Advances in Applied Mathematics 应用数学进展, **2021, 10(2), 603-616**
Published Online February 2021 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2021.102066

Hans汉斯

基于熵权TOPSIS的深圳未来医疗和养老保障问题的评价及对策

范卓然[*]**，徐纪隆，李宇宸**

中国海洋大学数学科学学院，山东 青岛
Email: [*]fanzhuoranhd@126.com

收稿日期：2021年1月23日；录用日期：2021年2月17日；发布日期：2021年2月26日

摘 要

本文从人口数量与结构、经济收入与消费水平、医疗资源与水平、社会保障制度与能力五个方面出发，建立未来深圳市医疗和养老资源合理配置的评价体系。首先对原始数据标准化，借助于Matlab软件编程对16个指标进行R-型聚类，根据聚类图可以分为7类，进而选取7个指标构建评价体系。然后通过分析深圳市近十几年的各项数据，采用熵权法确定医疗和养老资源评价体系的系数。通过TOPSIS法确定评价系统的正、负理想，从而计算评价系统与理想解之间的欧氏距离，进而确定人口数量与结构、经济收入与消费水平、医疗资源与水平、社会保障制度与能力以及综合能力的排序。最后利用多元线性回归模型对人均预期寿命进行预测，进而得到人均预期寿命的回归方程。

关键词

指标聚类，熵权法，TOPSIS方法，残差分析，多元线性回归

Evaluation and Countermeasures of Future Medical and Pension Security Problems in Shenzhen Based on Entropy Weight and TOPSIS

Zhuoran Fan[*]**, Jilong Xu, Yuchen Li**

College of Mathematical Sciences, Ocean University of China, Qingdao Shandong
Email: [*]fanzhuoranhd@126.com

Received: Jan. 23rd, 2021; accepted: Feb. 17th, 2021; published: Feb. 26th, 2021

[*]通讯作者。

文章引用: 范卓然, 徐纪隆, 李宇宸. 基于熵权 TOPSIS 的深圳未来医疗和养老保障问题的评价及对策[J]. 应用数学进展, 2021, 10(2): 603-616. DOI: 10.12677/aam.2021.102066

Abstract

In this paper, from the population and structure, economic income and consumption level, medical resources and level, social security system and ability of five aspects, we established the future rational allocation of medical and pension resources in Shenzhen evaluation system. Firstly, the original data were standardized, and the R-type clustering was carried out on the 16 indicators with the help of MATLAB software programming. According to the clustering diagram, the indicators could be divided into 7 categories. Then 7 indexes were selected to construct the evaluation system. Then, by analyzing the data of Shenzhen in recent ten years, the entropy weight method was adopted to determine the coefficient of the medical and pension resource evaluation system. The positive and negative ideals of the evaluation system were determined by TOPSIS method, so as to calculate the Euclidean distance between the evaluation system and the ideal solution, and then determine the order of population size and structure, economic income and consumption level, medical resources and level, social security system and ability and comprehensive ability. Finally, multiple linear regression model was used to predict the average life expectancy, and the regression equation of the average life expectancy was obtained.

Keywords

Index Clustering, Entropy Weight Method, TOPSIS Method, Residual Analysis, Multiple Linear Regression

1. 引言

十九大报告提出中国特色社会主义进入新时代，我国社会主要矛盾已经转化为人民日益增长的美好生活需要和不平衡不充分发展之间的矛盾。深圳是中国改革开放的前沿阵地，是中国经济建设和对外开放的“试验场”。借助于历史数据如何根据深圳的人口状况，经济地位，历史特点，社会环境，发展趋势，人民的期待来构建深圳市医疗和养老保障评价指标，根据评价的结果找出影响未来 5 年、10 年和 15 年深圳市医疗和养老保障目标实现的主要因素，并进一步分析原因，找出实现目标的方案，合理配置深圳市城市资源，建设“民生幸福标杆”和“可持续发展先锋”城市。未来深圳如何利用好政策红利，建立可持续发展的医疗和养老保障体系，来实现建设国家“先行示范区”成为人们关注的热点之一[1]。

2. 评价指标体系的建立

借助于健康中国 2030 年建设主要指标及参考国际上先进标准，根据中国国情和深圳市经济发展现状给出未来 5 年、10 年和 15 年深圳医疗和养老保障需要实现的目标的量化描述。以国际、国内权威机构经典观点的高频指标为重点，结合文献梳理和调查研究进行指标的海选。根据可观测性原则将数据无法获得的海选指标删除，保证初步筛选后的指标体系可以量化。

本文采用了深圳市、广东省和国家有关政府部门发布的医疗、养老资源以及健康发展的定量数据，包括 2000~2019 年《中国卫生统计年鉴》、2000~2019 年《中国卫生和计划生育统计年鉴》、2019 年《中国卫生健康统计年鉴》、2000~2019 年《深圳市统计年鉴》、2000~2019 年《广东省卫生年鉴》数据，以及各年的卫生统计公报或者政府工作报告等[2]。

为了量化描述深圳医疗和养老保障我们选取了每千人口床位数、每千人口卫生工作人员数、每千人口卫生技术人员数、每千人口医师数、城乡居民达到《国民体质测定标准》合格以上的人数比例、居民年末常住人口、基本养老保险人数、养老保险覆盖率、人均地区生产总值(人均 GDP)、人均可支配收入、人均消费支出、恩格尔系数等 16 个指标[3]，建立深圳市未来 5 年、10 年和 15 年合理配置医疗和养老资源评价指标体系，见表 1。

这些指标数量较大，且这种大系统的指标体系中的各指标之间往往存在着联系，若直接使用这些指标来分析其与医疗和养老保障之间的关系工作难度较大，下面采用聚类方法对指标进行 R 型聚类。

