

区域人口教育水平多因素 灰色决策模型

刘启明 陈 萍

随着科学技术的发展,描述和设计系统的理论和方法层出不穷,其焦点主要集中在解决系统描述和分析的精确性与系统的复杂性的矛盾上。信息论与控制论试图解决事物的随机性;模糊集合试图解决系统复杂性 with 精确性的矛盾,指出了事物的模糊性。然而系统除了随机性与模糊性外,还存在着一种范围更为广泛、内容更为深刻的特性,这就是系统的灰色性,即信息不完全和不确知现象。灰色系统理论认为,系统的随机性和模糊性是灰色性的两个不同方面的不确定性。以人口统计为例,指标中绝对精确数据很少,绝大多数是这个指标灰色区间值的白化值。如经常采用的人均指标。同时更由于人口普查收集来的各种统计、观测数据,由于技术方法、人为因素、自然环境等的变化和影响,造成各种误差、虚假、短缺等现象也在所难免。尽管有些数字还精确到小数点以后,但它所表示的仍然是一个信息不完全的灰色量。因此对很多人口现象的分析,传统的统计学方法存在无法弥补的缺陷。灰色系统分析用生成的方法对原始数据进行处理,可以弱化不确定性,强化规律性,因而可以用少量数据及存在灰色信息的情况下,建立系统的连续微分方程模型。与传统的模型相比,灰色建模不存在误差累积问题,因而用于人口分析、人口预测具有广阔前景。本文利用第四次人口普查1%的数据资料及10%抽样汇总资料,分析了反映中国人口文化教育水平的多项综合性指标,并以此为基础,建立了多区域评价的灰色聚类决策模型,最后对模型计算结果作了初步分析。

一、人口文化教育水平指标体系及初步分析

目前采用的衡量人口文化教育水平的指标,如平均受教育年限、平均受教育水平、文盲率等,受所研究地区人口的年龄结构、性别差异、城乡差异的影响,不利于做对比研究。我们希望文化教育指标既能消除年龄结构的影响,也能够从在学状况、升学状况、学业完成情况等方面考察。本文综合上述各因素选择下列五个综合指标来反映中国人口文化教育水平的空间差异。

(一) 15岁以上人口分性别年龄标准化文盲率(AIR^*)。在衡量发展中国家人口受教育程度时,对文盲半文盲的考察是极重要的一项。粗文盲率(定义为15岁以上文盲半文盲人口占总人口的比重)受到人口年龄结构的影响。因此本文在进行不同地区比较时,进行了年龄标准化处理,其计算公式为

$$AIR^* = \frac{\sum (I_i/P_i) \times P_i^*}{P^*} \times 100\% \quad (i \geq 15)$$

这里 I_i/P_i 为年龄别文盲率,即 i 岁文盲人口除以 i 岁人口, P_i^* 为 i 岁标准人口数, P^* 为标准总人口数。本文 P_i^* 和 P^* 取全国人口,并且算出了各省分性别的 AIR^* 。

(二) 城乡文盲率比值($RIR_{\frac{1}{2}}$)。该项指标可以反映一个地区城乡教育水平的差异,其

计算公式为
$$RIR_u = \frac{CIR_u}{CIR_r} \times 100\% \quad (RIR_u \leq 1)$$

其中: CIR_u 为城市粗文盲率, CIR_r 为农村粗文盲率。 RIR_u 越大, 说明城乡人口的文化水平差异性越小。

(三) 大学文盲比 (BIR)。由于各地区就读小学、初中、高中的人口比例较高, 因此小学、初中、高中文化程度人口比例在各地的差异并不太明显。大学文化程度人口的比例可以集中反映高等教育的水平, 文盲半文盲人口比例则集中反映了教育的普及程度。各地区的总体文化水平的差异可以从这两方面得到反映。大学文盲比是指每100名文盲、半文盲人口相对应的大学文化程度人口数, 计算公式为

$$BIR = \frac{P_b}{P_{15+}} \times 100\%$$

P_b 为具有大学文化程度的人口数, P_{15+} 为15岁及15岁以上文盲半文盲人口数。

(四) 6~14岁在学率 ($ASER_{6\sim14}$)。该指标主要反映小学、初中的在学情况, 公式为

$$ASER_{6\sim14} = \frac{E_{6\sim14}}{P_{6\sim14}} \times 100\%$$

$E_{6\sim14}$ 为6~14岁在学人口数, $P_{6\sim14}$ 为6~14岁总人口数。

(五) 平均受教育年限综合均值。该指标的 settings 是把文化程度这种按品质分组的指标转化为数量标志, 从而计算其均值。从计算上讲是以大学、大专、中专、高中、初中、小学人数占总人数的百分比为权数, 对不同文化程度的平均受教育年限做加权处理, 求出综合均值。

