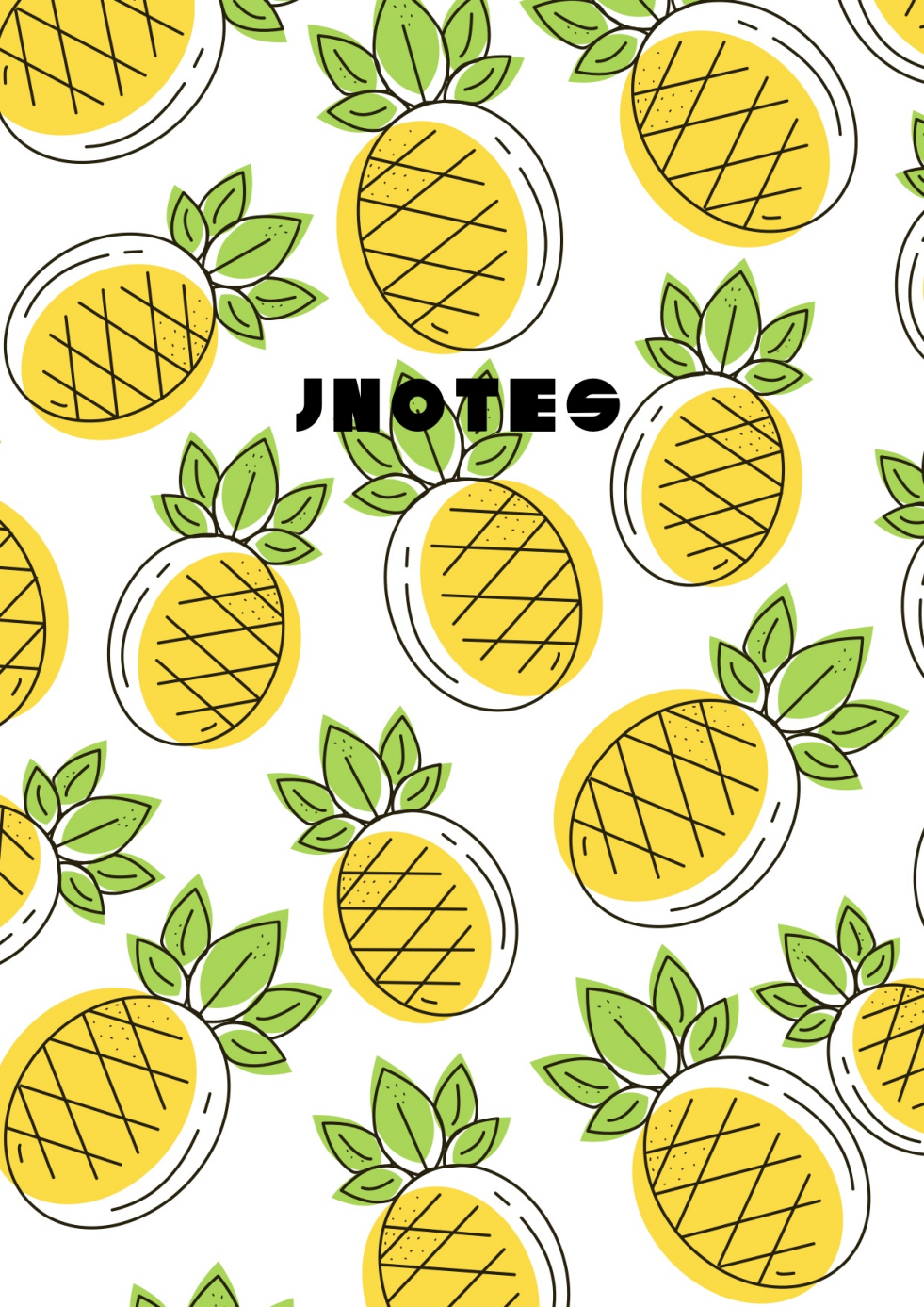


The background of the entire page is a repeating pattern of stylized pineapples. Each pineapple is depicted with a bright yellow body, a black outline, and a crown of five green leaves. The pineapples are arranged in a scattered, overlapping manner across the white background.

**NOTES**

日期: /

# 第1章 概论.

## 信息的概念与分类.

广义信息(音讯.消息) 本质: 事物运动状态的不同描述

1948. 哈特莱 信息量. 1948 年的 非负值非解

1948. 香农 科学定义/定量描述

性质: ① 语法 ② 语义 ③ 语用

## 研究内容.

1. 范围: 语义(信息测度. 编码理论) 一般含噪声/信道/保密

广义(模式识别. 心理学)

2. 主题: 压缩理论(信源编码). 传输理论(信道编码). 保密理论(密码)

3. 模型: 信源. 信宿. 信道. 干扰源

主要功能模块: 信源编码/解码. 信道编码/解码. 加密/解密.

日期: /

## 第二章: 离散信源熵

信源描述方法:

① 随机变量描述 (单符号): 离散/连续信源

符号集:  $A: \{a_1, a_2, a_3, \dots, a_q\}$

离散:  $\begin{bmatrix} x \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \dots & a_q \\ p(a_1) & p(a_2) & \dots & p(a_q) \end{bmatrix} \quad \sum_{i=1}^q p(a_i) = 1$

连续:  $\begin{bmatrix} x \\ p(x) \end{bmatrix} = \begin{bmatrix} (a, b) \\ p(x) \end{bmatrix} \quad \int_a^b p(x) dx = 1$

② 随机矢描述 (多符号, 时间离散) 序列的统计特性  
平稳/非平稳信源 与时间无关

$N$  维随机矢量  $X = (x_1, x_2, \dots, x_N)$  随机序列

平稳信源又分为 无记忆信源, 有记忆信源 (马尔科夫)

离散无记忆信源: 符号是彼此独立的,  $N$  维联合概率密度:

$$p(x) = p(x_1, x_2, \dots, x_N) = p(x_1)p(x_2)\dots p(x_N) = \prod_{i=1}^N p(x_i)$$

$N$  次扩展信源: 一组长度为  $N$

$$\begin{bmatrix} x^N \\ p(x^N) \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{q^N} \\ p(\alpha_1) & p(\alpha_2) & \dots & p(\alpha_{q^N}) \end{bmatrix} \quad \alpha_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iN})$$
$$p(\alpha_i) = p(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iN}) = \prod_{k=1}^N p(a_{ik})$$

不同时刻发出的符号之间是相互依赖的, 有记忆信源

记忆长度为  $m+1$  时  $m$  阶马尔科夫信源

日期: /

贝叶斯符号:  $p(x_i | x_{i-1} x_{i-2} \dots x_{i-m} \dots) = p(x_i | x_{i-1}, x_{i-2}, \dots, x_{i-m})$

$m=1$ 时,  $p(x_i | x_{i-1} = a_i)$

③ 随机过程 (时间连续, 取值连续)

离散信源基本概念和性质.

