

# 自信息量

---

武汉理工大学

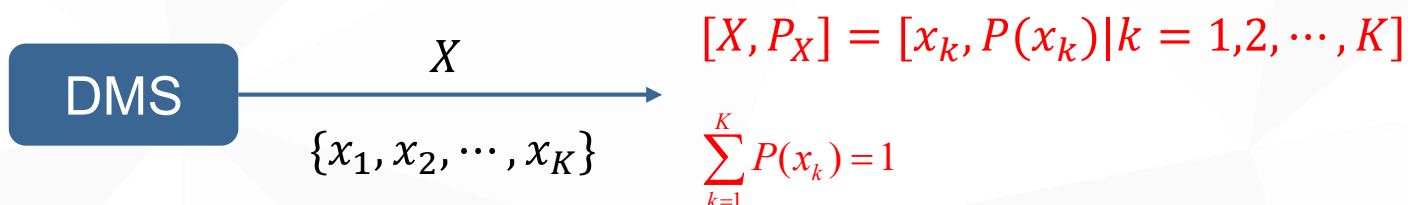
Information theory  
and  
coding



武汉理工大学

## »» 自信息量

---



$I(x_k)$  :  $x_k$  的 (先验) 不确定性, 也称为  $x_k$  的自信息量。

$$I(x_k) = \log \frac{1}{P(x_k)} = -\log P(x_k) \quad k = 1, 2, \dots, K$$

注：自信息量与信息有联系，但不是信息，而是符号的先验不确定性。

## »» 自信息量



例1：某地二月份天气的概率分布统计如下：

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{bmatrix} a_1(\text{晴}), a_2(\text{阴}), a_3(\text{雨}), a_4(\text{雪}) \\ \frac{1}{2}, \quad \frac{1}{4}, \quad \frac{1}{8}, \quad \frac{1}{8} \end{bmatrix}$$

这四种气候的自信息量分别为：

$$\begin{aligned} I(a_1) &= 1 \text{ bit}, & I(a_2) &= 2 \text{ bit}, \\ I(a_3) &= 3 \text{ bit}, & I(a_4) &= 3 \text{ bit}. \end{aligned}$$

## »» 自信息量的单位

自信息量的单位与公式中**对数底**的选取有关。

$$I(x_k) = \log \frac{1}{P(x_k)} = -\log P(x_k)$$

对数底	对数符号	单位
2	$-\log P(x_k)$	二进制单位, 比特 (bit, binary digit)
e	$-\ln P(x_k)$	自然单位, 奈特 (nat, natural digit)
10	$-\lg P(x_k)$	十进制单位, 迪特 (dit, decimal digit的缩写), 也可用哈特 (Hart)
正整数r	$-\log_r P(x_k)$	r进制单位

## »» 自信息量的单位

---

### 单位换算

$$1\text{bit} = \ln 2 \text{ nat} = \lg 2 \text{ dit} = \log_r 2 \text{ r 进制单位}$$

为了强调是符号的不确定性，我们将单位写成：

- bit/符号
- nat/符号
- dit/符号
- r进制单位/符号

## »» 自信息量单位的物理含义说明



例2: 随机变量  $X = \{x_1, x_2, x_3, x_4\}$  , 各符号的概率相等,  
则各符号的自信息量相等:

$$\begin{aligned} I(x_1) = I(x_2) = I(x_3) = I(x_4) &= -\log \frac{1}{4} = 2 \quad \text{bit/符号} \\ &= -\log_4 \frac{1}{4} = 1 \quad \text{四进制单位/符号} \end{aligned}$$

$x_k$  的不确定性可用2位二进制数字来度量(00 01 10 11),  
或1位四进制数字来度量(0 1 2 3)。

意义

The background of the slide features a complex network diagram. It consists of numerous nodes, represented by small circles in various shades of blue and grey, interconnected by thin, light blue lines. The network is dense and spans the entire width and height of the slide, creating a subtle, technical backdrop.

---

# 感谢观看！

---

Information theory  
and  
coding



武汉理工大学