信息论导论

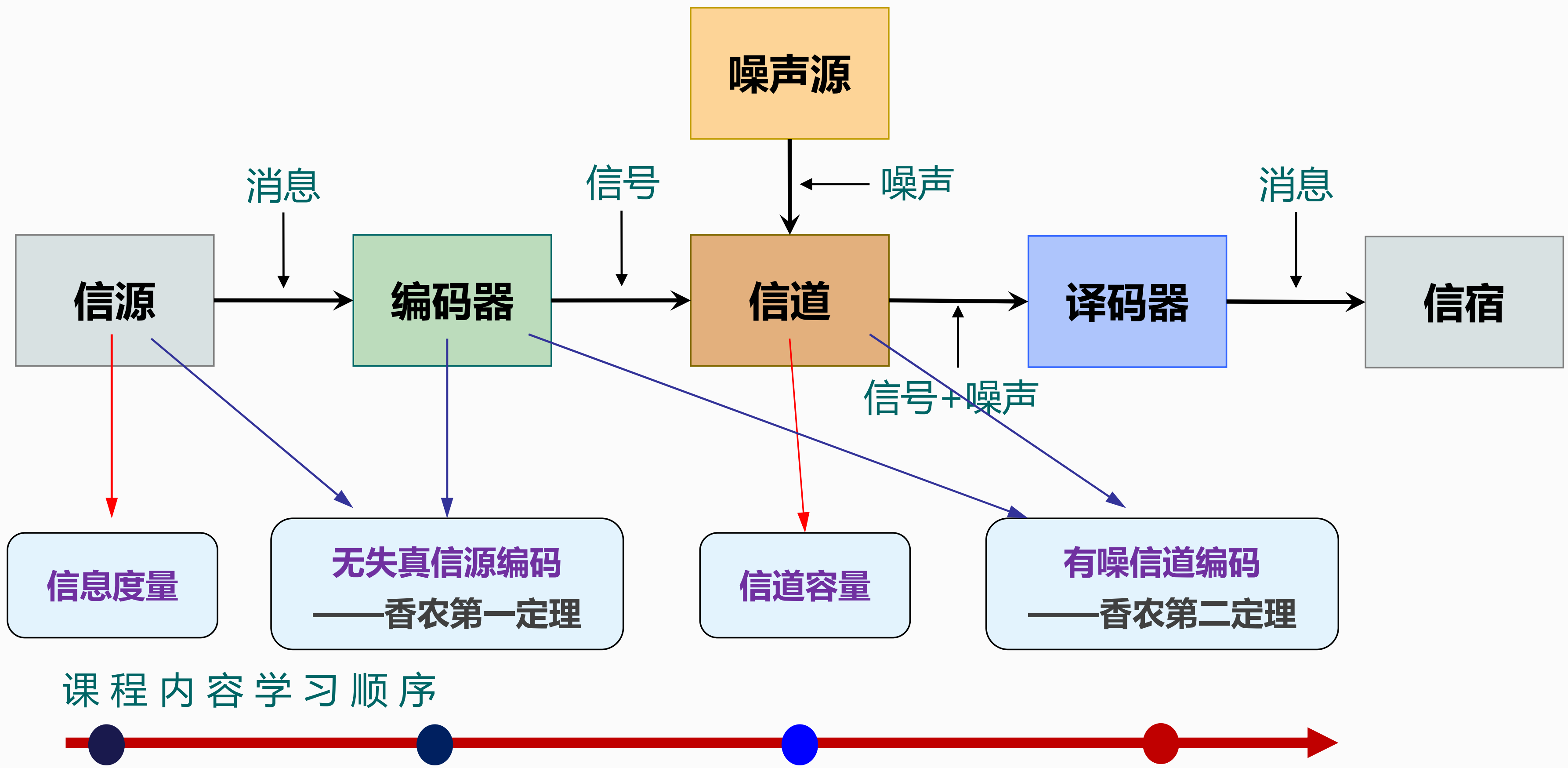
第6讲 离散无记忆信道、信道容量

[信息论教材中页码范围] DMC（离散无记忆信道）、信道容量：p183~p191，互信息的凹凸性：p.33

信息学部-信息科学与技术学院 吴绍华
hitwush@hit.edu.cn

课程内容进度安排

哈爾濱工業大學(深圳)
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN



作为导论课，本课程只讨论“离散”信源、信道

离散无记忆信道（DMC, discrete memoryless channel）

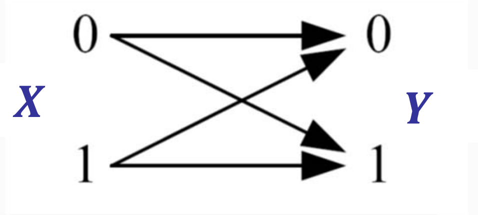
- 输入为 $X \in \mathscr{X}$；输出为 $Y \in \mathscr{Y}$
- 概率转移矩阵 $\mathbf{Q}_{Y|X}$（时不变）

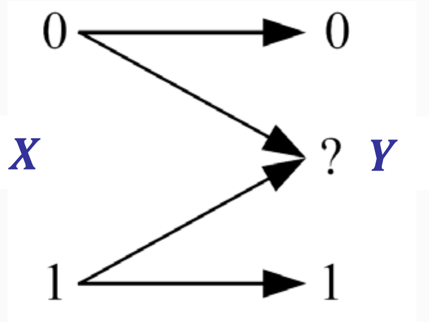
$$(\mathbf{Q}_{Y|X})_{i,j} = p(y = y_j | x = x_i)$$

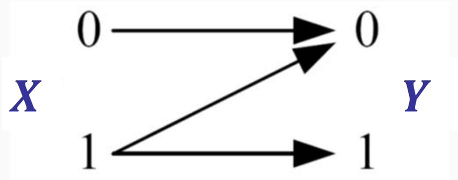
  - 可得 $\mathbf{p}_Y = \mathbf{Q}_{Y|X}^{\mathrm{T}} \mathbf{p}_X$
  - $\mathbf{Q}_{Y|X}$：行和为 1，平均列和为 $|\mathscr{X}||\mathscr{Y}|^{-1}$
  - 简单起见，我们将使用 $\mathbf{Q}$ 来表示 $\mathbf{Q}_{Y|X}$
- 无记忆信道：$\mathbf{p}(Y_n|Y_{1:(n-1)}, X_{1:n}) = \mathbf{p}(Y_n|X_n)$

