

# 基于评价方法属性层次的组合评价研究

陈衍泰<sup>1</sup>, 陈国宏<sup>2</sup>, 李美娟<sup>2</sup>

(1. 复旦大学管理学院, 上海 200433; 2. 福州大学管理学院, 福州 350002)

**[摘要]** 针对现存的多种单一综合评价方法机理、属性层次相异引起的对同一个对象评价结论不相容的问题, 提出一种新的组合评价方法。定义了“具有相容性”与“特殊属性”的两类综合评价方法; 建立了两个组合评价模型, 分别将这两个属性层次方法所得的不一致的评价值与排序结论进行组合; 在此基础上提出“复组合”的概念。并应用此模型对我国欠发达地区 15 个城市的综合实力进行评价, 使得系统评价更具有科学性和实用性, 为欠发达地区开发战略制订宏观经济政策提供实证基础。

**[关键词]** 评价方法的属性层次; 组合模型; 复组合

**[中图分类号]** C931 **[文献标识码]** A **[文章编号]** 1009-1742(2005)08-0056-04

## 1 引言

由于各种单一评价方法的机理不同、方法的属性层次相异, 在应用中多种评价方法之间评价结论会存在差异。学术界提出“组合评价”的概念, 用以提高评价结论的信度<sup>[1]</sup>。

目前的组合评价方法有: 基于对象属性的组合赋权法<sup>[2]</sup>, 适合将多指标进行组合; 对单一评价法的综合评价进行组合的方法<sup>[3]</sup>, 用对多方法的标准评价值取算术平均进行组合等, 其不足点在于, 组合过于简单。多方法评价组合的研究有待于进一步改进。

笔者曾经从基于方法集角度的组合评价进行分析<sup>[4]</sup>, 建立了组合评价系统分析框架<sup>[1]</sup>, 并利用计算机仿真试验测度单一方法评价结论的漂移度, 以期从评价方法的内在机理角度解决组合评价问题<sup>[5]</sup>。以上的研究偏重于定量分析, 利用模拟分析来取证结论。

下面, 笔者试图在基于方法集的评价上做进一步探索。

## 2 基于评价方法属性层次的组合评价原理与模型

### 2.1 具有属性层次差异的评价方法

1) 一般意义上具有相容性的综合评价方法包括: 多属性决策方法 (MODM)、运筹学方法、统计分析方法、智能化评价方法、信息论方法、灰色系统理论与灰色综合评价、其他新的理论方法 (例如粗糙集评价方法) 等。

2) 特殊属性的综合评价方法: 由于其工作机理的特殊性, 只能适用于某些特定属性的评价对象, 而不具有一般的适用性。常用的具有特殊属性的评价方法有系统工程方法、模糊数学方法、新的评价方法, 例如物元分析方法、可拓集分析方法以及交合分析等。

### 2.2 基本思路

基于评价方法属性层次的组合评价, 解决利用相容性与特殊属性层次上多种方法对同一个对象评

**[收稿日期]** 2004-10-08; **修回日期** 2005-05-13

**[基金项目]** 国家自然科学基金资助项目 (70171026)

**[作者简介]** 陈衍泰 (1979-), 男, 福建三明市人, 复旦大学管理学院博士研究生; 陈国宏 (1953-), 男, 福建福州市人, 博士, 福州大学教授, 博士生导师

价结论不相容问题，为决策提供科学依据。首先在给定的两个层次的评价方法集合里，利用统计学中 KENDALL 协和系数，检验多种单一方法的评价结论是否具有相容性<sup>[6]</sup>；其次，分别利用两个数学模型将相容性和特殊属性两个层次方法的评价结论进行组合；第三，将组合的结论进行复组合；最后用统计学方法检验组合模型相对于原来单一方法评价结论的有效性，并进行后续处理，得到最终有效的评价结论，为科学决策提供理论基础。

2.3 KENDALL 协和系数统计检验

用统计学中 KENDALL 一致性系数检验各种单一方法排序结果。选用表 1 中  $b$  (这里  $b$  分别是 7 和 3) 种综合评价方法分别对  $n$  个对象进行评价；检验这些评价结果之间是否一致。

假设  $H_0$ :  $b$  种方法的评价等级不具有不一致性；

$H_1$ :  $b$  种方法的评价等级具有一致性。

表 1 两种组合方法评价得分以及复组合最终排序结果

Table 1 Conclusion of two different combination evaluation methods & the recombination results

