Mathematical Modeling

数学建模

同济大学数学科学学院

第十五章

信息论模型

目录/Contents ▌第十五章 信息论模型

一司机开车，被警察抓住，非说他超速，并责令交出驾驶证。

司机：我没有驾驶证！

警察：没有驾驶证，还开车？这车是谁的？

司机：抢来的！

警察(警惕)：抢谁的？

司机：我哪知道啊，人都杀了也没办法问了！

警察(急急呼叫总台后)：杀的人在哪了？

司机：扔后备箱了！

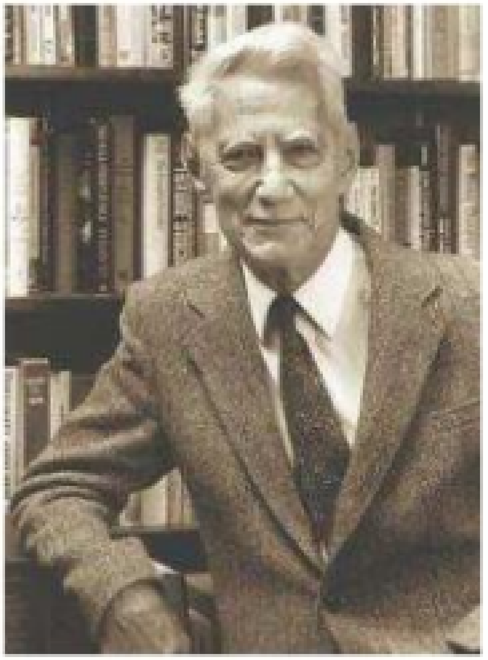
警察命其下车，并等待其他警察... 警察大队来了，带队的警察命令搜车，但是搜查了半天也没找到死尸，于是就问。

警察：死人呢？

司机：没有啊。

另一名警察：那他报告说你杀人了？

司机(感叹)：咳... 他还说我超速呢！



Claude E. Shannon (香农)

1916.4.30 -- 2001.2.24

信息论奠基人

In his words “*I just wonder how things put together*."

问题: 如何度量信息量?

- 他不坐在前3排, 也不坐在后3排。
- 他坐在第5排。
- 他刚才被查到了无证驾驶。
- 他在后备箱藏了尸体。

目录/Contents ▍第十五章 信息论模型

- 信息量是事件发生概率的连续函数;
- 如果事件A发生必有事件B发生，则得知事件A发生的信息量不小于得知事件B发生的信息量;
- 如果事件A,B的发生是相互独立的，则获知两事件同时发生的信息量为单独获知两事件发生的信息量的和;
- 任何信息的信息量都是有限的。

定理1 (Shanno) 记事件 M 发生的概率为 p，满足四条公理的获知 M 发生的信息量计算公式为 $I(M)=-C\log_a p,\ a>1.$

证明： 由公理1，$I(M)=f(p), f$ 是连续函数。

由公理2，A发生必有B发生，$p_A \le p_B, f(p_A)\ge f(p_B), f$ 单调不增。

若A，B独立，由公理3，$f(p_A p_B)=f(p_A)+f(p_B)$。记 $p=a^{-q}, f(a^{-q})=g(q)$，则

$$g(q_A+q_B)=g(q_A)+g(q_B)$$

(1) $g(0)=0$;
(2) 若 $g(1)=C, g(n)=nC,\ g(1/n)=C/n,\ g(m/n)=mC/n,\ g(x)=xC$;
(3) 若 $x<0$, 因$g(x)+g(-x)=0$, 因此对所有x，$g(x)=xC$;

$$I(M)=f(p)=-C\log_a p.$$

- $a=2,\ C=1$, 信息量单位称为比特(bit);
- $a=10,\ C=1$, 信息量单位称为迪吉特(dit);
- $a=\mathrm{e},\ C=1$, 信息量单位称为奈特(nat);

例:

- 随机抛硬币, 结果是正面。
- 你的学号个位数是3。
- 把大家交上来的数学建模论文随机发回, 每个人都没有拿到自己的.