DMC举例：二元信道

- 二元对称信道
  - $\mathscr{X} = [0;1]$, $\mathscr{Y} = [0;1]$
  - $\mathbf{Q} = \begin{bmatrix} 1-f & f \\ f & 1-f \end{bmatrix}$
- 二元擦除信道
  - $\mathscr{X} = [0;1]$, $\mathscr{Y} = [0;?;1]$
  - $\mathbf{Q} = \begin{bmatrix} 1-f & f & 0 \\ 0 & f & 1-f \end{bmatrix}$
- Z 信道
  - $\mathscr{X} = [0;1]$, $\mathscr{Y} = [0;1]$
  - $\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ f & 1-f \end{bmatrix}$
- 对称信道：$\mathbf{Q}$ 的所有行互为置换，所有列也互为置换。







弱对称信道

弱对称信道：

1. Q 的每一列都具有同样的列和 $|\mathscr{X}||\mathscr{Y}|^{-1}$
   - 如果 X 是均匀分布的（即 $p(x)=|\mathscr{X}|^{-1}$），那么 Y 也是均匀分布的

$$\begin{aligned} p(y) &= \sum_{x\in\mathscr{X}} p(y|x)p(x) = |\mathscr{X}|^{-1}\sum_{x\in\mathscr{X}} p(y|x) \\ &= |\mathscr{X}|^{-1}\times|\mathscr{X}||\mathscr{Y}|^{-1} = |\mathscr{Y}|^{-1} \end{aligned}$$