城市	组合模型 1 得分	排序	组合模型 2 得分	排序	复组合 评价排序	四种排序 的方差	城市	组合模型 1 得分	排序	组合模型 2 得分	排序	复组合 评价排序	四种排序 的方差
重庆	0.248 519	1	14	1	1	—	乌鲁木齐	0.041 445	10	-8	12	11	1.25 831 <sup>①</sup>
成都	0.159 632	2	12	2	2	—	南昌	0.043 748	9	0	8	8	0.5 000 <sup>②</sup>
西安	0.088 158	3	10	3	3	—	兰州	0.035 381	11	-6	11	10	0.50 000 <sup>③</sup>
郑州	0.083 834	4	6	5	4	0.81 650 <sup>①</sup>	海口	0.024 463	12	-4	10	12	2.38 048 <sup>③</sup>
昆明	0.07 806	5	8	4	5	1.25 831 <sup>①</sup>	呼和浩特	0.007 083	13	-10	13	13	—
南宁	0.076 527	6	4	6	6	—	西宁	0.002 767	14	-12	14	14	—
太原	0.052 073	7	2	7	7	—	银川	0	15	-14	15	15	—
贵阳	0.048 412	8	-2	9	9	1.5 000 <sup>②</sup>	—	—	—	—	—	—	—

①②③分别为计算多个城市应用二次组合模型排序相同时对比组别。

2.4 互补性集化原理与模型

2.4.1 基本模型 1 基于相容性评价结论的组合。

设多指标评价问题有  $n$  个评价对象,  $b_1$  种具有相容性的评价方法。经过标准化处理后, 第  $k$  种评价方法的评价结果权向量为

$$H_k = (h_{1k}, h_{2k}, \dots, h_{nk})^T, (k = 1, 2, \dots, b_1)$$

组合评价权向量为 (假设  $h_{i0}$  均为非负数) 为

$$H_0 = (h_{10}, h_{20}, \dots, h_{n0})^T$$

在权重偏差平方和最小的意义下, 构造最优化模型

$$\begin{cases} \min \sum_{k=1}^{b_1} \sum_{i=1}^n (h_{i0} - h_{ik})^2 \\ s. t. \sum_{i=1}^n h_{i0} = 1 \end{cases} \quad (2)$$

KENDALL 协和系数是在假设  $H_0$  条件下当  $n \rightarrow +\infty$  时的统计检验量,  $x^2 = b(n-1)W$  近似服从  $\chi^2(n-1)$ ; 在实际应用中一般  $n > 7$  就可以使用<sup>[6]</sup>。

$$W = \frac{12 \sum_{i=1}^n r_i^2 - 3b^2n(n+1)^2}{b^2n(n^2-1)} \quad (1)$$

其中,  $n$  为评价对象的数目;  $b$  为评价方法的数目;

$r_i = \sum_{j=1}^b y_{ij}$ ,  $y_{ij}$  为第  $i$  个对象在第  $j$  种评价方法中的排序值。

当  $X^2 \geq \chi_a^2(n-1)$  (当  $n > 7$ ) 时拒绝原假设, 认为  $b$  种单一评价方法的评价等级之间具有相容性, 说明可以对  $b$  种单一方法的评价结论进行下一步的组合研究; 否则, 只能逐步剔除其中的一些方法, 直到满足评价等级相容性条件方能组合。

可以构造 Lagrange 函数可以求  $h_{i0}$ ,  $\mu$  的偏导数后, 解方程可以得到

$$h_{i0} = \frac{1}{n} + \frac{1}{nb_1} (na_i - \sum_{i=1}^n \alpha_i), i = 1, 2, \dots, n \quad (3)$$

其中  $\alpha_i = \sum_{k=1}^{b_1} h_{ik}$ 。此时, 总方差达到最小, 这是多方法评价对加权均值的估计。定义 (2) ~ (3) 为最小方差的组合评价模型 1。最后对  $h_{i0}$  进行排序, 得到组合评价的结论。

2.4.2 基本模型 2 基于特殊属性评价结论的组合。基本原理是“少数服从多数”的方法。设多指标评价问题有  $n$  个评价对象,  $b_2$  种具有特殊属性的评价方法。这是针对  $b_2$  种评价方法能通过

