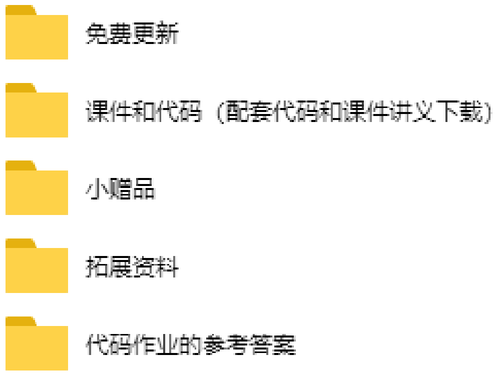
温馨提示

（1）视频中提到的附件可在**售后群的群文件**中下载。

包括**讲义、代码、优秀的作业、我视频中推荐的资料**等。



（2）关注我的**微信公众号《数学建模学习交流》**，后台发送**“软件”**两个字，可获得常见的建模软件下载方法；发送**“数据”**两个字，可获得建模数据的获取方法；发送**“画图”**两个字，可获得数学建模中常见的画图方法。另外，也可以看看公众号的历史文章，里面发布的都是对大家有帮助的技巧。

（3）**购买更多优质精选的数学建模资料**，可关注我的微信公众号《数学建模学习交流》，在后台发送**“买”**这个字即可进入店铺进行购买。

（4）视频价格不贵，但价值很高。单人购买观看只需要**58元**，和另外两名队友一起购买人均仅需**46元**，视频本身也是下载到本地观看的，所以请大家**不要侵犯知识产权**，对视频或者资料进行二次销售。

数学建模学习交流

基于熵权法对Topsis模型的修正

有n个要评价的对象，m个评价指标的标准化矩阵：

$$Z=\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nm} \end{bmatrix}$$

可以使用层次分析法给这m个评价指标确定权重：

$$\sum_{j=1}^{m}\omega_j=1$$

层次分析法最大的缺点

判断矩阵的确定依赖于专家，如果专家的判断存在主观性的话，会对结果产生很大的影响。

（主观性太强）

熵权法是一种客观赋权方法

依据的原理：指标的变异程度越小，所反映的信息量也越少，其对应的权值也应该越低。（客观 = 数据本身就可以告诉我们权重）

（一种极端的例子：对于所有的样本而言，这个指标都是相同的数值，那么我们可认为这个指标的权值为0，即这个指标对于我们的评价起不到任何帮助）

如何度量信息量的大小

小张和小王是两个高中生。小张学习很差，而小王是全校前几名的尖子生。

高考结束后，小张和小王都考上了清华。小王考上了清华，大家都会觉得很正常，里面没什么信息量，因为学习好上清华，天经地义，本来就应该如此的事情。

然鹅，如果是小张考上了清华，这就不一样了，这里面包含的信息量就非常大。怎么说？因为小张学习那么差，怎么会考上清华呢？把不可能的事情变成可能，这里面就有很多信息量。

注：本例子来自微信公众号："小宇治水"