2. Q 的所有行互为置换
   - Q 的每一行都有着相同的行熵

$$H(Y|X) = \sum_{x\in\mathscr{X}} p(x)H(Y|X=x) = H(\mathbf{Q}_{i,:})$$

其中 $H(\mathbf{Q}_{i,:})$ 为 Q 第 i 行的行熵

对称信道 $\Rightarrow$ 弱对称信道

信道容量

哈爾濱工業大學(深圳)
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

定义

一个 DMC 的信道容量定义为

$$C = \max_{\mathbf{p}_X} I(X;Y)$$

需要在所有可能的输入分布 $\mathbf{p}_X$ 中找到使得 $I(X;Y)$ 取最大值的分布。

信道容量的性质

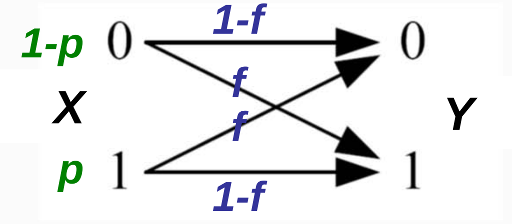
- 由于对于固定的 $\mathbf{p}_{Y|X}$，$I(X;Y)$ 是关于 $\mathbf{p}_X$ 的上凸函数（Concave）（后面将证明），因此必定存在一个唯一的最大值
- C 的上下界：

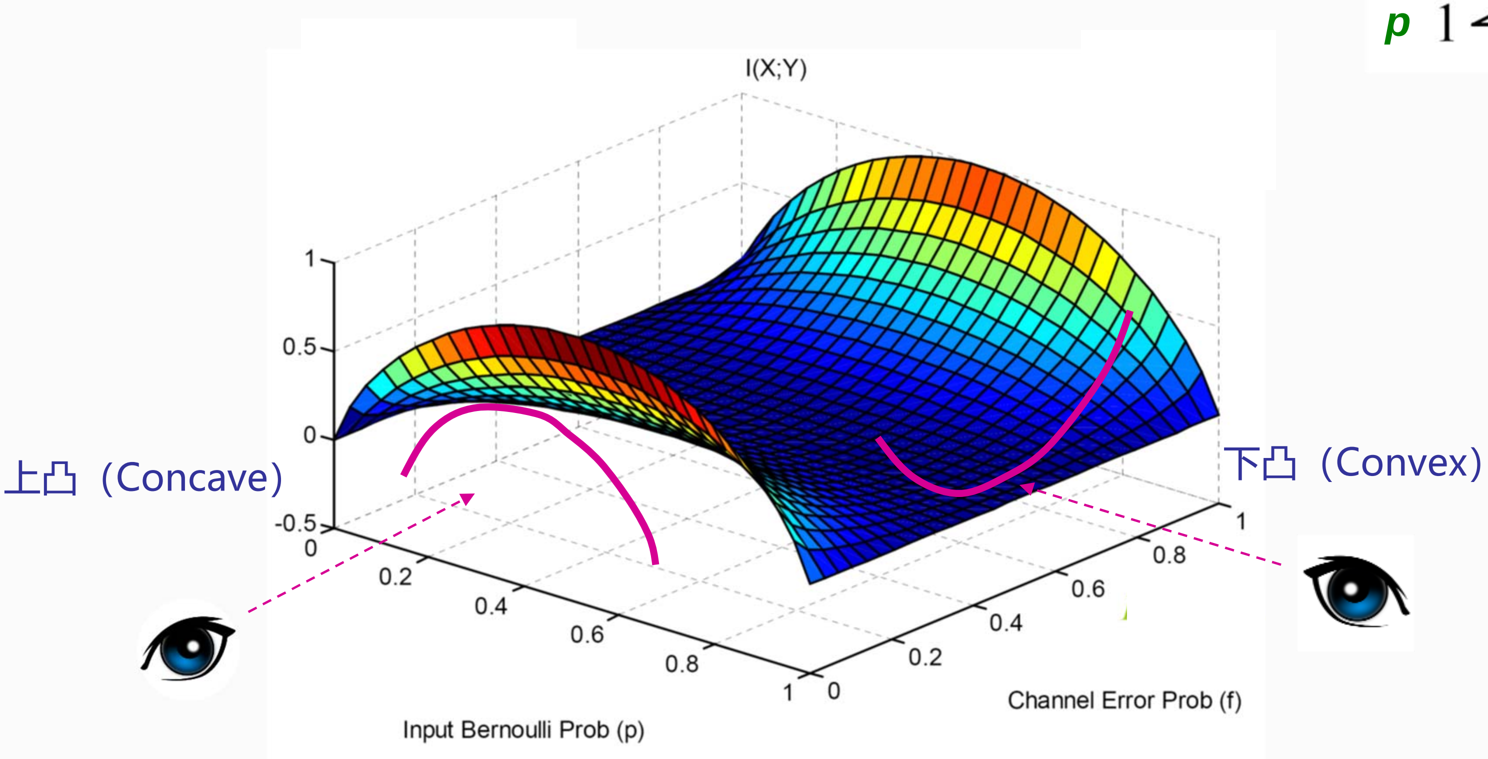
$$0 \leq C \leq \min(H(X), H(Y)) \leq \min(\log|\mathscr{X}|, \log|\mathscr{Y}|)$$

