

# 熵权法

Entropy Evaluation Method

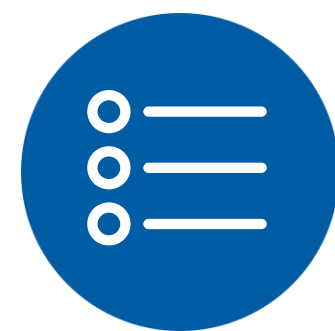
成都信息工程大学学生数学建模协会

牛右刀薛面csdn评价算法——熵权法  
“随机一个昵称” B站“信息熵详解”



# 目录

## CONTENTS



01 /

方法介绍



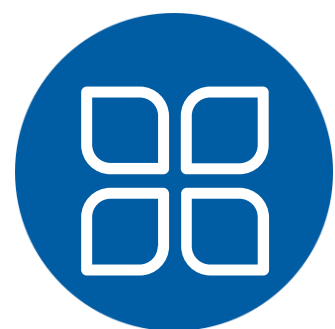
02 /

例题讲解



03 /

代码实现



04 /

章末总结



# 01 / 方法介绍

Introduction



熵权法



进入词条

搜索

首页

探索百科

用户

任务

合作

知识商城

帮助中心

# 熵权法

免费编辑

添加义项名

+ 添加义项 ?

所属类别：图书

熵权法，物理学名词，按照信息论基本原理的解释，信息是系统有序程度的一个度量，熵是系统无序程度的一个度量;根据信息熵的定义，对于某项指标，可以用熵值来判断某个指标的离散程度，其信息熵值越小，指标的离散程度越大，该指标对综合评价的影响(即权重)就越大，如果某项指标的值全部相等，则该指标在综合评价中不起作用。因此，可利用信息熵这个工具，计算出各个指标的权重，为多指标综合评价提供依据。



# 熵权法

研究：评价类问题

依据：利用信息熵计算各个指标的权重

目的：为多指标综合评价提供依据。



# 信息熵

信息论之父香农提出“信息熵”的概念，完成了对信息的量化。

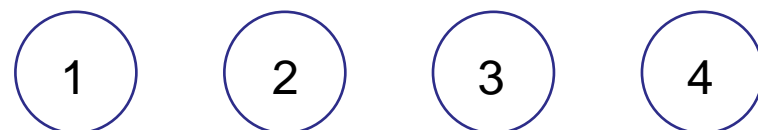


信息熵计算公式：

$$H(p) = - \sum_{i=1}^n p_i \log_2 p_i$$

信息的基本作用就是消除人们对事物了解的不确定性。

例：假设有四个大小形状完全相同的球，那么随机取出一个球，取出这个球是什么编号？



这是一个随机变量，这件事情有一定的不确定性，要消除这个不确定性，就需要引入信息，这些信息可以用“信息熵”量化。需要引入信息量越多，则信息熵越高，反之则越低。





**例：**假设有四个大小形状完全相同的球，那么随机取出一个球，取出这个球是什么编号？

现在我们想知道这个随机事件的信息量。

**信息熵计算公式：**

$$H(p) = - \sum_{i=1}^n p_i \log_2 p_i$$

根据以上的计算公式

那么上述随机变量（取出哪个编号的球）的信息量是：

$$H(p) = -(p_1 \log_2 p_1 + p_2 \log_2 p_2 + p_3 \log_2 p_3 + p_4 \log_2 p_4)$$

结果当抽出每个编号球的概率都相等的时候， $H=2$ 。





# 02 / 例题讲解

Example





## 评价类问题

编号	语文	数学	英语	总分
甲	90	10	60	160
乙	70	30	60	160
丙	10	90	60	160
丁	80	20	60	160

这四个同学谁的综合成绩最好？



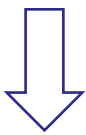
第一步：归一化

消除量纲的影响

$$\frac{x - \min}{\max - \min}$$



编号	语文	数学	英语	总分
甲	90	10	60	160
乙	70	30	60	160
丙	10	90	60	160
丁	80	20	60	160



甲	1	0	1	160
乙	0.75	0.25	1	160
丙	0	1	1	160
丁	0.875	0.125	1	160



## 第二步：计算信息熵

P 概率 在这里应该作为元素在特征列的占比

P的计算方式： $\frac{\text{元素的值}}{\text{一列之和}}$

再利用信息熵公式  $H(p) = - \sum_{i=1}^n p_i \log_2 p_i$

计算每个指标的信息熵，即每一列的信息熵

$$e_j = - \frac{1}{\ln(n)} \sum_{i=1}^n [p_{ij} * \ln(p_{ij})], j = \{1, 2, 3 \dots, m\}$$

编号	语文	数学	英语	总分
甲	90	10	60	160
乙	70	30	60	160
丙	10	90	60	160
丁	80	20	60	160

甲	1	0	1	160
乙	0.75	0.25	1	160
丙	0	1	1	160
丁	0.875	0.125	1	160



### 第三步：确定权重

我们把 $e$ 记作每一项指标的熵值。

则权重的计算方式： $d_j = 1 - e_j$

$$\omega_j = \frac{d_j}{\sum_{j=1}^m d_j}, (j = 1, 2, 3, \dots, m)$$

编号	语文	数学	英语	总分
甲	90	10	60	160
乙	70	30	60	160
丙	10	90	60	160
丁	80	20	60	160



03 /

代码实现

Code Implementation





04 /

# 章末总结

Summary



## 总结

熵权法：评价类问题

优点：

- ① 精度更高，客观性更强。
- ② 算法应用简单。
- ③ 熵权法可以反应指标的区分能力，确定较好的权重

缺点：

- ① 对样本的依赖性大。建模样本变化，权重也会随之变化。
- ② 无法考虑到指标与指标之间的横向影响。

归一化

计算信息熵

计算权重



# 感谢聆听