自信息:  $I(x_i) = -\log p(x_i)$  以2为底 bit  
e nat  
10 det

联合自信息: 二维联合结果  $XY$  上的元素  $(x_i, y_j)$  的联合自信息定义为

$I(x_i, y_j) = -\log p(x_i, y_j)$  其中  $p(x_i, y_j)$  为元素  $(x_i, y_j)$  的二维联合概率

条件自信息:  $I(x_i | y_j) = \log \frac{1}{p(x_i | y_j)}$

性质: 1) 自信息为非负值 2) 关于  $p(x)$  的递减函数

3) 可加性  $I(x_i, y_j) = I(x_i | y_j) + I(y_j)$

相互独立:  $I(x_i, y_j) = I(x_i) + I(y_j)$

信息熵 (平均不确定性/自信息)

$H(X) = \sum p(x_i) I(x_i)$

离散信源熵:  $H(X) = \sum p(x_i) I(x_i) = -\sum p(x_i) \log p(x_i)$  零概率事件,  $H(x) = 0$

熵的性质: 对称性, 非负性, 极值性 (等概率时取最大), 上凸 (严格凸)

条件熵:  $H(Y|X) = \sum_i p(x_i) \sum_j p(y_j | x_i) I(y_j | x_i) = -\sum_{x_i, y_j} p(x_i, y_j) \log p(y_j | x_i)$

熵是联合概率, 而不是条件概率 ( $I(y_j | x_i)$  为加权平均)

日期:

(东大商)

联合熵: 两个元素  $x_i, y_j$  的自信息的加权平均值.

$$H(XY) = \sum_{x,y} P(x_i, y_j) I(x_i, y_j) = - \sum_{x,y} P(x_i, y_j) \log P(x_i, y_j)$$

关系 1)  $H(XY) = H(X) + H(Y|X)$     2)  $H(XY) = H(Y) + H(X|Y)$

3)  $H(XY) \leq H(X) + H(Y)$  ( $X$ 与 $Y$ 独立, 等号成立)

4)  $H(Y|X) \leq H(Y)$

多符号离散平稳信源熵

- 1位平稳:  $P(x_i) = P(x_j), i \neq j$

- 2位 ~:  $P(x_i x_{i+1}) = P(x_j x_{j+1}), i \neq j$

...

$N$ 位:  $P(x_i \dots x_{i+N}) = P(x_j \dots x_{j+N}), i \neq j$  ( $N \rightarrow \infty$ ) (完全离散平稳信源)

数字类型:

$$\begin{bmatrix} x \\ p(x) \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{2^n} \\ p(\alpha_1) & p(\alpha_2) & \dots & p(\alpha_{2^n}) \end{bmatrix} \quad \begin{aligned} X &= x_1 x_2 \dots x_N \\ x_i &\in \{a_1, a_2, \dots, a_{2^n}\} \end{aligned}$$

$$\alpha_i = (a_{i1}, a_{i2}, \dots, a_{iN}) \quad 0 \leq p(\alpha_i) \leq 1$$

$$p(\alpha_i) = p(a_{i1} a_{i2} \dots a_{iN}) \quad \sum_{i=1}^{2^n} p(\alpha_i) = 1$$

$N$ 位离散信源熵:  $H(X) = H(X^N) = - \sum_{x^N} P(x^N) \log P(x^N) = - \sum_{x^N} P(\alpha_i) \log P(\alpha_i)$

$$H(X^N) = N H(X)$$

$N$ 位信源:  $H(X) = H(x_1 \dots x_N) = H(x_1 \dots x_{N-1}) + H(x_N | x_1 \dots x_{N-1})$

$$= H(x_1) + H(x_2 | x_1) + \dots + H(x_N | x_1 \dots x_{N-1})$$

日期: /

平均符号熵:  $N$  长信源符号序列中, 平均的信源符号所携带的信息量, 即,

$$H_N(x) = \frac{1}{N} H(x_1 x_2 \dots x_N)$$

当  $H(x) < \infty$ , 则有如下的性质:

① **条件熵**  $H(x_N | x_1 x_2 \dots x_{N-1})$  随  $N$  增加是非递增的。

$$H(x_N | x_1 x_2 \dots x_{N-1}) \leq H(x_{N-1} | x_1 x_2 \dots x_{N-2}) \leq \dots \leq H(x_2 | x_1) \leq H(x)$$

② 当  $N$  充分大, **平均符号熵** 大于等于 **条件熵**  $H_N(x) \geq H(x_N | x_1 x_2 \dots x_{N-1})$

③ **平均符号熵** 随  $N$  的递增也是非递增的  $H_i(N) \geq H_{i+1}(N)$

④  $H_\infty = \lim_{N \rightarrow \infty} H_N(x)$  存在, 并且  $H_\infty = \lim_{N \rightarrow \infty} H_N(x) = \lim_{N \rightarrow \infty} H(x_N | x_1 x_2 \dots x_{N-1})$

称为 **极限熵** 或 **极限信息量**

熵公式:  $H_N(x) = \frac{1}{N} H(x_1 x_2 \dots x_N) = \frac{1}{N} \cdot N H(x) = H(x)$

$$H_\infty = \lim_{N \rightarrow \infty} H_N(x) = H(x)$$

马尔可夫信源: **非平稳**

**稳定状态是平稳**

$X \in A = \{a_1, a_2, \dots\}$  输出

$S \in E = \{E_1, E_2, \dots\}$  状态空间

若信源输出的符号序列和信源所处的状态满足以下两个条件,

(1) 某时刻信源符号输出只与此信源所处状态有关, 而与以前状态及以前的输出符号都无关, 即:

$$P(X_n = a_k | S_n = E_i, X_{n-1} = a_{k-1}, S_{n-1} = E_j, \dots) = P(X_n = a_k | S_n = E_i) \stackrel{\text{统计}}{=} p(a_k, E_i)$$

