

# 多元统计分析方法在我国各地区消费数据中的应用

## Application of Multivariate Statistical Analysis Method in Consumption

### Data of Various Regions in my country

卢庆钧

Qingjun Lu

广西大学应用数学专业 中国·广西 南宁 530004

Guangxi University Applied Mathematics Nanning City, Guangxi Zhuang Autonomous Region, 530004, China

**摘要:** 改革开放以来,随着经济不断发展,社会不断进步,我国居民的消费结构和消费水平也不断变化。各个地区之间经济发展的差异,也反映在消费领域中。本文主要分析了全国 31 个地区的消费水平和消费结构。利用多元统计分析中学习到的方法,采用 SPSS 进行数据分析,并且引入经济学的一些观点和概念,对于我国 31 个地市的居民消费情况进行了分析和对比。主要得到了这 31 个省市的居民消费情况和 8 个关键的居民消费指标之间的联系,并且采用因子得分的方法,计算并排列了消费情况的综合得分,得出了各个地区消费水平的排序并且进行了分析。然后用聚类分析的方法对我国 31 个地区的居民消费情况做了大致分类,得到了这些地区的消费结构。随后对分类后的情况做出了分析和总结,最后结合了相关文献资料和经济学知识得到了一些意见和看法。

**Abstract:** Since the reform and opening up, with the continuous development of the economy and the continuous progress of the society, the consumption structure and consumption level of Chinese residents have also been changing. The differences in economic development between various regions are also reflected in the consumption field. This article mainly analyzes the consumption level and consumption structure of 31 regions across the country. Using the methods learned in multivariate statistical analysis, using SPSS for data analysis, and introducing some viewpoints and concepts of economics, the consumption situation of residents in 31 cities in my country was analyzed and compared. Mainly obtained the relationship between the consumption situation of the residents in these 31 provinces and cities and 8 key consumption indicators, and used the factor scoring method to calculate and arrange the comprehensive score of the consumption situation, and obtained the consumption level of each region. Sorted and analyzed. Then use the cluster analysis method to roughly classify the consumption situation of residents in 31 regions of our country, and get the consumption structure of these regions. Afterwards, it analyzed and summarized the situation after classification, and finally obtained some opinions and opinions by combining relevant literature and economics knowledge.

**关键词:** 多元统计;分析方法;消费数据

**Key words:** multivariate statistic; analysis methods; consumption data

**DOI:** 10.36012/emr.v2i5.2539

**【作者简介】**卢庆钧(1999~),男,湖南郴州人,本科在读。

## 1 引言

消费作为宏观经济学中的一个重要概念,对于经济情况的描述和分析起到了重要的作用。消费作为社会再生产的最终阶段,为生产提供了方向和动力,消费是生产的目的和导向,没有消费,生产就没有意义。消费促进货币流通,促进地区经济发展,提供就业岗位,推动科技进步,提高了人民生活水平。

所以,对我国各地区的消费情况的分析,可以用于研究我国各地区的消费水平和消费结构。对不同地区的消费水平的分析可以反映出我国消费链条和经济发展之间的关系,同时,对不同地区的消费结构的研究还可以反应我国各地区的经济发展情况,促进产业结构协调发展,拉动投资,促进就业,对于我国宏观经济的发展也可以起到重要作用。

本文引用了2018年统计年鉴中八项主要消费指标,食品烟酒、衣着、居住、生活用及服务、交通通信、教育文化娱乐、医疗保健、其他用品及服务,这八项指标涵盖了居民社会生活的方方面面,兼顾了不同地区,不同经济情况的消费状况,通过分析,整合,对比这些数据,就可以大致分析得到我国的总体和不同地区的消费水平,消费结构和各地区的相同的以及不同点。在此基础上结合经济学知识和相关资料文献,对我国的消费情况和经济情况做出大致判断。

## 2 数据概述

6-20 分地区居民人均消费支出(2018年)