某剧院有1280个座位，分为32排，每排40座。
欲从中找出一人，求以下信息的信息量。

i) 他在第10排；ii) 他在第15座；iii) 他在10排15座。

$$-\log_2\frac{1}{32}=5,\quad -\log_2\frac{1}{40}=5.32,\quad -\log_2\frac{1}{1280}=10.32$$

得知 N 个等概率事件的某一个发生的信息量为

$$-\log_2\frac{1}{N}=\log_2 N$$

目录/Contents ▌第十五章 信息论模型

某一实验按照概率 $p_1, p_2, \cdots, p_N$ 出现 N 种结果, 得知第 i 种结果的信息量为 $-\log_2 p_i$, 实验的不确定性可由平均信息量(熵, entropy)度量:

$$H = -\sum_{i=1}^{N} p_i \log_2 p_i$$

抽一张扑克牌, 出现某张牌的概率各为1/54, 故熵 $H = \log_2 54 = 5.75$

投掷一枚骰子的结果有6种, 即1--6点, 每一种概率为1/6, 熵 $H = \log_2 6 = 2.585$

投掷一枚硬币的结果有2种, 即正反面, 每一种概率为1/2, 熵 $H = \log_2 2 = 1$

在石头上投掷一个鸡蛋, 只有一个结果(鸡蛋摔破), 熵 $H = \log_2 1 = 0$

熵反映结果的模糊度, 熵越大事情越模糊。

当 $p_1 = p_2 = \cdots = p_N$ 时, 事件具有最大熵。

问题：有12枚外表相同的硬币，已知其中有一个是假的，可能轻些也可能重些。用一个没有砝码的天平几次能够找出假币？

每枚硬币都可能是假的，可能轻也可能重，总计有24种情况。

确定是哪一种结果的信息量是 $-\log_2 \frac{1}{24} = \log_2 24.$

每一次把硬币放上天平，可以得出三种不同的结果，信息量有 $-\log_2 \frac{1}{3} = \log_2 3.$

若最少需要称 k 次，则 $k \log_2 3 \geq \log_2 24$

因此，$k \geq 3.$

5,7,9,11	6,8,10,12
2,9,10,12	3,4,8,11
1,4,11,12	3,6,7,9

目录/Contents ▌第十五章 信息论模型

人类活动中，大量信息是通过文字或者语言来表达的，它们是一串符号的组合。可以通过计算熵给出一种语言的每一个符号的平均信息量。

例如，英文

符号	概率	符号	概率	符号	概率	符号	概率
空格	0.2	R	0.0054	U	0.0225	B	0.005
E	0.105	S	0.052	M	0.021	K	0.003
T	0.072	H	0.047	P	0.0175	X	0.002
O	0.0654	D	0.035	Y	0.012	J	0.001
I	0.065	L	0.029	W	0.012	Q	0.001
A	0.063	C	0.023	G	0.011	Z	0.001
N	0.059	F	0.0225	V	0.008		

➢ 汉字的高频字

的一了是我　不在人们有　来他这上着　个地到大里　说就去子得
也和那要下　看天时过出　小么起你都　把好还多没　为又可家学
只以主会样　年想能生同　老中十从自　面前头道它　后然走很像
见两用她国　动进成回什　边作对开而　己些现山民　侯经发工向

…… ……

一个反例

XX

英文中，每个符号含比特

$$H = -\sum_{i=1}^{27} p_i \log_2 p_i = 4.3$$

对于有27个符号得信息源，每个符号得最大信息可达

$$H_{\max} = \log_2 27 = 4.75$$

因此，英文表达的冗余度为

$$\frac{H_{\max} - H}{H_{\max}} = 0.094 = 9.4\%$$

验证：Q后面总是U，T后面很可能是H

如何为一些抽象的概念建模?

比如：满意程度，复杂程度，稳定程度?

阅读：福尔摩斯探案集之跳舞的小人

Mathematical Modeling

数学建模

同济大学数学科学学院

学海无涯，祝你成功！

人民邮电出版社
POSTS & TELECOM PRESS