(2) 信源在  $n$  时刻所处的状态由前  $n-1$  时刻的输出符号和前一时刻  $(n-1)$  状态唯一确定

日期: /

即:  $P(S_n = E_j | X_n = a_k, S_{n-1} = E_i) = \begin{cases} 1 & E_j \cdot E_j \in E \\ 0 & a_k \in A \end{cases}$

马尔可夫链。

$$[P^{(n)}] = [P(a_{k_{n+1}} | a_{k_1} a_{k_2} \dots a_{k_n})] \quad (k_1, k_2, \dots, k_n, k_{n+1} = 1, 2, \dots, q)$$

当  $n=1$  时，与前一符号有关。 - 马尔可夫链

一般马尔可夫链的熵:  $H_{\infty} = H_{\infty}(X) = \lim_{N \rightarrow \infty} \frac{1}{N} H(X_1 X_2 \dots X_N)$

从状态  $E_i$  下输出一个符号下平均熵为:

$$H(X | S = E_i) = - \sum_{k=1}^q P(a_k | E_i) \log P(a_k | E_i)$$

马尔可夫链遍历的马尔可夫链:

$$H_{\infty} = \sum_{i=1}^I Q(E_i) H(X | E_i) = - \sum_{i=1}^I \sum_{k=1}^q X(E_i) P(a_k | E_i) \log P(a_k | E_i)$$

各状态的平稳分布概率  $\downarrow$  条件熵

其中:  $\sum_{E_j \in E} Q(E_j) P(E_i | E_j) = Q(E_i) \quad E_i = (a_{k_1}, a_{k_2}, a_{k_3}, \dots, a_{k_n})$

马尔可夫链遍历的马尔可夫链:

$$H_{\infty} = H_{n+1}(X) = H(X_{n+1} | X_1 X_2 X_3 \dots X_n)$$

信源相关性定理名称。

$H_{\infty} = 1.978$

1. 熵的相对率:  $\eta = \frac{H_{\infty}}{H_0}$ : 实际信息熵与具有同样符号数一最大熵的比值。

2. 信源冗余度:  $\sigma = 1 - \eta = 1 - \frac{H_{\infty}}{H_0}$

3. 信息冗余:  $I_{\infty} = H_0 - H_{\infty}$

日期: /

无失真离散信源编码

有效性  $\rightarrow$  信源编码. 提高信息传输速率

可靠性  $\rightarrow$  信道编码 抗干扰能力

安全性  $\rightarrow$  密码

码的分类: 等长码, 变长码, 非奇异码, 奇异码  $\rightarrow$  一个码中有相同的码字

$\swarrow$   
所有码字都不相同

唯一可译码: 要求: ① 不同信源符号变换成不同的码字.

② 信源字母表有限长  $N$  个码字都是非奇异码

分为即时码, 非即时码: 看任意一个码字是否是其他码字的后缀部分

$\downarrow$   
都可以用树图法来表示

当码字长度给定时, 即时码不是唯一的

唯一可译码存在 的充要条件.

$\rightarrow$  进制数

$k$ : 信源符号数

定理 (Kraft不等式): 对于  $r$  元字母表上的即时码, 码字长度  $k_i$  必须满足不等式

$$\sum_{i=1}^n r^{-k_i} \leq 1$$

(只能用来判断是否存在唯一可译码, 而不能作为判据)

日期: /

## 离散无失真信源编码定理

信源码  $X_1 X_2 \dots X_{L-1} X_L$

码字  $Y_1 Y_2 \dots Y_k \dots Y_K$

使平均码率最小, 即

$$F = \frac{1}{L} \log m = \frac{1}{L} \log M \text{ 最小 } M = m^k$$

求信源熵.

码长  $k$  是变量  $\rightarrow$  寻找最小的  $k$

$$\text{输入 } 2^L \quad \text{输出 } m^k \quad 2^L \leq m^k \quad \frac{k}{L} \geq \frac{\log 2}{\log m}$$

定理1 (渐近等分割性定理)

$$N \rightarrow \infty, \quad -\frac{1}{N} \log p(X_1 X_2 \dots X_N) \xrightarrow{P} H(X)$$

$$\text{依概率收敛 } \lim_{N \rightarrow \infty} P \left( \left| \frac{1}{N} \log p(X_1 X_2 \dots X_N) - H(X) \right| < \epsilon \right) = 1$$

定理2:  $\epsilon$ -典型集:

$$A_{\epsilon N} \equiv \{ \alpha_i : \left| -\frac{1}{N} \log p(\alpha_i) - H(X) \right| \leq \epsilon \}$$

性质: ① 高概率  $P(A_{\epsilon N}) > 1 - \delta$

$$\text{② 几乎等概 } \alpha_i \in A_{\epsilon N} \text{ 则 } 2^{-N(H(X)+\epsilon)} < p(\alpha_i) < 2^{-N(H(X)-\epsilon)}$$

$$\text{③ 有限性 } (1-\delta) 2^{N(H(X)-\epsilon)} \leq |A_{\epsilon N}| \leq 2^{N(H(X)+\epsilon)}$$

日期: /

变长信源编码定理:

设符号  $H(X)$  的无记忆平稳信源  $X_1 X_2 \dots X_L$  由  $k$  个

$Y_1 \dots Y_k \dots Y_k$  (每个符号有  $m$  可能值) 进行变长编码. 对  $\forall \epsilon > 0, \delta > 0$ .

满足:  $\frac{k}{L} \log m \geq H(X) + \epsilon$   $L$  足够大, 编码差错小于  $\delta$  元无失真编码.

反之  $\frac{k}{L} \log m \leq H(X) - 2\epsilon$  不可实现无失真编码.  $L$  足够大, 必定出错.