本文假定取得各级学历所需的受教育年限为大学16年、大专15年、中专高中均为12年, 初中9年、小学6年, 计算公式为

$$\bar{X} = \sum_{i=1}^6 X_i \cdot P_i$$

\bar{X} 为平均受教育年限, X_i 为*i*类教育的平均受教育年限, P_i 为具有*i*类文化程度人口占总人口的比例。

根据上述五项指标计算的30个省、市、自治区人口文化教育状况的指标值如表1所示。

从表1中我们可以看出:

1. 从标准化文盲率看, 男性最低的前3个地区是上海、天津、北京; 最高的3个地区是西藏、青海、甘肃。女性标准化文盲率最低的前3个地区是北京、上海、天津; 最高的3个地区仍然是西藏、青海、甘肃。从各地区看, 女性文盲率分别是男性的2~3倍, 其规律是经济发展水平较高的地区这一比值均较高。比值大于3的有广东、上海、天津、北京, 西藏、新疆的比值最低。这种现象一方面与各地区的文盲率整体水平有关, 另一方面也反映了在经济发达地区男女两性文盲率存在较大的差异。

2. 从城乡文盲率比值看, 城乡差异较小的几个省区是山东、浙江、广东、山西, 差异较大的是青海、宁夏、陕西、云南、西藏等边远地区。值得提出的是, 北京、天津、上海这种经济发达且文化程度综合水平很高的地区, 城、乡的发展水平差异性明显, 乡村地区的文化教育水平明显低于城市。

3. 从6~14岁在学率看, 基本上可分成三个层次, 指标值大于85%的有10个省区, 指标

表1

1990年30个省、市、自治区人口文化教育状况综合指标值

地区	标准化文盲率(男) 比值(%) 位次	标准化文盲率(女) 比值(%) 位次	城乡文盲率 比值(%) 位次	大学文盲比 比值(%) 位次	6~14岁在学率 比值(%) 位次	平均受教育年限综合 均值(年) 位次
北京	3.40 3	10.73 1	43.69 20	106.89 1	88.76 2	7.78 1
天津	3.32 2	11.87 3	53.84 11	52.31 3	85.68 9	7.05 3
河北	9.07 15	22.02 14	40.95 23	6.28 20	88.35 4	5.45 17
山西	7.92 13	17.10 6	60.19 4	12.24 9	88.21 5	5.96 7
内蒙古	10.38 21	22.78 17	43.42 21	9.58 11	79.29 17	5.71 10
辽宁	4.71 5	12.69 4	57.07 8	29.45 4	86.44 7	6.67 4
吉林	7.30 9	16.19 5	58.03 7	20.54 5	84.21 13	6.35 5
黑龙江	7.90 12	18.12 8	58.73 6	19.67 6	83.57 14	6.32 6
上海	3.11 1	11.80 2	39.94 24	59.20 2	88.57 3	7.54 2
江苏	7.89 11	22.31 15	50.80 12	8.55 15	90.08 1	5.77 9
浙江	8.53 14	22.66 16	64.41 2	6.70 19	75.80 23	5.51 14
安徽	14.65 24	34.56 24	45.36 18	3.61 26	76.96 19	4.63 25
福建	7.86 10	25.36 21	50.60 13	7.85 16	76.65 21	5.20 22
江西	9.54 19	27.66 23	41.99 22	6.11 21	74.19 24	5.17 23
山东	9.40 18	23.21 18	64.47 1	5.78 23	85.27 10	5.44 18
河南	6.26 7	18.17 9	44.10 19	5.25 25	86.81 6	5.56 16
湖北	9.08 16	23.41 20	49.38 14	9.91 10	86.16 8	5.56 13
湖南	10.38 20	23.26 19	49.34 15	9.41 13	84.46 11	5.67 11
广东	4.39 4	17.34 7	61.39 3	12.81 8	82.15 15	5.77 8
广西	5.66 6	18.96 10	46.30 16	7.46 17	76.76 20	5.38 20
海南	6.29 8	20.18 11	54.56 10	8.91 14	78.20 18	5.47 15
四川	9.12 17	21.71 13	59.96 5	5.92 22	79.35 16	5.36 21
贵州	16.55 25	40.97 28	55.54 9	3.20 28	66.00 27	4.19 28
云南	17.86 26	37.43 25	33.64 27	3.17 29	64.64 28	4.13 29
西藏	46.55 30	64.77 30	37.27 26	1.29 30	25.05 30	1.92 30
陕西	12.47 23	27.18 22	30.92 28	9.49 12	84.45 12	5.14 19
甘肃	20.47 28	40.31 27	38.70 25	3.95 27	71.48 26	4.30 27
青海	22.28 29	44.15 29	26.43 30	5.38 24	60.35 29	4.34 26
宁夏	18.83 27	39.24 26	29.75 29	7.29 18	72.66 25	4.91 24
新疆	11.98 22	20.33 12	45.62 17	14.46 7	76.30 22	5.61 12