2.1. 原始数据的标准化处理

深圳市合理配置医疗和养老资源评价指标体系中有两类指标。一类是人均地区生产总值(人均 GDP)等指标值越大表示医疗和养老资源合理配置越好的指标。另一类是居民恩格尔系数等指标值越小表示医疗和养老资源合理配置越好的指标。

Table 1. Rational allocation of medical and old-age resources evaluation index system
表 1. 合理配置医疗和养老资源评价指标体系

序号	准则	指标层
1	x_1 医疗资源与水平	x_{11} 每千人口床位数
2		x_{12} 每千人口卫生工作人员数
3		x_{13} 每千人口卫生技术人员数
4		x_{14} 每千人口医师数
5	x_2 居民健康水平	x_{21} 人均预期寿命
6		x_{22} 达到国民体质标准合格以上的人数比例
7		x_{23} 基本保险养老人数
8		x_{24} 养老保险覆盖率
9	x_3 经济收入与消费水平	x_{31} 人均 GDP
10		x_{32} 城镇居民可支配收入
11		x_{33} 居民人均消费支出
12	x_4 人口数量与结构	x_{41} 恩格尔系数
13		x_{42} 居民年末常住人口
14		x_{43} 非户籍人口
15	x_5 社会保障制度与能力	x_{51} 医疗卫生事业费
16		x_{52} 人均医疗事业费
17		x_{53} 地方财政支出

为了消除量纲和指标正负向不同的影响。本文利用模糊隶属度方法对指标进行标准化。设 y_{ij} 为第 j 个评价对象第 i 个指标标准化后的值；w_{ij} 为第 j 个评价对象第 i 个指标值；n 为被评价的对象数。则可通过正向指标标准化公式(1)，负向指标标准化公式(2)对指标进行标准化。

$$y_{ij} = \frac{w_{ij} - \min_{1 \le j \le n} w_{ij}}{\max_{1 \le j \le n} w_{ij} - \min_{1 \le j \le n} w_{ij}} \tag{1}$$

$$y_{tj}=\frac{\max_{1\le j\le n} w_{tj}-w_{tj}}{\max_{1\le j\le n} w_{tj}-\min_{1\le j\le n} w_{tj}} \tag{2}$$

下面用 R-型聚类方法对指标进行分类，借助于 Matlab 软件编程[4] [5]，利用公式(1)和(2)标准化方法将前面收集到的 16 个数据作标准化处理，得到原始数据的标准化矩阵，见表 2。

Table 2. Standardization of raw data
表 2. **原始数据的标准化**

年份	x_{11}	x_{12}	x_{13}	x_{14}	x_{22}	x_{23}	x_{24}	x_{31}
2009	0	0.075	0.0772	0.1042	0.0857	0	0	0
2010	0.0298	0	0	0	0.9714	0.1704	0.0293	0.0977
2011	0.0893	0.125	0.1382	0.1146	0.6	0.2662	0.2396	0.2083
2012	0.2976	0.2643	0.2683	0.2292	0.4	0.3621	0.4503	0.341
2013	0.3571	0.425	0.3984	0.3542	0.4857	0.3994	0.5163	0.4673
2014	0.4345	0.5357	0.5203	0.4583	0.3714	0.4552	0.6047	0.5721
2015	0.7143	0.5786	0.5569	0.5208	0	0.5881	0.7306	0.6488
2016	0.7976	0.575	0.5732	0.5417	0.5	0.7077	0.8362	0.7334
2017	0.8036	0.6429	0.6504	0.6354	1	0.8727	1	0.8287
2018	0.8929	0.8179	0.8049	0.7708	0.3429	0.9112	0.9324	0.8804
2019	1	1	1	1	0.6	1	0.9931	1
年份	x_{32}	x_{33}	x_{41}	x_{42}	x_{43}	x_{51}	x_{52}	x_{53}
2009	0	0	0.2267	0	0	0	0	0
2010	0.0942	0.0593	0.16	0.322479	0.7436	0.0099	0.0026	0.0736
2011	0.2182	0.1183	0	0.343555	0.6476	0.0273	0.0132	0.1642
2012	0.3455	0.241	0.0533	0.361228	0.5888	0.0435	0.0336	0.1573
2013	0.463	0.3375	0.08	0.379233	0.515	0.0893	0.0932	0.1866
2014	0.3517	0.3394	0.48	0.412372	0.4791	0.1728	0.1994	0.3244
2015	0.4624	0.5018	0.6267	0.54488	0.6679	0.2582	0.2867	0.7012
2016	0.5845	0.6928	0.8267	0.661902	0.7854	0.3216	0.3423	0.6379
2017	0.712	0.778	0.8933	0.798851	0.8445	0.4113	0.4189	1
2018	0.8624	0.8806	1	0.908936	0.9943	0.5257	0.5211	0.9134
2019	1	1	0.9733	1	1	1	1	0.9886

2.2. 医疗和养老资源合理配置评价指标的选取

对上述 16 个评价指标进行聚类分析，利用 MATLAB 软件编程将指标分类，分类情况如图 1 所示。

从聚类图可以看出深圳市医疗卫生事业费和人均医疗事业费两者有较强的相关性，都是涉及到医疗事业费的指标，最先聚为一类。深圳市每千人口卫生工作人员数、卫生技术人员数、医师数三者之间相关性较强，三者均涉及到医疗资源与水平的问题，聚为一类。每千人口床位数、养老保险覆盖率、人均 DGP、基本养老保险人数、城镇居民可支配收入、居民人居消费支出和年末居民常住人口这 7 个指标相