(1) 定理中发现元无失真编码的条件为:

$$k \log m > L H(X)$$

(2) 信源的**信息传输率**: 编码器平均每个信源符号能携带的最大信息量

$$R' = \frac{k}{L} \log m = \frac{1}{L} \log M$$

对于变长编码, 若要实现元无失真编码, 当允许错误概率小于  $\delta$ , 则

信源序列长度必须满足:

$$L \geq \frac{\sigma^2(X)}{\epsilon^2 \delta} \quad \sigma^2(X) = E [E I(X_i) - H(X)]^2$$

**编码效率:**  $\eta = \frac{H(X)}{R'} = \frac{H(X)}{\frac{k}{L} \log m}$

**最佳效率:**  $\eta = \frac{H(X)}{H(X) + \epsilon} \quad (\epsilon > 0) \quad \text{接近于 } 1$

日期: /

变长编码定理:

码长  $k$  是变长的. 目的是找到 **最小平均值**.

概率大  $\rightarrow$  短码. 概率小  $\rightarrow$  长码

$$\text{信源: } \begin{bmatrix} x \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \dots & a_g \\ p(a_1) & p(a_2) & \dots & p(a_g) \end{bmatrix}$$

编码后:  $w_1, w_2, \dots, w_g$  码长:  $l_1, l_2, \dots, l_g$

$$\bar{L} = \sum_{i=1}^g p(a_i) l_i \quad \text{熵: } R = H(Y) = \frac{H(x)}{\bar{L}}$$

单符号变长编码定理:

无失真编码方法, 构成一行一列的码. 其码字平均长度满足下列不等式:

$$\frac{H(x)}{\log m} \leq \bar{L} < \frac{H(x)}{\log m} + 1 \quad H_m(x) \leq \bar{L} < H_m(x) + 1$$

无失真变长信源编码定理 (香农第一定理)

无记忆平稳信源  $X$  在  $N$  次独立同分布信源  $X^N$   $H(X^N)$   $m$  进制编码

$$\text{满足: } \frac{H(x)}{\log m} \leq \frac{\bar{L}_N}{N} < \frac{H(x)}{\log m} + \frac{1}{N} \quad \text{当 } N \rightarrow \infty \quad \lim_{N \rightarrow \infty} \frac{\bar{L}_N}{N} = \frac{H(x)}{\log m} = H_m(x)$$

$$\text{有证: } \lim_{N \rightarrow \infty} \frac{\bar{L}_N}{N} = \frac{1}{\log m} \lim_{N \rightarrow \infty} \frac{H(X_1 X_2 \dots X_N)}{N} = \frac{H_m(x)}{\log m}$$

日期: /

信源的信息传输率:  $R' = \frac{L_N}{N} \log m < H(X) + \frac{\log m}{N}$

香农第一定理: 若  $R' > H(X)$  有唯一-强变长码, 当  $R' < H(X)$  则不存在

信道的信息传输率:  $R = \frac{H(X)}{L}$

$$L = \frac{L_N}{N} \geq \frac{H(X)}{\log m} \quad R = \frac{H(X)}{L} \leq \log m$$

编码效率:  $\eta = \frac{H(X)}{\log m} = \frac{H(X)}{L \log m}$

无记忆:  $\eta = \frac{H_m(X)}{L} = \frac{H(X)}{L \log m} > \frac{H(X)}{H_m(X) + \frac{\log m}{N}}$

冗余:  $1 - \eta = 1 - \frac{H(X)}{L \log m}$

日期: /

香农编码:

每个码字长度  $k_i$  满足:  $k_i = \lceil \log_2 \frac{1}{p(x_i)} \rceil$

$$即: -\log_2 p(x_i) \leq k_i < -\log_2 p(x_i) + 1$$

这种编码即为 Shannon 编码.

二元香农编码步骤:

(1) 按概率排序  $p(x_1) \geq p(x_2) \geq \dots \geq p(x_n)$

(2) 令  $p_1 = p(x_1) = 0$ ,  $p_i$  表示  $i$  个码字的累积概率.  $p_i = \sum_{k=1}^i p(x_k)$

(3) 求出  $k_i$

(4) 将  $p_i$  用二进制表示, 并取小数点后  $k_i$  位作为  $x_i$  的编码

$$例: \begin{bmatrix} x \\ p(x) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ 0.4 & 0.3 & 0.2 & 0.05 & 0.05 \end{bmatrix}$$

$p_i$  用二进制表示, 取小数点后  $k_i$  位作为  $x_i$  的编码

$k_i$     $p(x_i)$     $p_i$     $-\log_2 p(x_i)$     $k_i$    码字

$x_1$    0.4   0   1.32   2   00

0.000

$x_2$    0.3   0.4   1.73   2   01

0.0110

$x_3$    0.2   0.7   2.32   3   101

0.101

$x_4$    0.05   0.95   4.3   5   11100

0.11101

$x_5$    0.05   0.95   4.3   5   11101

11110

日期:

/

费诺编码: 属于概率函数编码. 步骤如下:

(1) 按概率排序

(2) 按编码进制将概率分组 每组概率尽量接近或相等. (以进制为基数)

(3) 每一组分配一码元

(4) 重复上述步骤

$$\text{例: } [P_X] = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ 0.4 & 0.3 & 0.2 & 0.05 & 0.05 \end{bmatrix}$$

| $x_i$ | $p(x_i)$ | 第1次分组 | 2次 | 3次 | 码字 | 码长 |
|-------|----------|-------|----|----|----|----|
|-------|----------|-------|----|----|----|----|

|       |     |   |  |  |  |  |
|-------|-----|---|--|--|--|--|
| $x_1$ | 0.4 | 0 |  |  |  |  |
|-------|-----|---|--|--|--|--|

|       |     |  |   |  |  |  |
|-------|-----|--|---|--|--|--|
| $x_2$ | 0.3 |  | 0 |  |  |  |
|-------|-----|--|---|--|--|--|

|       |     |  |   |   |  |  |
|-------|-----|--|---|---|--|--|
| $x_3$ | 0.2 |  | 1 | 1 |  |  |
|-------|-----|--|---|---|--|--|

|       |      |  |  |  |  |  |
|-------|------|--|--|--|--|--|
| $x_4$ | 0.05 |  |  |  |  |  |
|-------|------|--|--|--|--|--|

|       |      |  |  |  |  |  |
|-------|------|--|--|--|--|--|
| $x_5$ | 0.05 |  |  |  |  |  |
|-------|------|--|--|--|--|--|

日期: /

霍夫曼编码: 文字中, 比较高的字符无失真信源编码方法

步骤: (1) 依概率排序

(2) 取两个 **概率最小** 的符号分别配以 0 和 1 两个码元, 并将这两个 **概率相加**

作为一个 **新符号** 的概率, 与未分配的 **二进制符号** **重新排列**.