互信息的凹凸性 - 图例

考虑一个输入服从伯努利分布的二元对称信道

$$I(X;Y) = H(p+f-2pf) - H(f)$$



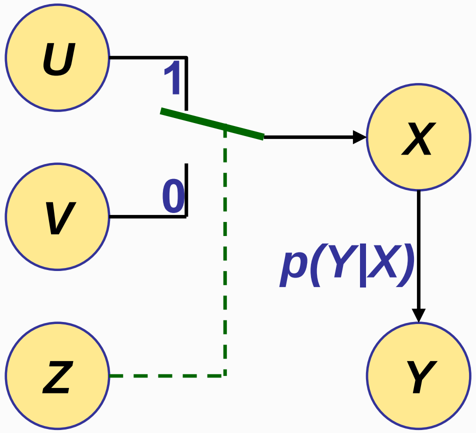


I(X;Y)的凹凸性 – 定理

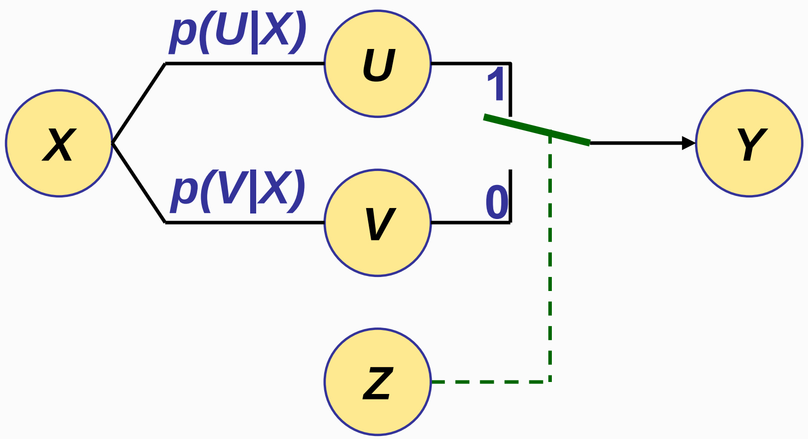
定理 2.7.4

互信息 $I(X;Y)$

- (a) 若固定 $\mathbf{p}_{Y|X}$，是关于 $\mathbf{p}_X$ 的上凸函数（Concave）
- (b) 若固定 $\mathbf{p}_X$，是关于 $\mathbf{p}_{Y|X}$ 的下凸函数（Convex）



(a) 固定 $\mathbf{p}_{Y|X}$



(b) 固定 $\mathbf{p}_X$

证明 - I(X;Y)是关于p_X的上凸函数

证明.

(a) 令 U 与 V 的概率分布向量分别为 $\mathbf{p}_U$ 与 $\mathbf{p}_V$；Z 是一个服从伯努利分布的随机变量且 $p(1)=\lambda$。从而 $\mathbf{p}_X=\lambda\mathbf{p}_U+(1-\lambda)\mathbf{p}_V$

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

而 $I(Z;Y|X)=H(Y|X)-H(Y|Z,X)=0$，因此

$$
\begin{aligned}
I(X;Y) &\geq I(X;Y|Z) \\
&= \lambda I(X;Y|Z=1)+(1-\lambda)I(X;Y|Z=0) \\
&= \lambda I(U;Y)+(1-\lambda)I(V;Y)
\end{aligned}
$$

□

证明 - I(X;Y)是关于$p_{Y|X}$的下凸函数

证明.