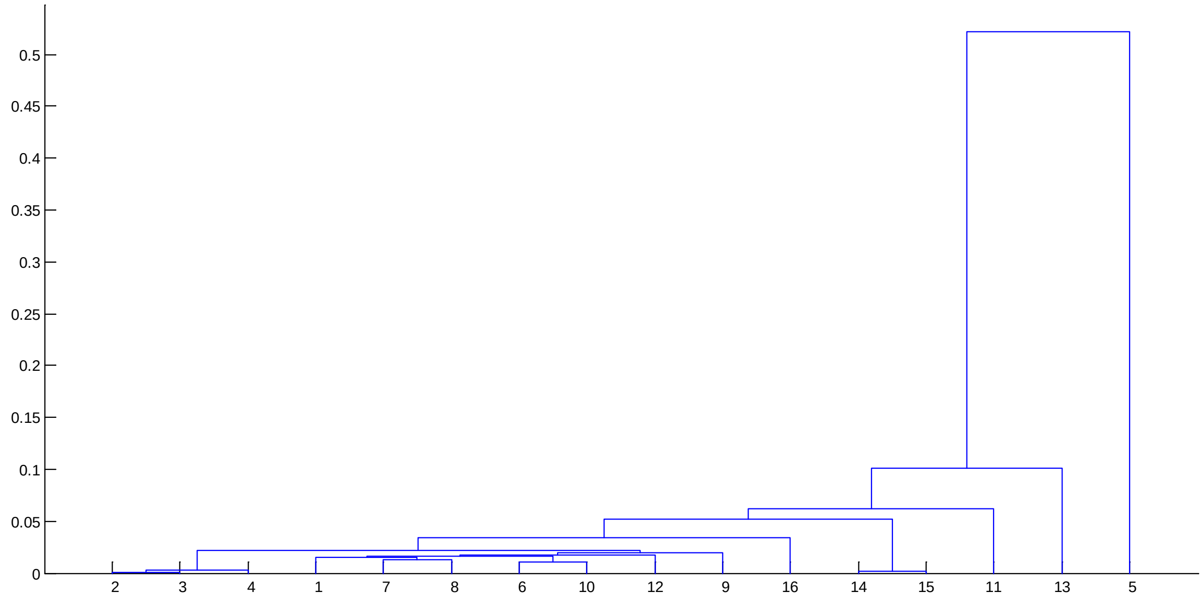

Figure 1. Clustering pedigree of sixteen indexes
图 1. 十六个指标的聚类谱系图

关性较强，都涉及到经济发展和养老保障问题，聚为一类也很正常。这样就从 16 个指标中选出 7 个指标对医疗和养老资源合理配置进行评价。这七个评价指标分别为：

x_{14} 每千人口医师数；

x_{22} 达到国民体质标准合格以上的人数比例；

x_{31} 人均 GDP；

x_{41} 恩格尔系数；

x_{43} 非户籍人口；

x_{51} 医疗卫生事业费；

x_{53} 地方财政支出。

这七个指标的标准化数据，见表 3。

Table 3. Standardization of evaluation index data
表 3. 评价指标数据的标准化

年份	x_{14}	x_{22}	x_{31}	x_{41}	x_{43}	x_{51}	x_{53}
2009	0.1042	0.0857	0	0.2267	0	0	0
2010	0	0.9714	0.0977	0.16	0.7436	0.0099	0.0736
2011	0.1146	0.6	0.2083	0	0.6476	0.0273	0.1642
2012	0.2292	0.4	0.341	0.0533	0.5888	0.0435	0.1573
2013	0.3542	0.4857	0.4673	0.08	0.515	0.0893	0.1866
2014	0.4583	0.3714	0.5721	0.48	0.4791	0.1728	0.3244
2015	0.5208	0	0.6488	0.6267	0.6679	0.2582	0.7012
2016	0.5417	0.5	0.7334	0.8267	0.7854	0.3216	0.6379
2017	0.6354	1	0.8287	0.8933	0.8445	0.4113	1
2018	0.7708	0.3429	0.8804	1	0.9943	0.5257	0.9134
2019	1	0.6	1	0.9733	1	1	0.9886

3. 医疗和养老资源合理配置的熵权 TOPSIS 评价方法

下面用熵权 TOPSIS 法对医疗和养老资源合理配置进行评价[6] [7]。一是用熵权法确定医疗和养老资源评价体系的系数。二是通过 TOPSIS 法，即逼近理想解的技术，确定被评价对象的排序。TOPSIS 法的核心思想是定义决策问题的理想解和负理想解，然后比较评价方案与理想解和负理想解的距离，最后计算各方案与理想解的相对贴近度，进行方案的优劣排序。

3.1. 利用熵权法求指标的客观权重

设 w_{ij} 为第 i 个系统中的第 j 项指标的观测值；这里 $i=2009,2010,\cdots,2019$； $j=1,2,\cdots,7$ 。对给定的指标 j， w_{ij} 的差异越大，该项指标对系统的比较作用就越大，亦即该指标包含和传输的信息越多。熵权法确定指标权重的具体步骤如下：

1) 计算深圳市每千人口医师数等 7 个指标的熵值。

设 e_j 为第 j 个指标的熵值，求解过程如下

$$p_{ij}=\frac{w_{ij}}{\sum_{i=1}^{11}w_{ij}} \text{和} e_j=-\frac{1}{lnn}\sum_{i=1}^{11}p_{ij}\ln p_{ij} \tag{3}$$

其中 p_{ij} 为第 i 个系统中的第 j 项指标的特征比重；这里 $i=1\text{表示}2009\text{年}, i=2\text{表示}2010\text{年}, \text{依次类推}, i=11\text{表示}2019\text{年}$； $j=1,2,\cdots,7$ 、 $\sum_{i=1}^{11}w_{ij}$ 为第 j 项指标的所有系统观测数据之和。

2) 指标 j 的差异系数 g_j 的计算：

$$g_j=1-e_j \tag{4}$$

3) 确定 7 项指标的权重系数：

$$s_j=\frac{g_j}{\sum_{i=1}^{7}g_j} \tag{5}$$

按照上述步骤，采用表 3 中的数据，借助于 Matlab 编程[8] [9]，可求得特征比重 p_{ij} 、熵值 e_j 、差异系数 g_j 和权重系数 s_j 值，具体结果见表 4。