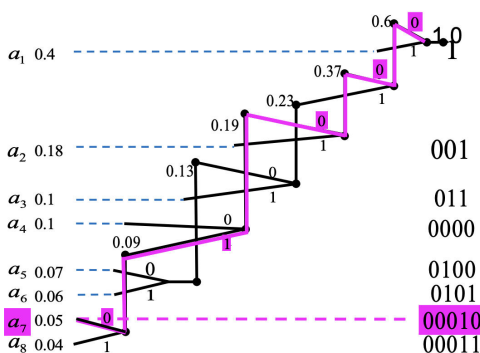
(3) 对重新排列的两个 **概率最小** 符号重复步骤 (2) 的过程

(4) 不断重复上述过程, 直到最后两个符号配以 0 和 1

(5) 从最后一级开始, 向前返回得到各个信源符号所对应的码元序列.

$$p(x) = [a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8]$$

$$p(x) = [0.4, 0.18, 0.1, 0.1, 0.07, 0.06, 0.05, 0.04]$$



单符号信源: 编码效率主要取决于

信源熵和平均码长之比

001

011

0000

0100

0101

00011

$$\eta = \frac{H(X)}{K}$$

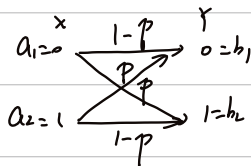
冗余符号 **新符号** 尽可能排在靠后

位置

日期: /

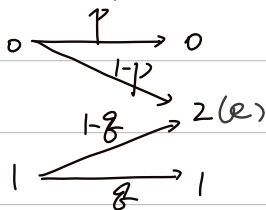
离散信道容量

二元对称信道



$$[P] = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

二元删除信道



p. 无错误传输概率.

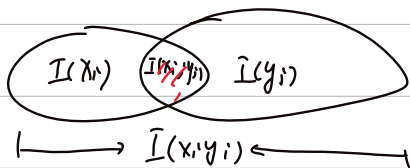
$$[P] = \begin{bmatrix} 1-p & p & 0 \\ 0 & 1-p & p \\ 0 & p & 0 \end{bmatrix}$$

互信息: 两个离散随机事件  $X$  和  $Y$ , 事件  $y_i$  给出关于事件  $x_i$  的信息量定义为

互信息量, 表达式:

$$\begin{aligned} \text{互信息量 } I(x_i; y_j) &= \log \frac{P(x_i; y_j)}{P(x_i)} = \log P(x_i; y_j) - \log P(x_i) \\ &= \underline{I(x_i) - I(x_i; y_i)} \end{aligned}$$

$$I(x_i; y_i) = \log \frac{P(x_i; y_i) P(y_i)}{P(x_i) P(y_i)} = \underline{I(x_i) + I(y_i) - I(x_i; y_i)}$$



性质: ① 对称性 (互易性) ② 互信息量  $\geq 0$  (两者独立) ③ 非负 ④ 不会大于熵

日期: /

## 多变量的互信息

符号  $x_i$  与符号  $y_j z_k$ :  $I(x_i; y_j z_k) = \log \frac{p(x_i; y_j z_k)}{p(x_i; z_k)}$

$$I(x_i; y_j z_k) = I(x_i; z_k) + I(x_i; y_j | z_k)$$

平均互信息量:

$x_i$  求和

1)  $I(x_i; y_j)$  在  $x$  集合上的统计平均值为:

$$I(X; Y) = \sum_i p(x_i; y_j) I(x_i; y_j) = \sum_i p(x_i; y_j) \log \frac{p(x_i; y_j)}{p(x_i)}$$

$y_j$  提供关于集合  $x$  的平均互信息量

$y_j$  求和

2) 上述:  $I(X; Y)$  在集合  $Y$  上的概率加权统计平均值即:

$$I(X; Y) = \sum_j p(y_j) I(X; y_j) = \sum_j p(y_j) \log \frac{p(x_i; y_j)}{p(x_i)}$$

→  $I(X; Y)$  平均互信息量 (交互熵)

$$I(X; Y) = \sum_{i=1}^n \sum_{j=1}^m p(a_i b_j) [\log p(a_i b_j) - \log p(a_i)] = H(X) - H(X|Y)$$

1) 互熵: ① 对称性:  $I(X; Y) = I(Y; X)$  ② 非负性:  $I(X; Y) \geq 0$  ③ 极值性:  $I(X; Y) \leq H(X)$   
 $I(Y; X) \leq H(Y)$

2) 若  $X$  与  $Y$  是一一对应的, 则有  $I(X; Y) = H(X)$

$$H(X|Y) = 0$$

3) 若  $X, Y$  相互独立, 则有  $I(X; Y) = 0$

$$H(X|Y) = H(X)$$

④  $I(X; Y)$  是信源分布  $\{p(a_i)\}$  的上凸函数 ⑤  $I(X; Y)$  是信源符号概率  $\{p(b_j|a_i)\}$  的下凸函数

日期: /

单符号离散信道容量.

信息传输率(信息率)  $R = I(X; Y) = H(X) - H(X|Y)$

信息传输速率  $R_t = \frac{1}{T} I(X; Y)$

$$R = H(X) - H(X|Y) = \int \underbrace{C p(a)}_{\text{信: 分布}} \cdot \underbrace{p(b|a)}_{\text{信道转移概率}} \quad \text{信道转移概率}$$

信道容量, 定义为最大信息传输率, 即:  $C = \max_{p(a)} R = \max_{p(a)} I(X; Y)$   
 $C_t = \frac{1}{T} \max_{p(a)} I(X; Y)$

几种特殊的离散信道

1. 一对一无噪无损信道. 对于  $\forall x_i, y_j$ , 总有  $p(x_i|y_j) = p(y_j|x_i) = 1$

$$C = \max R = \max H(X) = \log n$$

2. 有损无损信道, 一个输入对应多个输出. 对于  $\forall y_j$  则有  $p(x_i|y_j) = 1$  反之不成立.

$$C = \max R = \max H(X) = \log n$$

3. 无损有损信道: 多个输入对应一个输出. 对于  $\forall x_i$ , 则有  $p(y_j|x_i) = 1$  反之不成立.

$$C = \max R = \max H(Y) = \log m$$

日期: /

离散(独对独)信道的容量: 最佳输入概率分布为等概率分布 (均匀分布)

