

文章编号:1672-058X(2014)12-0001-04

城镇居民消费结构实证分析*

——以西部地区为例

王艳艳

(重庆师范大学 数学学院,重庆 401331)

摘要:应用因子分析模型,选取反映城镇综合消费水平的 8 个指标作为原始变量,运用 SPSS 软件,对西部 12 个省区市的城镇居民消费水平进行了综合评价和排序,并针对西部 12 个省区市消费水平特点,提出一些建议.

关键词:因子分析;西部地区;消费水平;综合评价

中图分类号:F224

文献标志码:A

1 评价方法

1.1 因子分析模型

因子分析是通过原始数据相关系数矩阵内部结构的研究,将多个指标转化为少量互不相关且不可观测的随机变量(即主因子),以提取原有指标绝大部分信息的统计方法.当因子载荷矩阵结构不够简化时,可以通过因子旋转使因子具有更为鲜明的实际意义,同时使用因子得分函数对样本给出相应的评价和排序.

设有 n 个样品,每个样品观测 p 个指标(变量), X_1, X_2, \dots, X_p ,得到原始数据矩阵 $\mathbf{X} = (X_1, X_2, \dots, X_p)^T$. 因子分析的数学模型为 $\mathbf{X} = \mathbf{A}\mathbf{F} + \boldsymbol{\varepsilon}$,其中 $\mathbf{F} = (F_1, F_2, \dots, F_m)^T$ 为 \mathbf{X} 的公共因子, \mathbf{A} 为因子载荷矩阵, $\boldsymbol{\varepsilon}$ 为特殊因子.论文使用主成分因子提取方法,其特点在于可以用方差贡献值 β_i 衡量第 i 个公因子的重要程度.

1.2 因子分析具体步骤^[1-3]

将原始数据标准化,为书写方便,仍为 \mathbf{X} .为了使综合评价的结果客观、合理,必须消除数量级和量纲不同带来的影响.采用标准化处理,建立指标的相关系数矩阵 \mathbf{R} ;计算特征方程 $|\mathbf{R} - \lambda\mathbf{E}| = 0$,得到相关系数矩阵 \mathbf{R} 的特征根 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0, \mu_1, \mu_2, \dots, \mu_p$ 为相应的标准正交化特征向量;根据累计贡献率的要求,譬如 $\sum_{i=1}^m \lambda_i / \sum_{i=1}^p \lambda_i \geq 85\%$ 时,取前 m 个主成分代替原来的 p 个指标,计算因子载荷矩阵 \mathbf{A} ;对 \mathbf{A} 施行方差最大正交旋转,建立因子模型的目的不仅在于找到公共因子,更重要的是知道每一个公共因子的意义;对主因子进行命名和解释,如需进行排序,则计算各个主因子的得分 $F_i = \alpha_i x$,以贡献率为权重,对 F_i 加权计算综合因子得分.

收稿日期:2014-04-19;修回日期:2014-05-09.

* 基金项目:国家社科基金项目(13BTJ008).

作者简介:王艳艳(1988-),女,河南开封人,硕士研究生,从事概率论与数理统计研究.

2 实证分析

2.1 消费水平评价指标与数据

选取了反映城镇居民综合消费水平的 8 个指标,分别为 X_1 人均食品支出(元/人)、 X_2 人均衣着支出(元/人)、 X_3 人均住房支出(元/人)、 X_4 人均家庭设备及服务支出(元/人)、 X_5 人均交通和通信支出(元/人)、 X_6 人均文教娱乐用品及服务支出(元/人)、 X_7 人均医疗保健支出(元/人)、 X_8 其他商品及服务支出(元/人).应用多元统计学中的因子分析法,选取中国统计年鉴中 2011 年的数据,对中国西部地区包括重庆、四川、贵州、云南、西藏自治区、陕西、甘肃、青海、宁夏回族自治区、新疆维吾尔自治区、内蒙古自治区、广西壮族自治区等经济情况进行实证分析^[5].

表 1 2011 年西部 12 个省(直辖市、自治区)城镇居民平均每人消费支出 元

地区	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
重庆	5 847.90	2 056.79	1 205.66	1 079.27	1 718.73	1 474.88	1 050.62	540.63
四川	5 571.69	1 483.54	1 226.14	1 020.16	1 757.52	1 369.47	735.26	532.52
贵州	4 565.85	1 209.88	1 102.99	857.55	1 395.28	1 331.43	578.33	311.57
云南	4 802.26	1 587.18	827.84	570.46	1 905.86	1 350.65	822.41	381.38
西藏	5 184.18	1 261.29	781.12	428.03	1 278.00	514.44	424.105	527.74
陕西	5 040.47	1 673.24	1 193.81	914.26	1 502.44	1 857.6	1 100.51	500.42
甘肃	4 182.47	1 470.26	1 139.85	660.48	1 289.80	1 158.30	874.055	413.37
青海	4 260.27	1 394.28	1 055.15	723.23	1 293.45	967.90	854.25	406.93
宁夏	4 483.44	1 701.73	1 247.14	885.36	1 637.61	1 441.18	978.125	521.47
新疆	4 537.46	1 715.94	888.16	791.43	1 377.67	1 122.18	912.99	493.56
内蒙古	4 962.40	2 514.09	1 418.60	1 162.87	2 003.54	1 812.07	1 239.36	765.13
广西	5 074.49	1 019.34	1 237.91	884.85	2 000.57	1 502.65	779.08	349.48