Table 4. Entropy value, difference coefficient and weight coefficient data
表 4. 熵值、差异系数和权重系数数据

	x_{14}	x_{22}	x_{31}	x_{41}	x_{43}	x_{51}	x_{53}
熵值 e_j	0.1427	0.1065	0.1318	0.1350	0.1185	0.1596	0.1376
差异系数 g_j	0.8573	0.8935	0.8682	0.8650	0.8815	0.8404	0.8624
权重系数 s_j	0.1414	0.1472	0.1431	0.1426	0.1453	0.1385	0.1421

3.2. 标准化数据的加权

设 y_{ij} 为第 i 年第 j 个指标标准化数据的加权值， w_{ij} 为第 i 年中的第 j 项指标观测值规范化处理后的值， s_j 为权重系数，根据加权方法可知：

$$y_{ij}=w_{ij}s_j \tag{6}$$

利用公式(6)以及表 3 中的数据可计算评价指标加权后的数据矩阵，见表 5。

Table 5. The weighted data matrix of evaluation indicators
表 5. 评价指标加权后的数据矩阵

年份	x_{14}	x_{22}	x_{31}	x_{41}	x_{43}	x_{51}	x_{53}
2009	0.0147	0.0126	0	0.0323	0	0	0
2010	0	0.1430	0.0140	0.0228	0.1080	0.0014	0.0105
2011	0.0162	0.0883	0.0298	0	0.0941	0.0038	0.0233
2012	0.0324	0.0589	0.0488	0.0076	0.0856	0.006	0.0224
2013	0.0501	0.0715	0.0669	0.0114	0.0748	0.0124	0.0265
2014	0.0648	0.0547	0.0819	0.0684	0.0696	0.0239	0.0461
2015	0.0736	0	0.0928	0.0894	0.0970	0.0358	0.0996
2016	0.0766	0.0736	0.1049	0.1179	0.1141	0.0445	0.0906
2017	0.0898	0.1472	0.1186	0.1274	0.1227	0.0570	0.1421
2018	0.1090	0.0505	0.1260	0.1426	0.1445	0.0728	0.1298
2019	0.1414	0.0883	0.1431	0.1388	0.1453	0.1385	0.1405

3.3. 确定评价系统的正、负理想

设 y_j^+ 为第 j 个指标观测数据的最大值， y_j^- 为第 j 个指标观测数据的最小值， $j=1,2,\cdots,7$ 。

$$y_j^+ = \max_{2009\le k\le 2019} y_{kj} \text{和} y_j^- = \min_{2009\le k\le 2019} y_{kj} \tag{7}$$

易知评价系统的正理想解为

$$y_j^+ = \left(y_1^+, y_2^+, \cdots, y_7^+\right) = (0.1414, 0.1472, 0.1431, 0.1426, 0.1453, 0.1385, 0.1421);$$

负理想解为 $y_j^- = \left(y_1^-, y_2^-, \cdots, y_7^-\right) = (0,0,0,0,0,0,0)$ 。

3.4. 计算评价系统与理想解之间的欧式距离

设 $d_{(i,t)}^+$ 为第 i 年第 t 个准则与相应分量正理想解的欧式距离， $d_{(i,t)}^-$ 为第 i 年第 t 个准则与相应分量负理想解的欧式距离， $t=1,2,\cdots,5$ 。例如准则层：社会保障制度与能力含有二级指标 2 个， x_{51} 医疗卫生事业费和 x_{53} 地方财政支出。则

$$d_{(2009,5)}^+ = \sqrt{\left(y_6^+ - y_{(2009,6)}\right)^2 + \left(y_7^+ - y_{(2009,7)}\right)^2} \tag{8}$$

$$d_{(2009,5)}^- = \sqrt{\left(y_6^- - y_{(2009,6)}\right)^2 + \left(y_7^- - y_{(2009,7)}\right)^2} \tag{9}$$

设 d_i^+ 为第i年与正理想解的欧式距离， d_i^- 为第 i 年指标与负理想解的欧式距离，则

$$d_i^+ = \sqrt{\left(y_1^+ - y_{i1}\right)^2 + \left(y_2^+ - y_{i2}\right)^2 + \cdots + \left(y_7^+ - y_{i7}\right)^2}; \tag{10}$$

$$d_i^- = \sqrt{\left(y_1^- - y_{i1}\right)^2 + \left(y_2^- - y_{i2}\right)^2 + \cdots + \left(y_7^- - y_{i7}\right)^2}. \tag{11}$$

3.5. 计算相对贴近度评价结果

设 f_i^j 为第 i 年第 t 个准则下的指标与理想解的相对贴近度， f_i 为第 i 年所有指标与理想解的相对贴近度，则

$$f_i^t = \frac{d_{(i,t)}^-}{d_{(i,t)}^- + d_{(i,t)}^+} \tag{12}$$

$$f_i = \frac{d_i^-}{d_i^- + d_i^+} \tag{13}$$

其中 $t = 1,2,\cdots,5$； $i = 2009, 2010,\cdots, 2019$ 。

通过计算贴近度来确定被评价指标的发展状况。相对贴近度 f_i^j 越大表示被评价指标与理想解越接近，发展状况越好。

利用表 5 中的数据以及公式(7)~(13)，借助于 Matlab 软件编程计算相对贴进度 f_i^j 和 f_i ，具体数值见表 6。

Table 6. Evaluation results and rankings for each criterion layer from 2009 to 2019
表 6. 2009~2019 年间各准则层的评价结果和排名