1. 离散独对独输入信道 (行对称)

每一行都是相同元素排列

2. 离散独对独输出信道 (列对称)

每一列都是相同元素排列

3. 对称信道: 行对称 + 列对称

强对称 (均匀信道): ① 行数 = 列数 ② 对于输入符号, 正确传输概率相等

错误传输概率相等且分配到  $(r-1)$  个符号 ③  $r \times r$  行对称矩阵

$$P = \begin{bmatrix} \bar{p} & \frac{p}{r-1} & \dots & \frac{p}{r-1} \\ \frac{p}{r-1} & \bar{p} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p}{r-1} & \frac{p}{r-1} & \dots & \bar{p} \end{bmatrix}$$

强对称信道: ① 不是对称信道 ② 可以按列 (按输出) 划分成对称子矩阵, 那对输出

果  $Y$  进行划分: 划分结果是  $-7m$ , 即对称信道

$$P = \begin{bmatrix} \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \vdots \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{3} & \vdots \\ \frac{1}{5} & \frac{1}{3} & \frac{1}{4} & \vdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

th: 对于对称信道, 若输入概率分布为等概时, 输出概率分布也是等概的

$$C = \max_{p(y)} [H(Y) - H(Y|X)] = \log 5 - H(p_1, p_2, \dots, p_5)$$

信道矩阵中任意一行元素

日期: /

平均信息下信道容量:

$$C = \log r - H(p_1', p_2', p_3', \dots, p_s') \leftarrow \text{对称信道}$$

$$= \log r - H\left(\bar{p}, \frac{p}{r-1}, \frac{p}{r-1}, \dots, \frac{p}{r-1}\right)$$

非对称信道的信道容量

$$C = \log r - H(p_1', p_2', \dots, p_s') - \sum_{k=1}^n V_k \log M_k$$

解法: (注) 对称信道达到信道容量时最佳输入分布为均匀分布

$$p(x_i) = \frac{1}{r} \quad p_Y = \underline{p_X} p_{Y|X} \quad \text{信源熵}$$

$$C = H(Y) - H(Y|X)$$

例. 求  $\begin{bmatrix} \bar{p}-\varepsilon & p-\varepsilon \\ p-\varepsilon & \bar{p}-\varepsilon \end{bmatrix}$  的信道容量.

$$\underline{p_X} = \left[\frac{1}{2}, \frac{1}{2}\right] \quad p_Y = p_X p_{Y|X} = \left[\frac{1}{2}, \frac{1}{2}\right] \begin{bmatrix} \bar{p}-\varepsilon & p-\varepsilon \\ p-\varepsilon & \bar{p}-\varepsilon \end{bmatrix}$$

等概

$$= \left[\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon, 2\varepsilon\right]$$

$$\therefore H(Y) = H\left(\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon, 2\varepsilon\right)$$

$$H(Y|X) = \sum_i \underline{p(x_i)} H(Y|x_i) = \sum_i \frac{1}{2} H(Y|x_i) = H(\bar{p}-\varepsilon, p-\varepsilon, 2\varepsilon)$$

$$C = H(Y) - H(Y|X)$$

等概

日期: /

纠错编码

差错图样  $E$   $E = C - R$  (模  $M$ )

二进制码: 模 2 加:  $E = C \oplus R$

发送码字  $C$  1 1 1 1 | 1 1 1 1 1

接收码字  $R$  1 0 0 1 0 0 | 1 1 1

差错图样  $E$  0 1 1 0 1 1 0 0 0 0

差错元个数叫作 **汉明距离**

→  $k$  个信息元,  $r$  个监督码元, 与前者并置构成  $(k, r, n)$

纠错码分类: **分组码** 和 **卷积码**

$k$  个信息元, 产生  $r$  个监督码元, 组成长度为  $(k+r)$  个码字

纠错编码原理:

纠错原理不仅与信道的统计特性有关, 还与 **译码规则** 有关

正码:  $P(F(b_j) | b_j) = P(a_i | b_j)$

\* 最大

似然误差:  $\sum_{j=1}^m P(b_j) P(e | b_j) = \sum_{j=1}^m P(b_j) [1 - P(a_i | b_j)]$  最小

$P(a^* | b_j) \geq P(a_i | b_j) \quad a_i \neq a^*$

$\frac{P(a^* | b_j)}{P(b_j)} \geq \frac{P(a_i | b_j)}{P(b_j)} \Rightarrow P(b_j | a^*) P(a^*) \geq P(b_j | a_i) P(a_i)$

等根号,  $P(b_j | a^*) \geq P(b_j | a_i)$  最大似然译码规则

$P_e = \sum_{j=1}^m P(b_j | a_i) P(a_i)$

日期: /

输入等概分布. 信源矩阵  $P = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.5 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$

$$F(b_1) = a_1 \quad F(b_2) = a_3 \quad F(b_3) = a_2$$

$$P_i = \frac{1}{3}(0.5 + 0.6 + 0.6) = 0.567$$

输入不等概.  $P(a_1) = \frac{1}{4} \quad P(a_2) = \frac{1}{4} \quad P(a_3) = \frac{1}{2}$

联合信源矩阵  $\begin{bmatrix} 0.15 & 0.075 & 0.05 \\ 0.05 & 0.075 & 0.125 \\ 0.15 & 0.15 & 0.2 \end{bmatrix}$

$$\begin{cases} F(b_1) = a_1 \\ F(b_2) = a_3 \\ F(b_3) = a_2 \end{cases}$$

$$P = [(0.15 + 0.05) + (0.075 + 0.075) + (0.05 + 0.125)] = 0.5$$

香农公式:  $H(X|Y) \leq H(P_e) + P_e \log(1/P_e)$   
 信息定义

错误概率与编码方法: 简单重复编码. 信息传输率  $R = \frac{\log M}{n}$

信道编码定理: 若离散系统在记忆信道, 容量  $C$ , 只要信源的熵率

$R < C$ , 总能找到一种编码. 当  $n$  足够长时, 译码差错概率  $P_e < \epsilon$ .