(b) 令 U 与 V 的概率分布向量分别为 $\mathbf{p}_{U|X}$ 与 $\mathbf{p}_{V|X}$； Z 是一个服从伯努利分布的随机变量且 $p(1)=\lambda$。从而 $\mathbf{p}_{Y|X}=\lambda\mathbf{p}_{U|X}+(1-\lambda)\mathbf{p}_{V|X}$

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

而 $I(X;Z)=0$ 且 $I(X;Z|Y)\geq 0$，因此

$$\begin{aligned} I(X;Y) &\leq I(X;Y|Z) \\ &= \lambda I(X;Y|Z=1)+(1-\lambda)I(X;Y|Z=0) \\ &= \lambda I(X;U)+(1-\lambda)I(X;V) \end{aligned}$$

□

“N次使用信道（*n*-use channel）”的信道容量

哈爾濱工業大學(深圳) HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

- 对于离散无记忆信道：

$$
\begin{aligned}
I(X_{1:n};Y_{1:n}) &= H(Y_{1:n}) - H(Y_{1:n}|X_{1:n}) \\
&= \sum_{i=1}^{n} H(Y_i|Y_{1:(i-1)}) - \sum_{i=1}^{n} H(Y_i|Y_{1:(i-1)}, X_{1:n}) \\
&= \sum_{i=1}^{n} H(Y_i|Y_{1:(i-1)}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\le \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) = \sum_{i=1}^{n} I(X_i;Y_i)
\end{aligned}
$$

  如要取到等号，则要求 X_i 相互独立 $\Rightarrow$ Y_i 相互独立
- $I(X_{1:n};Y_{1:n})$ 的最大化，可通过独立地最大化每个 $I(X_i,Y_i)$ 实现（而由于各次使用的信道是同一个 DMC，使得各个 $I(X_i,Y_i)$ 最大化的 X_i 必然 i.i.d.）
  - 所以：我们可专注于求解“单次使用信道对应的容量”，即 $\max\limits_{\mathbf{p}_X} I(X;Y)$

如何最大化I(X;Y)——求解信道容量

$$\boxed{C = \max_{p_X} I(X;Y)}$$

- 可以通过标准的非线性优化方法求取最大值，例如梯度搜索
- 一般来说，信道容量没有闭式解
- 但对于许多简单的信道，利用其存在的一些简单的性质（例如对称性），可以进一步求解出其信道容量

(弱)对称信道的信道容量

哈爾濱工業大學(深圳) HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

- 如果信道是弱对称的：

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(\mathbf{Q}_{1,:})$$
$$\leq \log|\mathscr{Y}| - H(\mathbf{Q}_{1,:})$$

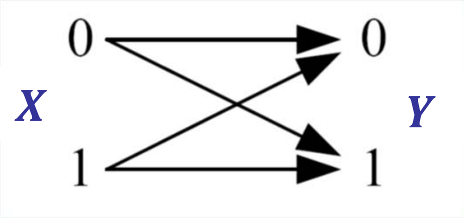
当且仅当输入为均匀分布时取得等号

因此，弱对称信道的信道容量是 $\log|\mathscr{Y}| - H(\mathbf{Q}_{1,:})$

例

对于一个二元对称信道（Binary Symmetric Channel, BSC）：

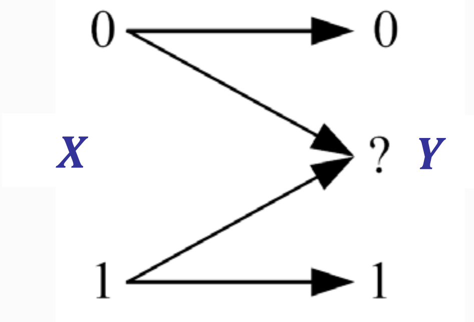
- $|\mathscr{Y}| = 2$
- $H(\mathbf{Q}_{1,:}) = H(f)$
- $I(X;Y) \leq 1 - H(f)$