上面的小例子告诉我们：

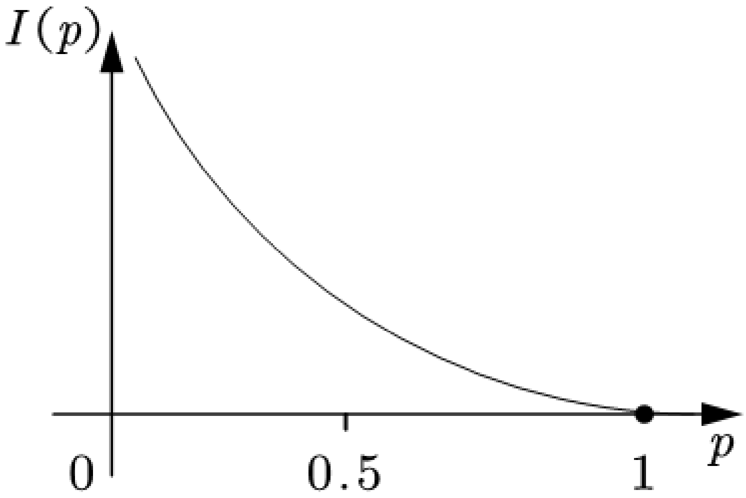
越有可能发生的事情，信息量越少，
越不可能发生的事情，信息量就越多。

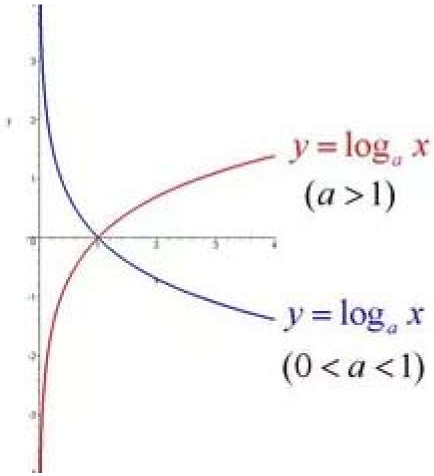
怎么衡量事情发生的可能性大小？

概率

如何度量信息量的大小

如果把信息量用字母I表示，概率用p表示，那么我们可以将它们建立一个函数关系：





假设x表示事件X可能发生的某种情况，

$p(x)$表示这种情况发生的概率

我们可以定义：$I(x) = -\ln(p(x))$

因为$0 \le p(x) \le 1$，所以$I(x) \ge 0$

信息熵的定义

假设x表示事件X可能发生的某种情况，$p(x)$表示这种情况发生的概率

我们可以定义：$I(x)=-\ln(p(x))$，因为$0\le p(x)\le 1$，所以$I(x)\ge 0$

如果事件X可能发生的情况分别为：$x_1,x_2,\cdots,x_n$

那么我们可以定义事件X的信息熵为：

$$H(X)=\sum_{i=1}^{n}[p(x_i)I(x_i)]=-\sum_{i=1}^{n}[p(x_i)\ln(p(x_i))]$$

从上面的公式可以看出，信息熵的本质就是对信息量的期望值。

可以证明的是：

当$p(x_1)=p(x_2)=\cdots=p(x_n)=\dfrac{1}{n}$时，$H(x)$取最大值，此时$H(x)=\ln n$

熵越大信息量越大还是越小?

知乎：信息熵越大，信息量到底是越大还是越小?
https://www.zhihu.com/question/274997106

有些说：熵越大，不确定性越大，包含的信息越多。

百科和一些资料中说：指标的信息熵越小，提供的信息越大。

还各举出了一些例子，感觉都很有道理。

甚至同一资料描述都相反，例如浅谈信息熵（熵权法的应用） - 不矜不伐的小学生 - 博客园：第四段说"高信息度的信息熵是很低的,低信息度的熵则高。"。而第六段的举例说"如果中国100%夺冠，那么熵是0，相当于没有任何信息。"

到底哪个正确？是我哪里理解错了吗

关注问题　写回答　邀请回答　添加评论　分享　···　收起

9 个回答　　默认排序

傅铁强
战！

3 人赞同了该回答

随机变量的信息熵越大，则它的值（内容）能给你补充的信息量越大，而知道这个值前你已有的信息量越小。

编辑于 2018-04-27

▲ 赞同 3　▼　1 条评论　分享　收藏　感谢

对于熵权法而言，因为我们关注的是已有的信息，所以答案是越小。（后面大家看到计算步骤就会明白）

数学建模学习交流

熵权法的计算步骤

（1）判断输入的矩阵中是否存在负数，如果有则要重新标准化到非负区间
（后面计算概率时需要保证每一个元素为非负数）

假设有n个要评价的对象，m个评价指标（已经正向化了）构成的正向化矩阵如下：

$$X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

那么，对其标准化的矩阵记为Z，Z中的每一个元素：$z_{ij}=x_{ij}\Big/\sqrt{\sum_{i=1}^{n}x_{ij}^2}$

判断Z矩阵中是否存在着负数，如果存在的话，需要对X使用另一种标准化方法

对矩阵X进行一次标准化得到$\tilde{Z}$矩阵，其标准化的公式为:

$$\tilde{z}_{ij}=\frac{x_{ij}-\min\{x_{1j},x_{2j},\cdots,x_{nj}\}}{\max\{x_{1j},x_{2j},\cdots,x_{nj}\}-\min\{x_{1j},x_{2j},\cdots,x_{nj}\}}$$

熵权法的计算步骤

（2）计算第j项指标下第i个样本所占的比重，并将其看作相对熵计算中用到的概率

假设有n个要评价的对象，m个评价指标，且经过了上一步处理得到的非负矩阵为：

$$\tilde{Z}=\begin{bmatrix}\tilde{z}_{11} & \tilde{z}_{12} & \cdots & \tilde{z}_{1m}\\ \tilde{z}_{21} & \tilde{z}_{22} & \cdots & \tilde{z}_{2m}\\ \vdots & \vdots & \ddots & \vdots\\ \tilde{z}_{n1} & \tilde{z}_{n2} & \cdots & \tilde{z}_{nm}\end{bmatrix}$$

我们计算概率矩阵P，其中P中每一个元素p_{ij}的计算公式如下：

$$p_{ij}=\frac{\tilde{z}_{ij}}{\sum_{i=1}^{n}\tilde{z}_{ij}}$$

容易验证：$\sum_{i=1}^{n}p_{ij}=1$，即保证了每一个指标所对应的概率和为1.