已知:  $R < C$ .  $P_e$  为大于 0

$\alpha_i = (a_{i1}, a_{i2}, \dots, a_{in})$ ,  $\beta_j = (b_{j1}, b_{j2}, \dots, b_{jn})$  为两  $n$  元组, 则

$\alpha_i$  和  $\beta_j$  之间的汉明距离

$$D(\alpha_i, \beta_j) = \sum_{k=1}^n a_{ik} \oplus b_{jk}$$

$$d_{\min} = \min \{D(\alpha_i, \alpha_j) \mid \alpha_i \neq \alpha_j\}$$

任意两码字之间的汉明距离的最小值称为码的最小距离

日期: /

th. 若以错码数为能力  $d_{min}$

1. 可以检测出  $(= d_{min} - 1)$  个差错

2. 可以纠正任意小于等于  $t = \lfloor \frac{d_{min} - 1}{2} \rfloor$  个差错

3. 可以检测出任意小于等于  $l$  同时纠正小于等于  $t$  个差错, 其中  $l, t$  满足,

$$l + t \leq d_{min} - 1, \quad t < l$$

线性码的基本概念

信道编码译码:  $\underbrace{\text{二元码信息位} \quad \text{监督位}}_{\text{码字}}$

汉明重量: 码字中非零码元的个数, 简称码重,

总线性码的基本参数如下: 码长  $n$  信息长  $k$  码字数  $M$  监督位  $r$   
最小码距  $d_{min}$  编码效率  $R = k/n$

公式:  $C = MG$   $M$ : 消息, 大小为  $k$  即  $1 \times k$  的行向量

$G$ : 生成矩阵, 大小为  $k \times n$

$C$ : 生成的码字, 大小为  $n$ , 即  $1 \times n$  的行向量

例: 循环码 (4,1) 线性码, 其生成矩阵为  $G = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$ ,  $\rightarrow M = (m_0, m_1, m_2) = (1, 0)$

求码字  $C = (C_0, C_1, C_2, C_3) = (1, 0, 1, 0) \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = (1, 0, 1, 0)$

日期: /

对于给定的线性分组码有以下性质:

- ① 零向量  $\mathbf{0}$  一定是一个码字, 称做零码字 ② 任意两个码字的和或差是一个码字 (线性)
- ③ 任意码字  $\mathbf{c}$  是  $G$  行的向量  $g_1, g_2, \dots, g_{n-k}$  的线性组合 ④ 最小距离即为最小非零码字的重量

线性分组码的译码

对于  $k \times n$  矩阵  $G$ , 存在  $(n-k) \times n$  矩阵  $H = \begin{bmatrix} h_{1,0} & \dots & h_{1,n-1} \\ \vdots & \ddots & \vdots \\ h_{n-k,0} & \dots & h_{n-k,n-1} \end{bmatrix}$ .

使得  $GH^T = [\mathbf{0}]_{k \times (n-k)}$

$H$ : 校验矩阵  $H^T$ : 转置  $CH^T = MGH^T = M[\mathbf{0}]_{k \times (n-k)} = \mathbf{0}$  ( $(n-k)$  维零向量)

生成矩阵  $G = G_s = [I_k \quad Q_{k \times (n-k)}]$ , (系统码)  $I_k = k$  阶单位阵

则校验矩阵  $H = H_s = [Q_{k \times (n-k)}^T \quad I_{(n-k)}]$

$G_s$  和  $H_s$  仍然满足  $G_s H_s^T = [\mathbf{0}]_{k \times (n-k)}$

注:  $G$  和  $H$  的角色可以互换, 即  $H$  可作为生成矩阵,  $G$  可作为校验矩阵.

由  $H$  生成码字作为  $C$  的伴随码

例: 一个 (5,3) 线性分组码的生成矩阵  $G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$ , 求  $G_s$  和  $H_s$ .

$$G_s = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad H_s = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

$\begin{matrix} \swarrow & \searrow & \swarrow & \searrow & \swarrow & \searrow \\ I & & Q & & Q^T & I \end{matrix}$

由  $G \rightarrow G_s$  可用行变换: 交换行, 将一行加到另一行, 将某一行乘一个常数

日期: /

若信道无干扰, 接收向量可表示为码矢之和差错的线性叠加

接收向量:  $r = (r_0, r_1, \dots, r_{n-1})$  码矢:  $c = (c_0, c_1, \dots, c_{n-1})$

错误同样:  $e = (e_0, e_1, \dots, e_{n-1})$  伴随式  $s = (s_0, s_1, \dots, s_{n-k-1})$

$$r = c \oplus e = (c_0 \oplus e_0, c_1 \oplus e_1, \dots, c_{n-1} \oplus e_{n-1})$$

$$s = rH^T \quad s \neq 0 \rightarrow \text{一定有错误} \quad s = 0 \rightarrow \text{无差错/错误同样为0} \quad \text{码字}$$

译码步骤:

1. 计算  $s$ . 构造  $(s-e)$  对接收向量上计算伴随式  $s = rH^T$

3. 查  $(s, e)$  得  $e$  4. 纠错  $\hat{c} = r \oplus e$

例: (6,3) 线性分组码. 生成矩阵  $G = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$   $r = (111001)$ , 对其译码

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \quad \text{由 } s = eH^T, \text{ 构造 } (s-e) \text{ 表}$$

$e$     000000    100000    010000    001000    000100    000010    000001

$$H_s^T = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$s$ :    000    110    011    101    100    010    001

$$s = rH_s^T = (001) \text{ 对于 } e = (0000001) \quad \hat{c} = r \oplus e = 111000$$

$m = (111)$ . 将校验码对照表来查找对应消息

日期: /

多码位的汉明码:

码长:  $n=2^r-1$

监督位长:  $r=n-k$

信息位长:  $k=n-r=2^r-r-1$  最小码距:  $d_{\min}=3$

码字数:  $M=2^k$

纠错能力  $t=1$

信息率函数

设无记忆离散信源  $(X) = \{a_1, a_2, \dots, a_n\}$   
 $p(x_i) = \{p(a_1), p(a_2), \dots, p(a_n)\}$

接收端  $(Y) = \{b_1, b_2, \dots, b_n\}$   
 $p(y_j) = \{p(b_1), p(b_2), \dots, p(b_n)\}$

信道转移概率:  $p(Y|X) = \begin{bmatrix} p(b_1|a_1) & p(b_2|a_1) & \dots & p(b_n|a_1) \\ p(b_1|a_2) & p(b_2|a_2) & \dots & p(b_n|a_2) \\ \vdots & \vdots & \ddots & \vdots \\ p(b_1|a_n) & p(b_2|a_n) & \dots & p(b_n|a_n) \end{bmatrix}$