因此，BSC 的信道容量为 $1 - H(f)$

二元擦除信道（BEC）的容量

- 二元擦除信道
  - $\mathscr{X} = [0;1]$, $\mathscr{Y} = [0;?;1]$
  - $\mathbf{Q} = \begin{bmatrix} 1-f & f & 0 \\ 0 & f & 1-f \end{bmatrix}$



例

对于一个二元擦除信道（Binary Erasure Channel, BEC）：

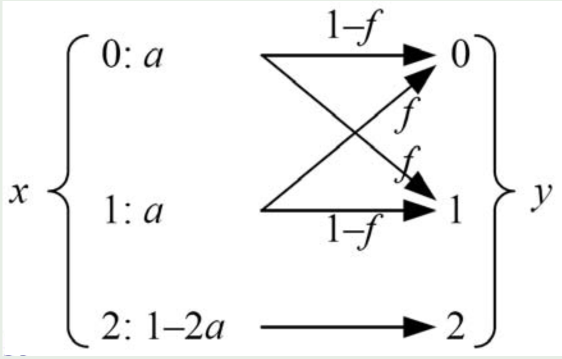
$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(X) - p(Y=0)\times 0 + p(Y=1)\times 0 + p(Y=?)H(X|Y=?) \\ &= H(X) - fH(X) = (1-f)H(X) \le 1-f \end{aligned}$$

因此，BEC 的信道容量为 $1-f$。备注：发送信息 $H(X)$ 中有占比为 f 的部分擦除/丢失了，因此信道容量应为 $(1-f)\max_{\mathbf{p}_X} H(X) = 1-f$（当 X 是均匀分布时达到）。

不对称信道的容量：举例

例

- $\mathscr{X} = \mathscr{Y} = [0; 1; 2], \mathbf{p}_X = [a; a; 1-2a]$
$$\mathbf{Q} = \begin{bmatrix} 1-f & f & 0 \\ f & 1-f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
- 因此 $\mathbf{p}_Y = \mathbf{Q}^{\mathrm{T}} \mathbf{p}_X = \mathbf{p}_X$



$$H(Y) = -2a \log a - (1-2a) \log(1-2a)$$

$$H(Y|X) = 2aH(f) + (1-2a)H(1) = 2aH(f)$$

为了求得 C，我们需要最大化 $I(X;Y) = H(Y) - H(Y|X)$

不对称信道的容量：举例

解

$$I = H(Y) - H(Y|X) = -2a\log a - (1-2a)\log(1-2a) - 2aH(f)$$

$$\frac{\mathrm{d}I}{\mathrm{d}a} = -2\log a - 2\log e + 2\log e + 2\log(1-2a) - 2H(f) = 0$$

$$\Rightarrow \log\frac{1-2a}{a} = \log(a^{-1}-2) = H(f) \Rightarrow a = \left(2 + 2^{H(f)}\right)^{-1}$$

$$\Rightarrow C = -2a\log a - (1-2a)\log(1-2a) - 2a\log\frac{1-2a}{a} = -\log(1-2a)$$

例如，

- $f = 0 \Rightarrow H(f) = 0 \Rightarrow a = 1/3 \Rightarrow C = \log 3 = 1.585$ bits
- $f = 1/2 \Rightarrow H(f) = 1 \Rightarrow a = 1/4 \Rightarrow C = \log 2 = 1$ bits

阶段小结

- 信道容量的定义：

$$C = \max_{\mathbf{p}_X} I(X;Y)$$

- 信道容量的性质：
  - C 的上下限

$$0 \le C \le \min(H(X), H(Y)) \le \min(\log|\mathscr{X}|, \log|\mathscr{Y}|)$$

  - 存在唯一的最大值，因为对于固定的 $\mathbf{p}_{Y|X}$，$I(X;Y)$ 是关于 $\mathbf{p}_X$ 的上凸函数 (Concave)

- 对称信道的容量：

$$\log|\mathscr{Y}| - H(\mathbf{Q}_{1,:})$$

结束