准则	医疗资源与水平		居民健康水平		经济收入与消费水平		人口数量与结构		社会保障制度与能力		综合能力	
年份	f_i^1	排名	f_i^2	排名	f_i^3	排名	f_i^4	排名	f_i^5	排名	f_i	排名
2009	0.10272537	10	0.0880503	10	0	11	0.150214509	11	0	11	0.0955	11
2010	0	11	0.9993012	1	0.09783368	10	0.467059927	7	0.0513798	10	0.3789	7
2011	0.11320755	9	0.617051	4	0.20824598	9	0.382390194	8	0.1130755	8	0.3145	9
2012	0.22641509	8	0.4116003	7	0.34102027	8	0.367242436	9	0.111859	9	0.298	10
2013	0.35010482	7	0.4996506	6	0.46750524	7	0.336214119	10	0.1421093	7	0.3385	8
2014	0.45283019	6	0.3822502	8	0.57232704	6	0.478505999	6	0.2508489	6	0.4153	6
2015	0.51432565	5	0	11	0.64849755	5	0.646194491	5	0.4757533	4	0.4897	5
2016	0.53529001	4	0.5143256	5	0.73305381	4	0.803572728	4	0.4723486	5	0.6106	4
2017	0.62753319	3	0.9729015	2	0.82879106	3	0.865368856	3	0.6399	3	0.7526	2
2018	0.7617051	2	0.3529001	9	0.88050314	2	0.995374552	1	0.6739802	2	0.7131	3
2019	0.9881202	1	0.617051	3	1	1	0.979049083	2	0.9674001	1	0.8581	1

4. 2009~2019 年间医疗和养老资源合理配置的评价分析

以 2009~2019 年的年份为横坐标，以表 6 中 6 个评价准则和 11 年间医疗和养老资源综合评价的相对贴近度得分为纵坐标画图，借助于 Matlab 编程得出 11 年间深圳市医疗和养老资源情况图，见图 2。

从图 2 可以看出，随着 11 年来深圳市经济的快速发展，深圳市的医疗资源水平、经济收入与消费水平、人口数量与结构、社会保障制度和能力以及各方面的综合能力基本上呈现出逐年递增趋势，这也表现出经济发展的确能带给老百姓物资方面实实在在的好处。从综合能力来看，2017 年的状况好于 2018 年。主要原因是 2018 年深圳市居民健康水平有所下降，比 2017 年度下降了 2.3 个百分点。设置这个指标的主要目的，就是增强市民健身意识，引导市民科学健身，推动深圳全民健身运动发展。

但居民健康水平波动性较大，这说明了经济的高速发展，并没有给人民的生活带来满足感，人民生活水平的发展远远落后于经济发展的速度。恩格尔系数虽逐年下降，但仍高于 29%，老百姓的民生问题，尤其是医疗问题还需要进一步的完善和改进。

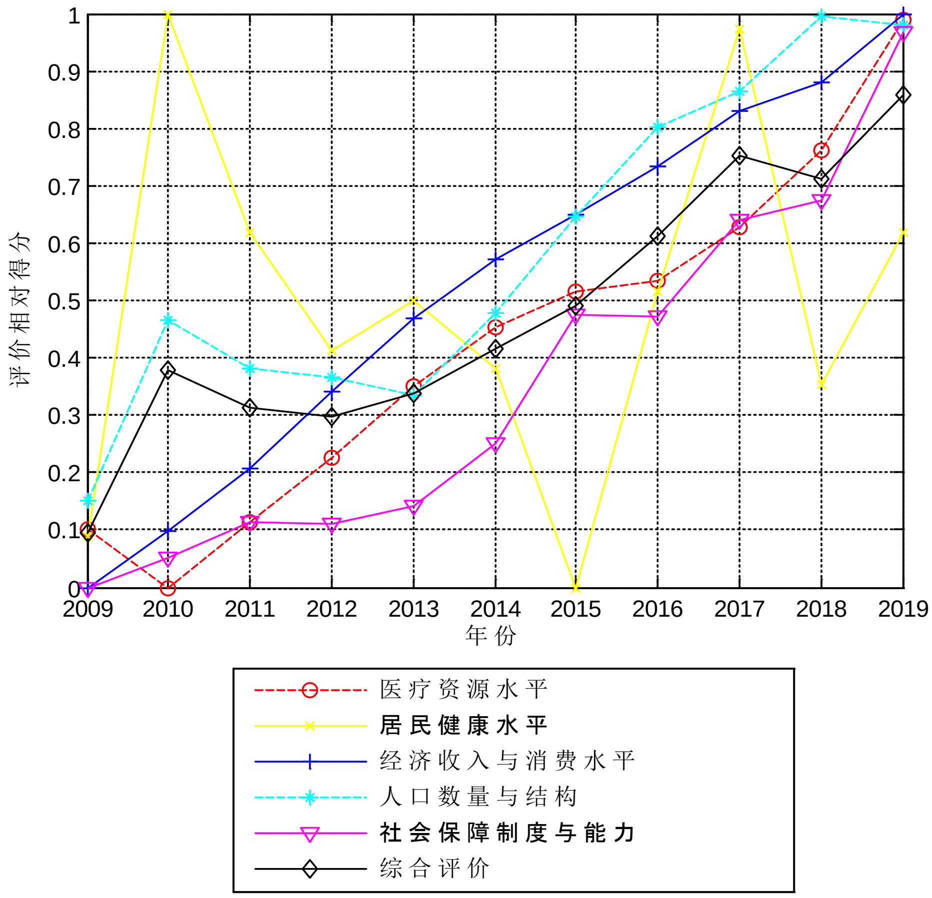

Figure 2. Line chart of comprehensive evaluation of medical and pension resources from 2009 to 2019
图 2. 2009~2019 年间医疗和养老资源综合评价折线图

5. 利用多元线性回归模型对人均预期寿命进行预测

从图 2 的分析可以看出，居民健康水平波动性较大，不稳定，而居民健康水平与 x_{21} 人均预期寿命、x_{22} 达到国民体质标准合格以上的人数比例、x_{23} 基本保险养老人数和 x_{24} 养老保险覆盖率相关，其中人均预期寿命是老百姓最为关心的指标。根据前面指标的相关性分析人均预期寿命主要受到人均医疗卫生事业费；每千人口医师数；每千人口床位数；达到国民体质标准合格以上的人数比例以及恩格尔系数有关。下面借助于未来 5 年、10 年和 15 年深圳医疗和养老保障需要实现的目标量化指标数据以及 2014~2019 年间数据用多元线性回归分析方法来研究人均预期寿命与上面 5 个指标的联系。