资料来源: ①四普1%原始数据带。②《中国1990年人口普查10%抽样资料》, 中国统计出版社 1991. 7。

值位于75%~85%的有13个省区, 指标值小于75%的有7个省区。6~14岁在学率最好的是江苏省(90.08%), 最差的仍是西藏(25.05%), 其余省区均处于60%~90%之间。

4. 大学文盲比与平均受教育年限综合均值的评价结果具有较好的一致性, 说明两者均具有较好的综合性。

以上五个指标从不同方面衡量了人口的文化教育水平。从计算结果可以看出, 五种指标的评价结果(排序)有所不同。为了避免单一指标所造成的偏差, 切合实际地综合评价人口文化教育状况, 文章在下一节引用了灰色聚类方法来达到这个目的。

二、灰色聚类决策模型及其结果评价

1982年, 中国著名学者邓聚龙教授创立了灰色系统理论。短短几年中, 灰色系统理论以其强大的生命力立于科学之林, 奠定了作为一门新的交叉学科的学术地位。其理论成果已在实践中获得精彩的应用和理论充实。灰色系统理论主要研究系统内部结构、要素及要素之间

关系中未知部分的理论。灰色聚类决策用于按照多个不同的决策指标对决策对象进行综合评价，以确定决策对象是否满足给定的取舍准则。其指标变权的特性，使其区别于一般聚类方法，具有决策的功能。

(一) 灰色聚类决策模型建模步骤

1. 几个基本概念：灰类、灰数白化与白化权函数

(1) 灰数：我们把只知道大概的范围而不知其确切值的数称为灰数。在应用中，灰数实际上指一个区间或一个一般的数集。灰色系统用灰色数、灰色方程、灰色矩阵等来描述。其中灰数是灰色系统的基本“单元”或“细胞”。

(2) 灰数白化：灰数白化即把灰数变为白化。对于一般的区间灰数 $\otimes \in [a, b]$ ，我们将白化值 $\hat{\otimes}$ 取为 $\hat{\otimes} = \alpha \cdot a + (1 - \alpha) \cdot b$ ， $\alpha \in [0, 1]$ 。

(3) 白化权函数：描述一个灰数对其取值范围内数值的分布程度的函数叫白化权函数。一般说来，一个灰色的白化权函数是研究者根据已知信息设计的，有一定的参考模式。

2. 数学方法概述

记 $k=1, 2, \dots$ 为典型类别； $i=I, II, \dots$ 为聚类元素； $j=1^*, 2^*, \dots$ 为聚类指标。

d_{ij} 表示第 i 类元素关于第 j 类指标的实际样本值， $i \in \{I, II, \dots\}$ ， $j \in \{1^*, 2^*, \dots\}$ 。所谓灰色聚类决策，即区分聚类元素在聚类指标下的所属类别。

3. 计算步骤

(1) 将聚类指标 j ， $j \in \{1^*, 2^*, \dots\}$ ，针对典型类别 k ， $k \in \{1, 2, \dots\}$ ，规定白化权函数。

如第 k 类对 j 个指标规定白化权函数 Y_{kj} 为

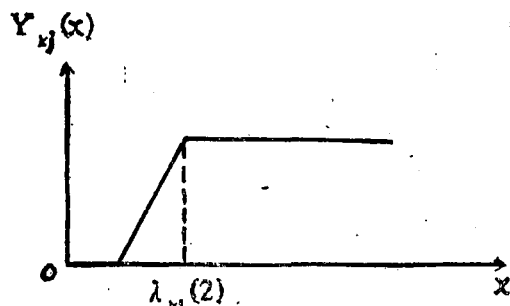


图1 白化权函数分布

$$Y_{kj}(x) = \begin{cases} a_{kj}(x), & x \in (0, \lambda_{kj}(2)) \\ Y_{kj}(s), & x \in (\lambda_{kj}(2), \infty) \end{cases}$$