KENDALL 检验、但排序结果不相容性的组合模型。

设在所选择的  $b_2$  种具有相容性方法所得的排序中,  $n$  个评价对象中对象  $i$  优于对象  $j$  的次数为  $s$ , 定义优良系数为

$$C_{ij} = \begin{cases} 1, & s > b_2/2 \\ 0, & s = b_2/2 \\ -1, & s < b_2/2 \end{cases} \quad (i, j = 1, 2, \dots, n) \quad (5)$$

再定义评价对象  $i$  的组合排序系数  $C_i$  为

$$C_i = \sum_{j=1}^n C_{ij} \quad (6)$$

现在, 对  $n$  个对象的组合排序系数进行再排序, 可以得到一个组合评价结论。组合排序系数越大, 说明这个评价对象的得分越高, 那么这个方案在系统中的性能就越好; 如果  $C_i = C_j$ , 就计算这两个方案在  $b_2$  种方法下的得分的方差, 方差越小的那个方案性能越好。

### 2.5 复组合评价法

以上两种组合评价方法的机理不同, 若二者仍存在差异, 需进行再组合, 从而确定最终组合评价结果。笔者采用两种组合排序的平均值作为复组合的评价结论, 如果复组合结论中的多个对象的排序号相同, 那么计算这多个对象在两种排序号的方差, 选优的准则是多个排序号的方差越小越好。

### 2.6 复组合评价结论相容性的检验

应用统计学中的 Spearman 等级相关系数检验组合评价的有效性<sup>[7]</sup>。

假设  $H_0$ : 复组合评价结论与原 2 种组合评价结论无关;

$H_1$ : 复组合评价结论与原来的 2 种组合方法密切相关。

$\rho_k$  表示复组合方法与第  $k$  种一次组合评价结论之间的 Spearman 相关系数, 表示两者的相关程度。

令  $\rho = \frac{1}{2} \sum_{k=1}^2 \rho_k$ , 表示复组合方法的评价结论与 2 种不同属性层次的组合评价结论的平均相关程度。

那么, 构造检验统计量  $t_k$ , 当  $n \geq 10$  时

$$t = \rho \sqrt{\frac{n-2}{1-\rho^2}} \sim t_{\alpha/2}(n-2) \quad (7)$$

给定显著性水平  $\alpha$ , 查  $t$  分布表得到临界值, 如果最终评价排序能够通过检验, 即可得到组合评价的最终排序。

## 3 实证研究

平衡发展是新世纪国家发展的一个重要战略。欠发达地区中心城市的综合发展是带动周边地区经济发展的重要动力, 分析与评价欠发达地区的中心城市综合发展水平具有十分重要的意义。

### 3.1 综合评价方法的抽取

笔者从具有相同属性的评价方法集合中, 任意抽取三种评价方法: 基于信息熵的评价方法(方法 1)、灰色关联度评价方法(方法 2)和主成分分析评价方法(方法 3)对我国 15 个西部开发的中心城市的综合发展水平做评价。

同时从具有特殊属性层次方法集合选取改进的数据包络分析方法(方法 4)、模糊综合评价方法(方法 5), 对相同对象进行评价。即上面公式中  $b_1 = 3, b_2 = 2$ 。

### 3.2 资料及其来源

选取了反映城市综合发展水平的 10 个社会经济指标和 8 个城市公共设施水平的指标。指标的选择参考了《中国统计年鉴》中的统计指标的设置。数据来源于《中国统计年鉴(2002)》<sup>[8]</sup>, 用于欠发达地区的中心城市发展综合水平的评价。

### 3.3 组合评价模型的应用

分别应用相容性评价方法中的三种单一的评价方法对 15 个中心城市的综合发展水平进行综合评价, 各方法的得分与排序存在很大的差异性。限于版面问题, 这里省略具体数据。对评价结果进行一致性检验。运用 KENDALL 协和系数法检验, 通过 SPSS 软件进行检验, 计算得到协和系数  $W$  为 15.635, 查表  $\chi_{0.05}^2(n-1)$  为 23.685。其统计意义是在置信度 95% 的情况下三种单一综合评价方法的结果不具有相容性。所以必须剔除其中的一种。将方法 3 主成分分析评价结论剔除后, 剩下熵值评价与灰关联评价的结论。运用 KENDALL 协和系数法检验, 通过 SPSS 软件进行检验, 计算得到协和系数  $W$  为 50.236, 查表  $\chi_{0.01}^2(n-1)$  为 29.141。其统计意义是在置信度 99% 的情况下三种单一综合评价方法的结果具有相容性。说明用这两种相同属性方法对待同一个对象, 结论是可以进行组合。