5.1. 多元线性回归分析预测

为方便期间下面用z表示因变量人均预期寿命；用 y_1 表示自变量人均医疗卫生事业费(元)；y_2 表示自变量每千人口医师数；y_3 表示自变量每千人口床位数；y_4 表示自变量达到国民体质标准合格以上的人数比例；y_5 表示自变量恩格尔系数。

基于表 2 中的数据，假设五个因素 $y_i\left(i=1,2,\cdots,5\right)$ 和人居预期寿命 z 值满足多元线性回归模型[10] [11]。则采用多元线性回归分析的模型为：

$$
\begin{aligned}
&z=\beta_0+\beta_1 y_1+\beta_2 y_2+\cdots+\beta_5 y_5+\varepsilon \\
&\varepsilon \sim N\left(0,\sigma^2\right)
\end{aligned} \tag{14}
$$

式中 $\beta_0,\beta_1,\cdots,\beta_5,\sigma^2$ 都是与 $y_i\left(i=1,2,\cdots,5\right)$ 无关的未知参数，这里 $\beta_0,\beta_1,\cdots,\beta_5$ 称为回归系数。

借助于已有数据，通过灰色系统预测方法得到10组独立观测数据 $\left(z_i,y_{i1},y_{i2},\cdots,y_{i5}\right)$，见表7。由(14)得

$$
\begin{aligned}
&z_i=\beta_0+\beta_1 y_{i1}+\beta_2 y_{i2}+\cdots+\beta_5 y_{i5}+\varepsilon_i \\
&\varepsilon_i \sim N\left(0,\delta^2\right),\ \left(i=1,2,\cdots,10\right)
\end{aligned} \tag{15}
$$

记

$$
Y=\begin{bmatrix} 1 & y_{11} & \cdots & y_{15} \\ 1 & y_{21} & \cdots & y_{25} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & y_{101} & \cdots & y_{105} \end{bmatrix},\ Z=\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{10} \end{bmatrix} \tag{16}
$$

$$
\varepsilon=\left[\varepsilon_1 \quad \cdots \quad \varepsilon_{10}\right]^{\mathrm{T}},\ \beta=\left[\beta_0 \quad \beta_1 \quad \cdots \quad \beta_{10}\right]^{\mathrm{T}}
$$

公式(15)用矩阵形式可以表示为：

$$
Z=Y\beta+\varepsilon,\ \varepsilon_i \sim N\left(0,\delta^2 E_n\right),\ \left(i=1,2,\cdots,10\right)
$$

其中 E_n 为 n 阶单位矩阵。

Table 7. Next 5, 10 and 15 years and known relevant data
表7. 未来5年、10年和15年和已知相关数据

年份	人均预期寿命	人均医疗卫生事业费(元)	每千人口医师数	每千人口床位数	达到国民体质标准合格以上的人数比例	恩格尔系数
2014	78.3	772.97	2.49	2.88	89.8	0.331
2015	79	960.87	2.55	3.35	88.5	0.32
2016	79.9	1080.45	2.57	3.49	90.25	0.305
2017	80.35	1245.38	2.66	3.5	92	0.3
2018	81.25	1465.45	2.79	3.65	89.7	0.292
2019	81.45	2496.35	3.01	3.83	90.6	0.294
2020	81.7	3009.6	3.12	4.0	89.905	0.279
2025	82.55	9541.1	3.95	4.91	90.395	0.242
2030	83.75	22960	5.02	5.66	89.750	0.209
2035	84.13	45824	6.33	6.11	91.467	0.181

数据来源：深圳市卫生健康委网站、《深圳统计年鉴 2019》、中国统计信息网。

5.2. 参数估计

模型(15)中的参数 $\beta_0,\beta_1,\cdots,\beta_5$ 用最小二乘法估计，即应选取估计值 $\hat{\beta}_j$，使当 $\beta_j=\hat{\beta}_j$，$j=1,2,\cdots,5$ 时，误差平方和

$$
Q=\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{10}\left(z_i-\beta_0-\beta_1 y_{i1}-\beta_2 y_{i2}-\cdots-\beta_5 y_{i5}\right)^2 \tag{17}
$$