2.2 判断数据是否适合因子分析

各变量应该具有相关性,如果变量间彼此独立,则无法从中提取公因子,也就谈不上因子分析法的应用.在 SPSS 中,KMO 和球形 Bartlett 检验用于因子分析的适用性检验,可以通过 Bartlett 球形检验来判断,如果相关阵是单位阵,则各变量相互独立,因子分析法无效.KMO 检验用于检查变量间的偏相关性,

取值在 0~1 之间,KMO 统计量越接近 1,变量间的偏相关性越强,因子分析的效果越好;当 KMO 统计量在 0.5 以下时,不适合应用因子分析法,应考虑重新设计变量结构或者采取其他统计分析方法.表 2 给出了因子分析 KMO 和 Bartlett 的检验结果.得知,原变量适合进行因子分析.

变量共同度是表示各变量中所含原始信息能被提取的公因子所表示的程度.由表 3 可知,几乎所有变量共同度都在 80% 以上,因此,提取出的这几个公因子对各变量的解释能力是较强的.

2.3 因子分析的数据结果及分析

图 1 被称为碎石图,用于显示各因子的重要程度,其横轴为因子序号,纵轴表示特征根大小.它将特征根从大到小排序,从而,可以很直观地看出哪些是主因子.

按照特征根大于 1 的原则,选入 3 个公共因子,其累计贡献率为 86.598%,特征根及累计贡献率见表 4.

表 5 为建立的因子载荷阵.此时得到的未旋转的公共因子的实际意义不好解释,为此,对公共因子施行方差最大化正交旋转,表 6 为经旋转得到的因子载荷阵,表 7 为因子载荷阵通过方差最大化正交旋转后得到的各主因子的方差贡献率.由旋转后的因子载荷阵表 6 可观察出,第三主因子 F_3 仅在 X_1 (人均食品支出)上的载荷较大,在此因子上的得分反映了人们最基础的消费情况;第二主因子 F_2 在 X_2 (人均衣着支出)、 X_8 (其他商品及服

表 2 KMO 和 Bartlett 的检验结果

取样足够度的 Kaiser-Meyer-Olkin	0.671
Bartlett 的球形度检验 近似卡方	58.882
Df	28
Sig0	0.001

务支出)上的载荷较大,在此因子上的得分反映了生活水平的提高;第一主因子 F_1 在 X_3 (人均住房支出)、 X_4 (人均家庭设备及服务支出)、 X_5 (人均交通和通信支出)、 X_6 (人均文教娱乐用品及服务支出)、 X_7 (人均医疗保健支出)上的载荷值较大,主要反映了生活消费水平进一步的提高.随着西部大开发,经济的快速增长,人们在住房、家庭设备及服务、交通和通信、文教娱乐用品及服务及医疗保健方面的消费越来越高.

表 3 共同度

	初始	Extraction
食品	1.000	0.923
衣着	1.000	0.942
住房	1.000	0.798
家庭设备	1.000	0.850
交通和通信	1.000	0.710
文教娱乐	1.000	0.898
医疗保健	1.000	0.891
其他	1.000	0.917

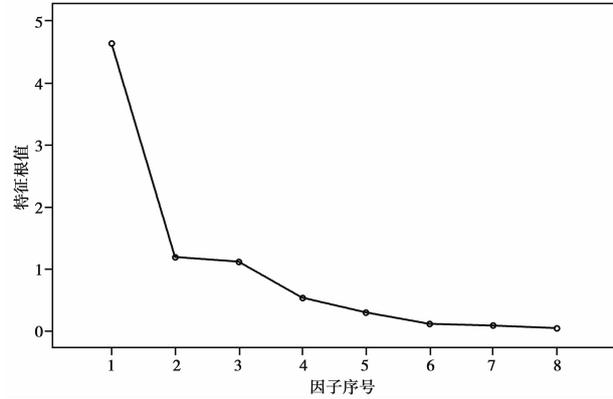


图 1 碎石图

表 4 因子变量的特征值、方差贡献率以及方差累积率

因子	初始特征值			提取公因子特征值		
	特征值	方差贡献率/%	累计贡献率/%	特征值	方差贡献率/%	累计贡献率/%
1	4.630	57.872	57.872	4.630	57.872	57.872
2	1.185	14.816	72.688	1.185	14.816	72.688
3	1.113	13.910	86.598	1.113	13.910	86.598
4	.535	6.690	93.288			
5	.297	3.714	97.002			
6	.111	1.389	98.390			
7	.088	1.094	99.484			
8	.041	.516	100.00			