熵权法的计算步骤

（3）计算每个指标的信息熵，并计算信息效用值，并归一化得到每个指标的熵权

对于第j个指标而言，其信息熵的计算公式为： $e_j=-\frac{1}{\ln n}\sum_{i=1}^{n}p_{ij}\ln(p_{ij})\ (j=1,2,\cdots,m)$

（1）为什么这里要除以$\ln n$这个常数?

在前面说过，当$p(x_1)=p(x_2)=\cdots=p(x_n)=\frac{1}{n}$时，$H(x)$取最大值，此时$H(x)=\ln n$

这里除以$\ln n$能够使得信息熵的始终位于$[0,1]$区间上面。

（2）e_j越大，即第j个指标的信息熵越大，表明第j个指标的信息越多还是越少?

答案是越少，当$p_{1j}=p_{2j}=\cdots=p_{nj}$时，$e_j=1$, 此时上面定义的信息熵达到最大，

但是，因为$p_{ij}=\tilde{z}_{ij}/\sum_{i=1}^{n}\tilde{z}_{ij}$, 所以$\tilde{z}_{1j}=\tilde{z}_{2j}=\cdots=\tilde{z}_{nj}$, 即所有样本的这个指标值都相同。

信息效用值的定义：$d_j=1-e_j$, 那么信息效用值越大，其对应的信息就越多。

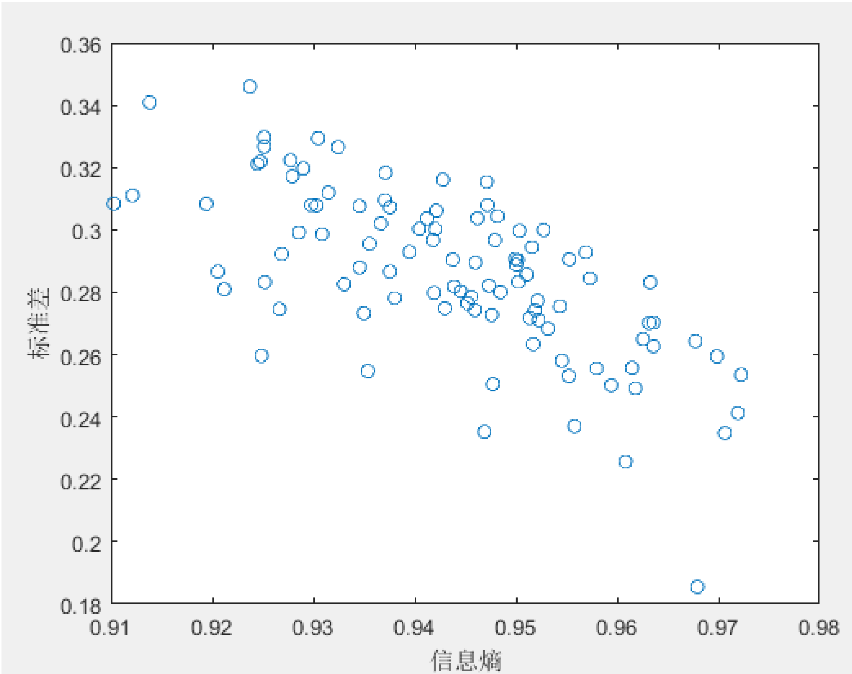
将信息效用值进行归一化，我们就能够得到每个指标的熵权: $W_j=d_j/\sum_{j=1}^{m}d_j\ (j=1,2,\cdots,m)$

熵权法背后的原理

熵权法是一种客观赋权方法

依据的原理：指标的变异程度越小，所反映的信息量也越少，其对应的权值也应该越低。（客观 = 数据本身就可以告诉我们权重）

我们可以用指标的标准差来衡量样本的变异程度，指标的标准差越大，其信息熵越小。



左图是蒙特卡洛的结果
随机生成一组有30个样本且位于区间[0,1]上的数据，计算其信息熵和标准差；将上述步骤重复100次，我们能够得到100组信息熵和标准差的取值，将其绘制成散点图。
可以发现，两个指标之间有很明显的负相关关系。

code_Monte_Carlo.m

熵权法的讨论


希殇
熵权法是不是能降低层次分析法的主观性?
群主清风
别用这个方法
群主清风
这玩意就是忽悠新手的
希殇
呃
星巴克西南交通车辆工程
怎么了
星巴克西南交通车辆工程
。。。