式中 x 亦常用 d_{ij} 或 d_i 等符号表示。

(2) 定标准权

从白化权函数看对应 $Y_{kj}(s)$ 有特定的 $\lambda_k(2)$ (见图1)，为此有下述变换

$$T_i: \{Y_{kj}(s)\} \rightarrow \{\lambda_{kj}(2)\}$$

式中 $\{Y_{kj}(s)\}$ 为 $Y_{kj}(s)$ 的全体， $\{\lambda_{kj}(2)\}$ 为 $\lambda_{kj}(2)$ 的全体。规定标准权为

η_{kj}

$$\eta_{kj} = \frac{\lambda_{kj}(2)}{\sum_{j=1^*} \lambda_{kj}(2)} \quad k=1, 2, \dots, m; j=1^*, 2^*, \dots, n$$

(3) 确定实际权

求出决策系数 σ_{ik} ， $i \in \{I, II, \dots\}$ ， $k \in \{1, 2, \dots, m\}$

$$\sigma_{ik} = \sum_{j=1}^{n*} Y_{kj}(d_{ij}) \eta_{kj}$$

构造决策系数向量 $\sigma(i) = \{\sigma_{i1}, \sigma_{i2}, \dots, \sigma_{in}\}$

(4) 划分类别, 在 $\sigma(i)$ 中求最大的元素 σ_{ik}^*

$$\sigma_{ik}^* = \max_k \{\sigma_{ik}\}$$

从 σ_{ik}^* 中判断聚类元素 i 属于第 k 个典型类别。在实际问题中, 常常会遇到多个决策对象同属于某一个决策灰类, 而该灰类所能容纳的对象个数又有一定限额的情况, 此时可根据决策系数的大小决定取舍。设

$$\max_{1 \leq k \leq s} \{\sigma_{i1}^k\} = \sigma_{i1}^{k*}, \quad \max_{1 \leq k \leq s} \{\sigma_{i2}^k\} = \sigma_{i2}^{k*}, \quad \sigma_{i1}^{k*} > \sigma_{i2}^{k*},$$

则称在灰类 k^* 中, 聚类元素 $i1$ 优于聚类元素 $i2$ 。

(二) 决策指标与白化权函数的确定

根据前文所述, 我们选择5类6项指标, 综合评价中国30个省、自治区、直辖市人口文化教育综合水平。

6个聚类指标 ($j=1, 2, \dots, 6$) 分别为男性标准化文盲率、女性标准化文盲率、城乡文盲率比值、大学文盲比、6~14岁在学率、平均受教育年限综合均值;

聚类元素 ($i=1, 2, \dots, 30$) 为30个省、自治区、直辖市;

灰类 ($k=1, 2, \dots, 5$) 按人口文化教育水平分为高、较高、中等、中下、低5类。

在全面考察各项指标差异的基础上确定白化权函数 $Y_{kj}(x)$, 按照 $Y_{kj}(x)$ 图型规定 λ_{kj} (如表2)。

根据计算机分析 结果, 分别计算最大 决策系数 (σ_{ik}^*) 及 其隶属灰类 (k), 计算结果如表3所示。	表2 文化教育水平灰类标准					
	指标1	指标2	指标3	指标4	指标5	指标6
灰类1	< 4	<15	>50	>20	>85	>7
灰类2	< 7左右	<20左右	>45左右	>10左右	>80左右	>6左右
灰类3	<10左右	<25左右	>40左右	> 7左右	>75左右	>5.6左右
灰类4	<15左右	<30左右	>35左右	> 5左右	>60左右	>5.5左右
灰类5	>20	>45	<30	< 4	<65	<5

第一, 区域发展的高度差异性。从三个地带划分看, 西部地区人口文化教育水平过低, 与中部、东部地区相比, 出现了一个“低谷”现象。从类型表现看, 四类地区为空白。西藏、青海、云南、宁夏、甘肃、贵州这些地区人口文化素质的低下, 将成为其经济脱贫的本质障碍。经济发展和人口文化素质在这些地区形成一种“马太效应”。人口素质差, 不仅造成了目前落后地区落后的结果, 而且还可能是造成将来落后地区更落后的原因。落后地区居民与发达地区居民人口文化素质的明显差距, 特别是这种差距日趋继续扩大的倾向最终预示着产值、国民收入等经济指标差异继续扩大的不良前景。