达到最小。

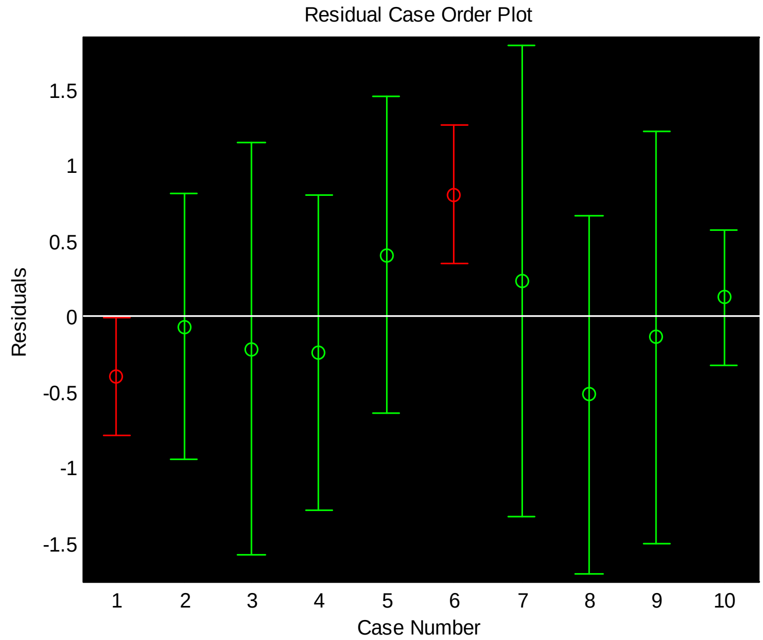

Figure 3. Residual diagram of original data
图 3. 原始数据残差图

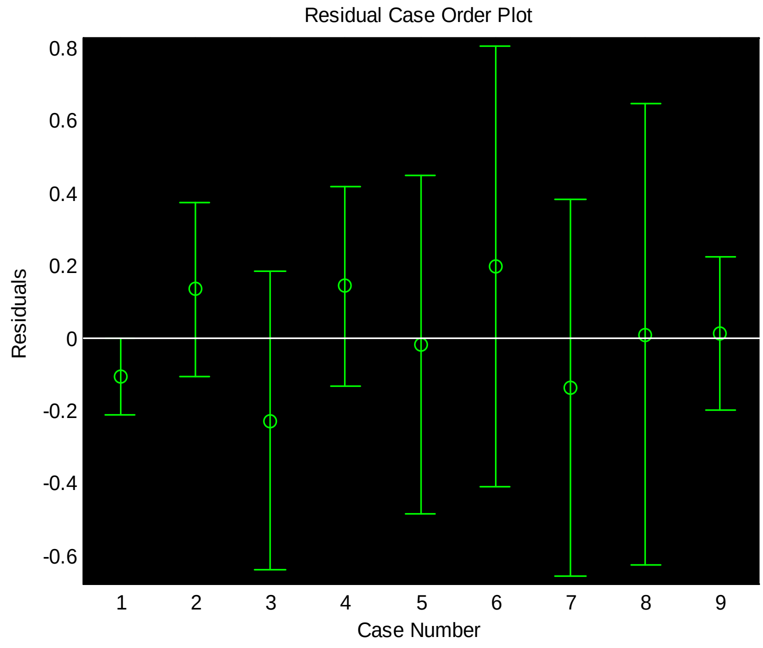

Figure 4. Residual diagram after removing outliers
图 4. 去掉异常点后的残差图

运用 MATLAB 软件运行程序，残差图 3 可以看出，除第一个和第六个数据外，其余数据的残差离零点均较近，且残差的置信区间均包含零点。去除第六个点后再次运行程序，从残差图 4 可以看出，数据的残差离零点均较近，且残差的置信区间均包含零点，可以认为第六个数据为异常点。

Table 8. Parameter values of multiple regression coefficients
表 8. 多元回归系数参数值

系数	系数估计值	系数置信区间
β_0	161.0511	[91.5622, 230.5399]
β_1	0	[−0.0002 , 0.0002]
β_2	−1.8413	[−5.3174, 1.6348]
β_3	−2.8335	[−6.4871, 0.8200]
β_4	−25.3407	[−64.8490, 14.1675]
β_5	−142.4757	[−226.9093, −58.0421]
$R^2 = 0.9951$; $F = 122.1485$; $p = 0.0012 < 0.01$; $s^2 = 0.0535$		

由图 4 和表 8 可以看出，残差图十分正常，该模型完全可用。

对应的回归方程为：

$$z = 161.0511 - 1.8413 y_2 - 2.8335 y_3 - 25.3407 y_4 - 142.4757 y_5 \tag{18}$$

从公式(18)可以看出，指标恩格尔系数前的绝对值系数最大，可以看出人居预期寿命受恩格尔系数的影响最大。一般来说，在其他条件相同的情况下，恩格尔系数较高，作为家庭来说则表明收入较低，作为国家来说则表明该国较穷。反之，恩格尔系数较低，作为家庭来说则表明收入较高，作为国家来说则表明该国较富裕。因此加快经济发展，增加广大城镇居民收入水平，是降低恩格尔系数的最重要手段。影响人居预期寿命第二大的指标就是达到国民体质标准合格以上的人数比例。要想提高人均预期寿命还需要更好的加强体育锻炼，并且要长久坚持。

