN CHÉO TRUNG BÌNH

Định nghĩa: *Lượng thông tin chéo trung bình I (X; Y)* giữa hai biến ngẫu nhiên với hàm phân bố xác suất kết hợp *(x, y)* được xác định theo công thức sau:

$$\boxed{I(X;Y) = \iint f(x,y)\log\frac{f(x,y)}{f(x)f(y)}dxdy}$$

Từ định nghĩa ta có:

$$I(X;Y) = h(X) - h(X|Y)$$
$$= h(Y) - h(Y|X)$$
$$= h(X) + h(Y) - h(X,Y)$$

KHẢ NĂNG THÔNG QUA CỦA KÊNH GAUSSE

- *Khả năng thông qua của kênh Gausse* trong trường hợp tín hiệu vào là hàm liên tục của thời gian liên tục với phổ hữu hạn F là:

$$\boxed{C' = F\,log\left(1+\frac{P}{N_0F}\right) = F\,log\left(1+\frac{P}{P_n}\right)}$$

Với N_0 : mật độ phổ công suất thực tế của nhiễu

P_n: công suất trung bình của nhiễu trong dải tần F.

P: công suất trung bình của tín hiệu hữu ích (tín hiệu phát)

- Nhận xét:
  - Khi $F \to 0$ thì $C' \to 0$
  - Khi $F \uparrow$ thì $C' \uparrow$
  - Khi $F \to \infty$ thì $C'_\infty = 1{,}443.\frac{P}{N_0}$

ĐỊNH LÝ MÃ HÓA THỨ HAI CỦA SHANNON ĐỐI VỚI KÊNH LIÊN TỤC

Định lý:

Các nguồn tin rời rạc có thể mã hóa và truyền theo kênh liên tục với xác suất sai bé tùy ý khi giải mã các tín hiệu nhận được nếu khả năng phát của nguồn nhỏ hơn khả năng thông qua của kênh. Nếu khả năng phát của nguồn lớn hơn khả năng thông qua của kênh thì không thể thực hiện được mã hóa và giải mã với xác suất sai bé tùy ý được.

BÀI TẬP

1. Tính độ rộng giải thông của 1 kênh vô tuyến truyền hình truyền hình ảnh đen trắng với 5.10^5điểm ảnh (pixel)/ảnh ; 25 ảnh/s và có 8 mức sáng đồng xác suất, với tỉ số tín/tạp = 15 . Coi rằng ảnh vô tuyến truyền hình xem như 1 dạng tạp âm trắng.

2. Tín hiệu thoại có băng tần W=3,4kHz.

   a. Tính khả năng thông qua của kênh với điều kiện SNR=30dB

   b. Tính SNR tối thiểu cần thiết để kênh có thể truyền tín hiệu thoại số có tốc độ 4800bps.