

自信息量的性质

武汉理工大学

Information theory
and
coding



武汉理工大学

»» 自信息量的性质和相互关系

$$I(x_k) = -\log P(x_k)$$

$$I(x_k, y_j) = -\log P(x_k, y_j)$$

$$I(x_k|y_j) = -\log P(x_k|y_j), I(y_j|x_k) = -\log P(y_j|x_k)$$



(1) 概率为0时，相应的自信息量无意义。

(2) 非负性。三种自信息量均非负。

»» 自信息量的性质和相互关系

公式

$$\begin{aligned} I(x_k) &= -\log P(x_k) & , & & I(x_k, y_j) &= -\log P(x_k, y_j) \\ I(x_k|y_j) &= -\log P(x_k|y_j) & , & & I(y_j|x_k) &= -\log P(y_j|x_k) \end{aligned}$$

联合概率、条件概率和边缘概率之间的乘法关系：

$$P(x_k, y_j) = P(x_k)P(y_j|x_k) = P(y_j)P(x_k|y_j)$$

自信息量的可加性：

$$\begin{aligned} I(x_k, y_j) &= I(x_k) + I(y_j|x_k) \\ &= I(y_j) + I(x_k|y_j) \end{aligned}$$

»» 可加性的物理解释

自信息量的可加性:
$$I(x_k, y_j) = I(x_k) + I(y_j|x_k)$$
$$= I(y_j) + I(x_k|y_j)$$

解释一

若 (x_k, y_j) 为信源发出的符号串, 则 (x_k, y_j) 的不确定性

(x_k, y_j) 的联合不确定性 $I(x_k, y_j)$

=符号 x_k 的不确定性 $I(x_k)$ +符号 y_j 在 x_k 出现的条件下的不确定性 $I(y_j|x_k)$

=符号 y_j 的不确定性 $I(y_j)$ +符号 x_k 在 y_j 出现的条件下的不确定性 $I(x_k|y_j)$

»» 可加性的物理解释

自信息量的可加性:

$$\begin{aligned} I(x_k, y_j) &= I(x_k) + I(y_j|x_k) \\ &= I(y_j) + I(x_k|y_j) \end{aligned}$$

解释二

若针对非理想观察模型, x_k 是观察输入符号, y_j 是观察输出符号。观察系统的不确定性来自于: **输入和随机干扰。**

(x_k, y_j) 的联合不确定性 $I(x_k, y_j)$

= 输入的不确定性 $I(x_k)$ + 干扰引入的不确定性 $I(y_j|x_k)$

= 输出的不确定性 $I(y_j)$ + 观察到 y_j 后还剩余的不确定性 $I(x_k|y_j)$

»» 自信息量相互关系推广

自信息量的可加性:

$$\begin{aligned} I(x_k, y_j) &= I(x_k) + I(y_j|x_k) \\ &= I(y_j) + I(x_k|y_j) \end{aligned}$$



推广到多维空间

自信息量可加性的链公式:

$$I(u_1, u_2, \dots, u_N) = I(u_1) + I(u_2|u_1) + I(u_3|u_1u_2) + \dots + I(u_N|u_1u_2 \dots u_{N-1})$$

»» 特殊情况下自信息量相互关系

公式

$$\begin{aligned} I(x_k) &= -\log P(x_k) & , & & I(x_k, y_j) &= -\log P(x_k, y_j) \\ I(x_k|y_j) &= -\log P(x_k|y_j) & , & & I(y_j|x_k) &= -\log P(y_j|x_k) \end{aligned}$$

当 x_k 和 y_j 统计独立时, 概率之间的乘法关系: $P(x_k, y_j) = P(x_k)P(y_j)$

自信息量的可加性: $I(x_k, y_j) = I(x_k) + I(y_j)$

可加性的链公式: $I(u_1, u_2, \dots, u_N) = I(u_1) + I(u_2) + \dots + I(u_N)$

DMS符号串不确定性等于各个独立符号的不确定性之和。

»» 特殊情况下自信息量相互关系



例题：设离散无记忆信源 $\begin{bmatrix} a_1 = 0 & a_2 = 1 & a_3 = 2 & a_4 = 3 \\ 3/8 & 1/4 & 1/4 & 1/8 \end{bmatrix}$

其发出的消息为202 120 130 213 001 203 210 110 321 010 021 032 011 223 210，求该消息的自信息量以及消息中平均每个符号的自信息量。

解
$$\begin{aligned} I(\bar{u}) &= 14I(a_1) + 13I(a_2) + 12I(a_3) + 6I(a_4) \\ &= 14 \times \log \frac{8}{3} + 13 \times \log 4 + 12 \times \log 4 + 6 \times \log 8 \\ &= 87.81 \text{ bit/消息} \end{aligned}$$

$$\frac{I(\bar{u})}{45} = 87.81/45 = 1.95 \text{ bit/符号}$$

思考

如果消息长度为无穷大，则该平均自信息量该如何计算？

A background network diagram consisting of numerous nodes (dots) connected by thin lines, forming a complex web. The nodes are colored in shades of blue and grey, and the lines are thin and light blue. The overall shape of the network is roughly triangular, with the base at the bottom and the apex at the top.

感谢观看！

Information theory
and
coding



武汉理工大学