群主清风
你可以理解把这个方法理解为方差大的指标权重就大
群主清风
你觉得这科学吗
希殇
emmmm
群主清风
记住这一句话 不管什么评价指标 实际上都不客观 因为模型的参数是你自己定的
希殇
希殇
哦哦

熵权法的讨论


星巴克西南交通车辆工程
方差大，对总体影响不就大了吗，权就该大一些吧
群主清风
熵权法就是不用我们定权重 样本数据自己的分布决定了权重
群主清风
这不扯淡吗
22:12
群主清风
举个很简单的例子
群主清风
X Y两个指标 用来评定班上谁是三好学生


希殇
我是之前看到一篇论文里这样子写😂
群主清风
X表示大家违纪上档案的次数 Y表示大家逃课的次数
22:14
群主清风
你们觉得哪个对与评定的影响程度大
星巴克西南交通车辆工程
x吧
群主清风
但是几乎所有人X都是0

熵权法的讨论

以上是我之前的看法
现在我给大家一个答复吧
如果大家的论文要发表，别用熵权法
如果大家只是用这个方法进行比赛
那么可以随便用
因为这个方法总比你自己随意定义好

熵权法的代码实现

20条河流的水质情况数据.xlsx	2019/07/11 20:25	Microsoft Excel ...	13 KB
code_Monte_Carlo.m	2019/08/18 18:50	M 文件	1 KB
data_water_quality.mat	2019/06/30 13:08	MATLAB Data	1 KB
Entropy_Method.m	2019/08/18 19:53	M 文件	1 KB
Inter2Max.m	2019/08/07 14:20	M 文件	1 KB
Mid2Max.m	2019/08/07 14:20	M 文件	1 KB
Min2Max.m	2019/08/07 14:20	M 文件	1 KB
mylog.m	2019/08/18 19:00	M 文件	1 KB
Positivization.m	2019/08/07 14:20	M 文件	2 KB
topsis.m	2019/08/18 19:55	M 文件	6 KB

```
function [W] = Entropy_Method(Z)
% 计算有n个样本，m个指标的样本所对应的的熵权
% 输入
% Z :  n*m的矩阵（要经过正向化和标准化处理，且元素中不存在负数）
% 输出
% W：熵权，m*1的行向量
```

运行结果

共有20个评价对象, 4个评价指标
这4个指标是否需要经过正向化处理，需要请输入1 ，不需要输入0： 1
请输入需要正向化处理的指标所在的列，例如第2、3、6三列需要处理，那么你需要输入[2,3,6]： [2,3,4]
请输入需要处理的这些列的指标类型（1：极小型， 2：中间型， 3：区间型）
例如：第2列是极小型，第3列是区间型，第6列是中间型，就输入[1,3,2]： [2,1,3]
第2列是中间型
请输入最佳的那一个值： 7
第2列中间型正向化处理完成
~~~~~~~~~~~~~~~~~~~~~~分界线~~~~~~~~~~~~~~~~~~~~~~
第3列是极小型，正在正向化
第3列极小型正向化处理完成
~~~~~~~~~~~~~~~~~~~~~~分界线~~~~~~~~~~~~~~~~~~~~~~
第4列是区间型
请输入区间的下界： 10
请输入区间的上界： 20
第4列区间型正向化处理完成

正向化后的矩阵 X =

4.6900	0.7172	3.0000	1.0000
2.0300	0.4069	35.0000	0.6940
9.1100	0.5241	8.0000	0.9058
8.6100	0.9655	8.0000	0.4443

标准化矩阵 Z =

0.1622	0.2483	0.0245	0.3065
0.0702	0.1408	0.2863	0.2127
0.3150	0.1814	0.0655	0.2776
0.2977	0.3342	0.0655	0.1361

运行结果

```
请输入是否需要增加权重向量，需要输入1，不需要输入0
请输入是否需要增加权重： 1
使用熵权法确定权重请输入1，否则输入0： 1
熵权法确定的权重为：
    0.1411    0.2267    0.4409    0.1913
```

```
最后的得分为：

stand_S =

    0.0390
    0.0552
    0.0411
    0.0428
    0.0362
    0.0441
    0.0489
    0.0525
```

```
sorted_S =

    0.0755
    0.0750
    0.0716
    0.0653
    0.0643
    0.0578
    0.0552
    0.0543
```

```
index =

    11
     9
    10
    12
    20
    15
     2
    13
```