

Word算法手册文档: <https://pubimage.spssau.com/algoword/entropy.docx>

熵值法计算

目录

SPSSAU操作	1
SPSSAU数据格式	2
算法	2
1. 数据标准化	2
2. 非负平移	2
3. 计算比例	2
4. 计算熵值	2
5. 计算冗余度	2
6. 计算权重	2
7. 综合得分	3
参考文献	3

熵值法是一种常用的多指标评价方法，通过计算各指标的信息熵来确定其权重。其位于SPSSAU-» 综合评价-» 熵值法。

SPSSAU 操作



将分析项拖拽至右侧框然后‘开始分析’即可。SPSSAU中涉及两项参数，分别是综合得分和非负平移。

- ✓ 综合得分：选中该参数后，SPSSAU会将综合得分保存为新的标题，标题名称类似为“CompScore_****”。
- ✓ 非负平移：选中该参数后，以列为单位，如果某列数据出现小于等于0，此时平移单位为：最小值的绝对值+0.01，此参数保证数据均为正数可正常计算。

SPSSAU 数据格式

	1 A	2 B	C	D	E
1	样本编号	指标1	指标2	指标3	指标4
2	1				
3	2				
4	3				
5	4				
6	5				
7					
8					
9					

熵值法用于指标的权重情况。1个指标占用1列数据。下图中样本编号只是个编号无实际意义，用于标识下样本的ID号，分析时并不需要使用。

算法

1. 数据标准化

首先，通常需要对原始数据进行标准化处理，以消除量纲的影响。常用的方法有极差标准化和Z-score标准化。

提示：SPSSAU在熵值法时并没有提供此处理，有需要的研究者可通过SPSSAU数据处理->生成变量功能进行设置。

◇ 极差标准化

$$x_{ij}' = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}$$

◇ Z-score标准化

$$x_{ij}' = \frac{x_{ij} - \bar{x}_j}{s_j}$$

其中， \bar{x}_j 为第j个指标的均值， s_j 为第j个指标的标准差。

2. 非负平移

$$x_{ij} = |\min(x_{*j})| + 0.01$$

如果选中该参数，且满足 $x_{*j} \leq 0$ 这一前提时，此系统会进行上式处理。

3. 计算比例

$$p_{ij} = \frac{x_{ij}'}{\sum_{i=1}^m x_{ij}'}$$

其中， m 为样本数量， p_{ij} 为第i个样本在第j个指标上的比例。

4. 计算熵值

$$E_j = -k \sum_{i=1}^m p_{ij} \ln(p_{ij})$$

其中， $k = \frac{1}{\ln(m)}$ ， E_j 为第j个指标的熵值。

5. 计算冗余度

$$d_j = 1 - E_j$$

6. 计算权重

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j}$$

SPSSAU数据科学分析平台

其中， w_j 为第 j 个指标的权重， n 为指标的数量。

7. 综合得分

$$S_i = \sum_{j=1}^n w_j x_{ij}'$$

其中， S_i 为第 i 个样本的综合得分。

提示：如果有进行非负平移，则SPSSAU以非负平移后的数据为准进行综合得分计算。选中该参数后，SPSSAU会将综合得分保存为新的标题，标题名称类似为“CompScore_****”。

参考文献

【1】 The SPSSAU project (2024). SPSSAU. (Version 24.0) [Online Application Software]. Retrieved from <https://www.spssau.com>.

【2】 周俊,马世澎. SPSSAU科研数据分析方法与应用.第1版[M]. 电子工业出版社,2024.

【3】 陈伟,夏建华. 综合主、客观权重信息的最优组合赋权方法[J]. 数学的实践与认识, 2007, 37(001):17-22.

spssau.com