借助于原始数据表中恩格尔系数的数据，利用 Matlab 软件编程，以年份为自变量，对恩格尔系数进行拟合，拟合图如图 5 所示。

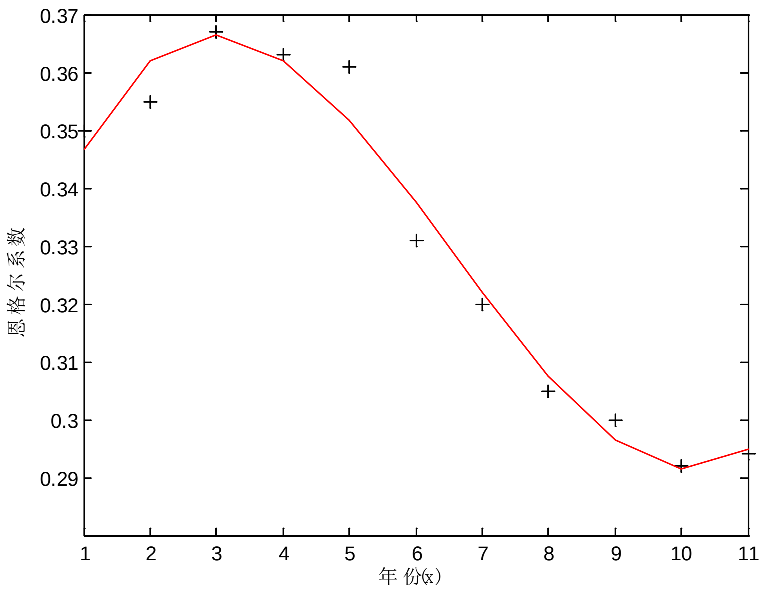

Figure 5. Fitting diagram of Engel's coefficient
图 5. 恩格尔系数拟合图

同时可得出残差为 0.0146；相关系数 0.9758；误差平方和 SSE 0.0002135。

效果较好，可以用该函数进行预测。拟合得到的函数关系为：

$$y_5 = 0.0004x^3 - 0.0079x^2 + 0.0364x + 0.3178 \tag{19}$$

由公式(1)、公式(2)、公式(3)、公式(18)以及公式(19)可得联立方程组：

$$\begin{cases} y_2 = 0.0002x^3 - 0.0005x^2 + 0.0629x + 2.0152 \\ y_3 = -0.0024x^3 + 0.0416x^2 - 0.0135x + 2.0988 \\ y_4 = 0.9025 + 0.003698\cos(0.6542x) + 0.00542\sin(0.6542x) + 0.0007838\cos(1.3084x) \\ \quad -0.004582\sin(1.3084x) - 0.00251\cos(1.9626x) - 0.01157\sin(1.9626x) \\ y_5 = 0.0004x^3 - 0.0079x^2 + 0.0364x + 0.3178 \\ z = 161.0511 - 1.8413y_2 - 2.8335y_3 - 25.3407y_4 - 142.4757y_5 \end{cases} \tag{20}$$

此方程组都是以“年份”作为自变量的，$x=1$ 对应 2009 年，$x=2$ 对应 2010 年，依此类推，$x=12$ 对应 2020 年，$x=17$ 对应 2025 年，$x=22$ 对应 2030 年。

方程组对 x 求导可得每个指标的年增长率。

$$\begin{cases} \dfrac{dy_2}{dx} = 0.006x^2 - 0.001x + 0.0629 \\ \dfrac{dy_3}{dx} = -0.0072x^2 + 0.0832x - 0.0135 \\ \dfrac{dy_4}{dx} = 0.0036\cos(0.6542x) - 0.0024\sin(0.6542x) - 0.006\cos(1.3084x) \\ \quad -0.001\sin(1.3084x) - 0.023\cos(1.9626x) + 0.005\sin(1.9626x) \\ \dfrac{dy_5}{dx} = 0.0012x^2 - 0.0158x + 0.0364 \\ \dfrac{dz}{dx} = -1.8413\dfrac{dy_2}{dx} - 2.8335\dfrac{dy_3}{dx} - 25.3407\dfrac{dy_4}{dx} - 142.4757\dfrac{dy_5}{dx} \end{cases}$$

合并同类项后可得，

$$\begin{cases} \dfrac{dy_4}{dx} = 0.0036\cos(0.6542x) - 0.0024\sin(0.6542x) - 0.006\cos(1.3084x) \\ \quad -0.001\sin(1.3084x) - 0.023\cos(1.9626x) + 0.005\sin(1.9626x) \\ \dfrac{dz}{dx} = -0.1616x^2 + 1.917x - 5.26367 - 25.3407\dfrac{dy_4}{dx} \end{cases}$$

由上述方程组可知，要想求出表 7 中未来 5 年、10 年和 15 年深圳医疗和养老保障需要实现的目标量化指标——人均预期寿命每年的增长率，只需要把相应的x值带进去即可。例如求人居预期寿命 2023 年的增长率，只需把 $x=15$ (2023 年对应的 x 为 15)带入上述方程组即可。通过 Mathematica 软件计算函数值，可得人均预期寿命 2023 年的年增长率为 0.0058。

6. 总结

为了对未来 5 年、10 年和 15 年深圳医疗和养老保障需要实现目标进行量化描述，本文首先以国际、国内主流观点的高频指标为重点，结合深圳市经济发展的实际情况，选取了人均预期寿命、恩格尔系数、每千人口医师数等指标，构建深圳市医疗和养老资源合理配置的评价体系。然后通过分析深圳市近十几年的各项数据，借助于熵权 TOPSIS 法对医疗和养老资源合理配置评价得分进行排序，并对目标实现过

程中遇到的困难给出合理的对策。通过分析可知深圳市建设“民生幸福标杆”和“可持续发展先锋”城市主要体现在两个方面：1) 大力发展经济，提高居民收入。2) 大力开展体育运动，加强体育锻炼，提高人均预期寿命。

熵权法是通过指标数据的变化大小确定指标权重的客观赋权方法。TOPSIS 方法主要是通过定义决策问题的正理想解和负理想解，然后计算决策变量与正负理想解的距离，再进一步计算各方案与理想解的相对贴近度，进行方案的优劣排序。熵权 TOPSIS 方法在评价问题建模中会有广泛的应用。利用该方法也可以评价深圳各区市医疗和养老保障的合理配置问题。

参考文献

[1] 2020 年“深圳杯”数学建模挑战赛 A 题——关于国家“先行示范区”建设中的医疗和养老保障问题[EB/OL]. http://www.m2ct.org/

[2] 中国统计信息网[EB/OL]. http://www.tjcn.org/

[3] 迟国泰，曹婷婷，张昆. 基于相关－主成分分析的人的全面发展评价指标体系构建[J]. 系统工程理论与实践, 2012, 32(1): 111-119.

[4] 卓金武. MATLAB 在数学建模中的应用[M]. 北京: 北京航空航天大学出版社, 2010.

[5] 周建兴. MATLAB 从入门到精通[M]. 北京: 人民邮电出版社, 2008.

[6] 杨惠敏，付萍. 基于熵权的多级模糊综合评价的应用[J]. 华北电力大学学报, 2005, 32(5): 104-107.

[7] 李刚，迟国泰，程砚秋. 基于熵权 TOPSIS 的人的全面发展评价模型及实证[J]. 系统工程学报，2011，26(3): 400-407.

[8] 龚纯，王正林. MATLAB 语言常用算法程序集[M]. 北京: 电子工业出版社, 2010.

[9] 司守奎，孙玺菁. 数学建模算法与应用[M]. 北京: 国防工业出版社, 2011.

[10] 汪应洛. 系统工程[M]. 北京: 机械工业出版社, 2008.

[11] 董大校. 基于 MATLAB 的多元非线性回归模型[J]. 云南师范大学学报, 2009, 29(2): 45-48.