同样, 分别应用方法 4 与方法 5 的两个排序结论存在差异性, 对评价结果进行一致性检验。运用 KENDALL 协和系数法检验, 通过 SPSS 软件进行检验, 计算得到协和系数  $W$  为 24.321, 查表  $\chi_{0.05}^2$

$(n - 1)$  为 23.685。其统计意义是在置信度 95% 的情况下三种单一综合评价方法的结果具有相容性。说明可以进行继续组合。

分别用组合评价基本模型 1 和模型 2 对两种不同组合方法所得的结论 (见表 1) 进行组合, 得到的排序见表 1 第 2~4 列。在两种组合所得的结论中仍然存在排序的细微差异, 这样决策者很难判断到底哪个评价结论有效, 所以有必要进行复组合。

### 3.4 组合评价的复组合与检验

计算两个组合结论排序的平均值, 进行复组合, 即得到最终评价结论 (见表 1 的第 6、第 7 列)。采用公式 (7) 进行组合结论的检验, 计算得到  $t = 2.732$ 。在给定显著性水平  $\alpha = 0.01$ , 查表得  $t_{\alpha/2}(n - 2) = 2.650$ , 说明此时组合方法评价的结论和原来两种不同的组合评价方法密切相关, 得到的组合评价的排序的结论是有效、可信的。

通过最终的评价结果可以知道西部大城市综合实力名列首位的是直辖市重庆, 成都和西安分居二、三位。这是和实际相符合的。

## 4 结论

笔者提出一种综合各种具有相容性评价方法的组合方法, 可以克服仅仅使用一种方法的片面性, 并解决了应用单一方法评价结果存在差异的现象。用此方法进行综合评价可以使得评价效果更加客

观、真实、全面、符合实际。不过该方法目前只适用于具有相容性的单一综合评价方法, 进行统计一致性检验。鉴于组合评价的研究刚刚起步, 其机理和应用还有待于进一步研究和改进。

### 参考文献

- [1] 陈国宏, 陈衍泰, 李美娟. 综合评价方法问题的系统研究 [J]. 复旦大学学报 (自然科学版), 2003, (5)
- [2] 宁自军. 多种综合评价方法的综合应用 [J]. 数理统计与管理, 2000, 19 (3): 13~16
- [3] 王坚强. 多目标组合决策方法研究 [J]. 系统工程与电子技术, 2002, 24, (10): 70~72
- [4] 陈国宏, 李美娟. 基于方法集的组合评价研究 [J]. 中国管理科学, 2004, (1): 37~42
- [5] 陈国宏, 李美娟, 陈衍泰. 单一方法评价结论漂移性的测度研究 [J]. 中国工程科学, 2004, 6(3): 58~63
- [6] Chen Yantai, Chen Guohong, Li Meijuan. Combination evaluation approach based on attributes difference of methods [C]. Proceedings of 2004' International Conference of Management. Macao: Macao university of science and technology.
- [7] 曾宪报. 关于组合评价的事先事后检验 [J]. 统计研究, 1997, (6): 56~58
- [8] 国家统计局编. 中国统计年鉴 (2002) [J]. 北京: 中国统计出版社, 2002

## Studies on Combination Evaluation Methods Based on Nature Differenc

Chen Yantai<sup>1</sup>, Chen Guohong<sup>2</sup>, Li Meijuan<sup>2</sup>

(1. Management School of Fudan University, Shanghai 200433, China;

2. Management School of Fuzhou University, Fuzhou 350002, China)

[Abstract] This paper firstly defines a group of system evaluation methods of attributes difference. Secondly, it puts forward one new combination evaluation method, in order to solve the problem of inconsistencies in the conclusions, which is caused by adopting many different single methods to evaluate the same objective. This method is used to evaluate the comprehensive competitive power of 15 cities in western China to make the system evaluation more scientific and useful.

[Key words] evaluation methods with different nature; combination evaluation; re-combination