表 5 因子载荷阵

	因子		
	1	2	3
食品	.408	.537	.685
衣着	.795	-.496	.253
住房	.815	.138	-.337
家庭设备	.900	.167	-.110
交通和通信	.672	.504	.069
文教娱乐	.848	.201	-.372
医疗保健	.843	-.359	-.225
其他	.689	-.426	.511

表 6 旋转后因子载荷阵

	因子		
	1	2	3
食品	.079	.155	.945
衣着	.333	.909	.068
住房	.861	.223	.078
家庭设备	.806	.350	.279
交通和通信	.623	.062	.564
文教娱乐	.924	.183	.103
医疗保健	.679	.635	-.164
其他	.124	.912	.265

表 7 旋转后公因子特征值、方差贡献率、累计贡献率的提取

特征值	方差贡献率/%	累计贡献率/%
3.227	40.337	40.337
2.294	28.679	69.016
1.407	17.582	86.598

2.4 因子得分和综合得分

以各公共因子的方差贡献率占 3 个公共因子总方差贡献率的比重作为权重进行加权汇总,得出各地区的综合得分情况,并且进行排名.综合得分模型:

$$F = (40.377 \times F_1 + 28.679 \times F_2 + 17.582 \times F_3) / 86.598$$

表 8 因子得分、综合得分及其排名表

地区	F_1	F_2	F_3	F	排名
重庆	0.290 51	0.803 99	1.357 99	0.68	2
四川	0.198 92	-0.240 49	1.491 57	0.32	4
贵州	0.210 80	-1.371 02	-0.450 30	-0.45	9
云南	-0.368 54	-0.496 43	0.273 36	-0.28	7
西藏	-2.563 97	0.187 63	0.889 68	-0.95	11
陕西	0.798 44	0.211 80	-0.295 63	0.38	3
甘肃	-0.151 10	-0.143 29	-1.536 04	-0.43	8
青海	-0.429 64	-0.163 44	-1.283 85	-0.51	10
宁夏	0.543 72	0.278 54	-0.721 78	0.20	5
新疆	-0.736 49	0.642 66	-0.733 64	-0.28	7
内蒙古	1.140 36	2.082 76	0.188 81	1.26	1
广西	1.066 99	-1.792 72	0.819 83	0.07	6

表 9 我国西部地区城镇居民消费情况分类表

类别	西部地区省、直辖市、自治区名称
第一类	内蒙古、重庆
第二类	陕西、四川、宁夏、广西
第三类	云南、新疆、甘肃、贵州、青海、西藏

3 结 论

对各地区在 3 个公共因子上的得分进行加权综合,就得到了综合得分.再根据统计软件 SPSS 进行 ward 聚类分析,由综合得分评价西部地区的城镇居民的消费水平,结果如表 9.

我国西部地区幅员辽阔,各个省份(直辖市、自治区)生产力发展水平、人口状况、地理环境、民族构成、交通设施的差异很大,各地区受地域特征和历史文化不同所形成的消费观念和消费习俗不同,形成了各个地区消费水平不平衡的显著特征.通过了解各地区的消费水平,可以采取相应的发展策略和技术水平,制定扩大居民消费、拉动地方经济发展的政策措施,这便是进行区域划分要解决的问题.

参考文献:

- [1] 克劳斯.巴克豪斯,本德.埃里克森,伍尔夫.普林克.多元统计分析方法:用 SPSS 工具[M].上海:格致出版社,2009
- [2] 方开泰.实用多元统计分析[M].上海:华东师范大学出版社,1989
- [3] 张尧庭,方开泰.多元统计分析引论[M].北京:科学出版社,1982
- [4] 国家统计局.中国统计年鉴 2012 [M].北京:中国统计出版社,2012
- [5] 徐科,张艳.重庆各区县经济发展水平的因子分析[J].重庆工商大学学报:自然科学版,2012,29(1):42-47

Empirical Analysis of Urban Residents Consumption Structure Taking West Part of China as an Example

WANG Yan-yan

(School of Mathematics, Chongqing Normal University, Chongqing 401331, China)

Abstract: By using factor analysis model, via taking 8 indicators reflecting urban comprehensive consumption level as primitive variables and through SPSS software, this paper makes comprehensive evaluation and ranking for urban resident consumption level in 12 west provinces and municipality and advances some suggestions according to the characteristics of consumption level in 12 west provinces and municipality.

Key words: factor analysis; west area; consumption level; comprehensive evaluation

责任编辑:李翠薇