数对  $(a_i, b_j)$  指定一个非负实数, 称为失真函数/失真度

$d(a_i, b_j) \geq 0$ . 矩阵形式  $[D] = \begin{bmatrix} d(a_1, b_1) & d(a_1, b_2) & \dots & d(a_1, b_n) \\ d(a_2, b_1) & d(a_2, b_2) & \dots & d(a_2, b_n) \\ \vdots & \vdots & \ddots & \vdots \\ d(a_n, b_1) & d(a_n, b_2) & \dots & d(a_n, b_n) \end{bmatrix}$

常见的失真函数:  $d(a_i, b_j) = \begin{cases} 0 & a_i = b_j \\ \alpha & a_i \neq b_j \end{cases}$

$\alpha=1 \rightarrow$  汉明失真/误差矩阵失真

日期: /

② 平方误差失真函数  $d(a_i, b_j) = (b_j - a_i)^2$

$$\bar{D} = E[d(a_i, b_j)] = \sum_{i=1}^n \sum_{j=1}^n \underbrace{p(a_i) p(b_j | a_i)}_{\text{联合概率}} d(a_i, b_j)$$

保真度准则  $\bar{D} \leq D \leftarrow$  允许失真上界。

序列  $x^n$  与  $y^n$  失真度表示为:

$$d(x^n, y^n) = \sum_{i=1}^n d(x_i, y_i)$$

$N$  维信源符号序列的平均失真度, 无论何时  $\bar{D}(N) = \frac{1}{N} \bar{D}$

信源平均失真度(单符号):  $\bar{D}_k = \frac{1}{N} \sum_{i=1}^N \bar{D}_i$

$$\bar{D}_k = \bar{D} \quad \bar{D}(N) = N \bar{D}$$

信息率失真函数

$\bar{D} < D$  取1/2. 让平均互信息量  $I(x; y)$  最小

率失真函数  $R(D) = \min_{p(a_i | b_j) \in P_0} I(x; y)$  无论何  $N$  次扩展:  $R_N(D) = NR(D)$

性质:

1)  $D=0$  (不允许失真),  $R(0) = H(x)$

$D = D_{\max}$  时,  $R(D) = 0$  且  $D > D_{\max}$  时,  $R(D)$  恒等于0.

2) 凸函数.

3) 单调递减/连续:

$D$  增大, 最低码率  $R(D)$  下降, 并且连续

日期: /

对  $D_{min}$  的讨论:

$$D_{min} = \min \left[ \sum_x \sum_y P(x_i) P(y_j | x_i) d(x_i, y_j) \right] = \left( \sum_i P(x_i) \right) \min \left[ \sum_j P(y_j | x_i) d(x_i, y_j) \right]$$

由信源引起  
每一项乘以

即: 构造试验信道满足:

$$\sum_j P(y_j | x_i) = 1 \quad d(x_i, y_j) \text{ 为失真矩阵每一行的最小值}$$

$$D_{min} = \sum_i P(x_i) \min d(x_i, y_j)$$

对  $D_{max}$  的讨论:

$D_{max}$  是满足信息传输率  $R(D) = 0$  的最小失真, 即对  $D > D_{max}$  有  $R(D) = 0$  成立

$\therefore I(X; Y) = 0$  即  $X$  与  $Y$  统计独立, 故  $P(y_j | x_i) = P(y_j)$

$$D_{max} = \min \sum_{i,j} P(x_i) P(y_j | x_i) d(x_i, y_j) \\ = \min \sum_{i,j} P(y_j | x_i) \left[ \sum_i P(x_i) d(x_i, y_j) \right] \text{ 最小值}$$

对于固定的  $x_i$ , 则  $\sum_j P(y_j | x_i) = 1$

$$\text{否则 } \sum_j P(y_j | x_i) = 0$$

例: 信道输入为  $X = \{a_1, a_2, a_3\}$ , 且为等概率分布, 已知失真矩阵为  $D = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$

试求  $D_{min}$  和  $D_{max}$  以及相应试验信道的转移概率矩阵:  $D = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$

$$D_{min} = \sum_i P(x_i) \min d(x_i, y_j)$$

$$= \frac{1}{3} \min(1, 2, 2) + \frac{1}{3} \min(2, 1, 2) + \frac{1}{3} \min(2, 2, 1)$$

$$= 1 \quad D = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$$

$$D_{max} = \sum_i \min_j P(x_i) d(x_i, y_j)$$

$$= \begin{bmatrix} P(a_1)x_1 & P(a_2)x_2 & P(a_3)x_3 \\ P(a_1)x_2 & P(a_2)x_1 & P(a_3)x_2 \\ P(a_1)x_3 & P(a_2)x_2 & P(a_3)x_1 \end{bmatrix}$$

$$P(X) = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$

并求每一行  
再按此列求和取最大值

$$= \frac{5}{3}$$

日期:

/

对于  $n$  元等概率信源的信息率失真函数为

$$R(D) = \ln n + \frac{D}{\alpha} \left( \ln \frac{D}{\alpha} + (1 - \frac{D}{\alpha}) \ln (1 - \frac{D}{\alpha}) \right)$$

$n$ :  $n$  元  $\alpha$ : 失真矩阵中  $d(x_i|y_i)$

$$R(W) = H(X) - H(X|\hat{X})$$

$p = \frac{D}{\alpha} \rightarrow$  1-p 还原正确  
 $p$  还原错误且数据在  $k-1$  个符号中.

$$H(X|\hat{X}) = \underbrace{\ln(p)}_{\text{对 0 个符号}} + \underbrace{p \ln(n-1)}_{\text{错那一个符号}}$$

熵:  $H(X) = \ln n$

$$\therefore R(D) = H(X) - H(X|\hat{X})$$

$$= \ln n + p \ln p + (1-p) \ln (1-p) - p \ln (n-1)$$

$$= \ln n + (1-p) \ln (1-p) + p \ln \frac{p}{n-1}$$

$$= \ln n + (1 - \frac{D}{\alpha}) \ln (1 - \frac{D}{\alpha}) + \frac{D}{\alpha} \ln \frac{\frac{D}{\alpha}}{n-1}$$