第二, 人口文化素质的地域差异, 基本上和沿海、内地、边远地区的经济梯度相吻合。沿海省区除福建外, 全部属于一二大类, 且以一类为主。中部内地省区人口文化水平全部属于一二两大类。北京、山西、黑龙江、吉林为一类, 其余均为二类。边远6省区除新疆、内蒙属二类外, 其余宁夏、青海、甘肃、西藏4省区均为人口文化素质最低的类别。由此可见, 人

表3

中国1990年30个省、自治区、直辖市人口文化教育水平综合评价分类结果

地区	σ_{i1}	σ_{i2}	σ_{i3}	σ_{i4}	σ_{i5}	σ^*_{ik}	隶属灰类(k)
北京	1.0000	0.8787	0.8172	0.6868	0.5771	1.0000	1
上海	0.9979	0.8750	0.8133	0.6824	0.5723	0.9979	1
天津	0.9839	0.8978	0.8397	0.7121	0.6046	0.9838	1
辽宁	0.9816	0.8861	0.8341	0.7058	0.5978	0.9816	1
吉林	0.9736	0.9028	0.8481	0.7216	0.6149	0.9736	1
黑龙江	0.9706	0.8781	0.8508	0.7250	0.6187	0.9706	1
广东	0.9596	0.9150	0.8361	0.7034	0.6010	0.9596	1
山西	0.9578	0.9051	0.8314	0.6960	0.5878	0.9578	1
山东	0.9147	0.9047	0.8886	0.8689	0.7389	0.6441	1
内蒙古	0.8896	0.9716	0.9545	0.8285	0.7298	0.9716	2
广西	0.9074	0.9678	0.9345	0.8040	0.6976	0.9678	2
海南	0.9225	0.9568	0.9030	0.7678	0.6607	0.9568	2
新疆	0.9034	0.9545	0.9280	0.8275	0.7304	0.9545	2
河南	0.9199	0.9500	0.9034	0.7817	0.6721	0.9509	2
湖南	0.9194	0.9482	0.9196	0.7894	0.6864	0.9482	2
湖北	0.9258	0.9478	0.9107	0.7765	0.6744	0.9478	2
江苏	0.9309	0.9345	0.8885	0.7515	0.6415	0.9345	2
四川	0.9032	0.9230	0.8981	0.7714	0.6620	0.9230	2
浙江	0.8948	0.8998	0.8996	0.7663	0.6738	0.8998	2
江西	0.8462	0.9244	0.9752	0.8832	0.7854	0.9752	3
福建	0.8933	0.9421	0.9451	0.8186	0.7153	0.9451	3
河北	0.8943	0.9344	0.9352	0.8095	0.7024	0.9352	3
陕西	0.8481	0.8981	0.9118	0.8518	0.7889	0.9118	3
安徽	0.8203	0.8938	0.9116	0.8677	0.8196	0.9116	3
西藏	0.6716	0.7226	0.7873	0.8787	0.9906	0.9906	5
青海	0.4492	0.4905	0.5264	0.5738	0.9695	0.9695	5
云南	0.7294	0.7968	0.8551	0.9248	0.9441	0.9441	5
宁夏	0.7482	0.8173	0.8751	0.8704	0.9751	0.9751	5
甘肃	0.7520	0.8214	0.8815	0.8732	0.9166	0.9166	5
贵州	0.7748	0.7968	0.8209	0.8376	0.8576	0.8576	5

资料来源: 同表1。

口文化素质的提高决非仅仅是引进和投资的问题, 其发展水平是与经济发展相适应的。从某种角度而言, 人口文化素质的提高取决于经济发展对它提出的要求。因此理顺人口质量与经济发展的各种关系, 使教育结构与经济发展的需求相适应, 成为问题的实质所在。

主要参考文献

1. 邓聚龙: 《灰色系统(社会、经济)》, 北京国防工业出版社, 1985; (本文责任编辑: 洪映)
2. 邓聚龙: 《灰色局势决策》, 《模糊数学》, 1985(2); (作者工作单位: 中国社科院人口所)
3. 邓聚龙: 《灰色控制系统》, 华中工学院出版社, 1985。

编者注: 本刊1992年第1期刊载《独生子女证在中国河北省接受情况及其后果(1979~1988)》一文, 作者: (美) R·库妮、M·鲍尔丝、(中国)魏进, 原文发表于“Population Research and Policy Review” 1991年第十卷第137~155页。本刊已得到荷兰 Kluwer Academic 出版社的许可登载该文中英文版。特此注明。