单位:元

| 地区  | 消费支出    | 食品烟酒    | 衣着     | 居住      | 生活用品及服务 | 交通通信   | 教育文化娱乐 | 医疗保健   | 其他用品及服务 |
|-----|---------|---------|--------|---------|---------|--------|--------|--------|---------|
| 全国  | 19853.1 | 5631.1  | 1288.9 | 4646.6  | 1222.7  | 2675.4 | 2225.7 | 1685.2 | 477.5   |
| 北京  | 39842.7 | 8064.9  | 2175.5 | 14110.3 | 2371.9  | 4767.4 | 3999.4 | 3274.5 | 1078.6  |
| 天津  | 29902.9 | 8647.5  | 1990.0 | 6406.3  | 1818.4  | 4280.9 | 3186.6 | 2676.9 | 896.3   |
| 河北  | 16722.0 | 4271.3  | 1257.4 | 4050.4  | 1138.7  | 2355.4 | 1734.5 | 1540.5 | 373.8   |
| 山西  | 14810.1 | 3688.2  | 1261.0 | 3228.5  | 855.6   | 1845.2 | 1940.0 | 1635.1 | 356.4   |
| 内蒙古 | 19665.2 | 5324.3  | 1751.2 | 3680.0  | 1204.6  | 3074.3 | 2245.4 | 1847.5 | 537.9   |
| 辽宁  | 21398.3 | 5727.8  | 1628.1 | 4169.5  | 1259.4  | 2968.2 | 2708.0 | 2257.1 | 680.2   |
| 吉林  | 17200.4 | 4417.4  | 1397.0 | 3294.8  | 899.4   | 2479.7 | 2193.4 | 2012.0 | 506.7   |
| 黑龙江 | 16994.0 | 4573.2  | 1405.4 | 3176.3  | 886.4   | 2196.6 | 2030.3 | 2235.3 | 490.4   |
| 上海  | 43351.3 | 10728.2 | 2036.8 | 14208.5 | 2095.5  | 4881.2 | 5049.4 | 3070.2 | 1281.5  |
| 江苏  | 25007.4 | 6529.8  | 1541.0 | 6731.2  | 1493.3  | 3522.8 | 2582.6 | 2016.4 | 590.4   |
| 浙江  | 29470.7 | 8198.3  | 1813.5 | 7721.2  | 1652.4  | 4302.0 | 3031.3 | 2059.4 | 692.6   |
| 安徽  | 17044.6 | 5414.7  | 1137.4 | 3941.9  | 1041.2  | 2082.1 | 1810.4 | 1224.0 | 392.8   |
| 福建  | 22996.0 | 7572.9  | 1212.1 | 6130.0  | 1223.1  | 2923.3 | 2194.0 | 1234.8 | 505.8   |
| 江西  | 15792.0 | 4809.0  | 1074.1 | 3795.2  | 1047.7  | 1872.1 | 1813.0 | 1000.0 | 381.0   |
| 山东  | 18779.8 | 5030.9  | 1391.8 | 3928.5  | 1394.3  | 2834.3 | 2174.4 | 1627.6 | 398.1   |
| 河南  | 15168.5 | 3959.8  | 1172.8 | 3512.0  | 1054.4  | 1838.0 | 1769.1 | 1541.5 | 321.0   |
| 湖北  | 19537.8 | 5491.3  | 1316.2 | 4310.6  | 1253.2  | 2584.1 | 2187.5 | 1907.9 | 487.0   |
| 湖南  | 18807.9 | 5260.0  | 1215.5 | 3976.1  | 1190.2  | 2322.9 | 2786.2 | 1705.5 | 351.5   |
| 广东  | 26054.0 | 8480.8  | 1135.3 | 6643.3  | 1440.8  | 3423.9 | 2750.9 | 1520.8 | 658.2   |
| 广西  | 14934.8 | 4545.7  | 616.7  | 3268.5  | 898.2   | 2150.1 | 1798.9 | 1364.6 | 291.9   |

数据来源:2019年《中国统计年鉴》中《分地区居民人均消费支出(2018)》指标:

| 变量名称    | 对应自变量 |
|---------|-------|
| 食品烟酒    | X1    |
| 衣着      | X2    |
| 居住      | X3    |
| 生活用品及服务 | X4    |
| 交通通信    | X5    |
| 教育文化娱乐  | X6    |
| 医疗保健    | X7    |
| 其他用品及服务 | X8    |

指标来源于统计年鉴中的相应项目。

## 3 分析方法

对研究的八个主要消费因素,利用统计年鉴中得到的31个地区的相应数据,在SPSS中运用因子分析和聚类分析的方法。首先,用因子分析的方法提取,简化数据,用主要因子对31个地区的消费支出情况得到大致的地区总体消费水平。然后,对简化的数据进行聚类分析,用聚类分析的方法进行区域性消费结构的研究,由此结合经济学相关知识分析我国消费支出的构成情况以及地区性差异,并且得到相关的统计规律和变化特点。

## 4 数据分析

### 4.1 基本数据

|    |         |        |        |        |        |        |        |        |       |
|----|---------|--------|--------|--------|--------|--------|--------|--------|-------|
| 海南 | 17528.4 | 6552.2 | 655.9  | 3744.0 | 826.6  | 1919.0 | 2185.5 | 1236.1 | 409.2 |
| 重庆 | 19248.5 | 6220.8 | 1454.5 | 3498.8 | 1338.9 | 2545.0 | 2087.8 | 1660.0 | 442.8 |
| 四川 | 17663.6 | 5937.9 | 1173.8 | 3368.0 | 1182.2 | 2398.8 | 1599.7 | 1568.6 | 434.5 |
| 贵州 | 13798.1 | 3792.9 | 934.7  | 2760.7 | 878.1  | 2408.0 | 1660.0 | 1083.5 | 280.1 |
| 云南 | 14249.9 | 3983.4 | 789.1  | 3081.1 | 859.9  | 2212.8 | 1772.7 | 1267.7 | 283.2 |
| 西藏 | 11520.2 | 4330.5 | 1285.2 | 2102.6 | 622.3  | 1847.7 | 609.3  | 460.1  | 262.6 |
| 陕西 | 16159.7 | 4292.5 | 1141.1 | 3388.2 | 1200.8 | 2005.8 | 2008.8 | 1749.4 | 373.2 |
| 甘肃 | 14624.0 | 4253.3 | 1111.5 | 3095.0 | 896.9  | 1640.7 | 1710.3 | 1573.9 | 342.4 |
| 青海 | 16557.2 | 4671.6 | 1350.6 | 2990.0 | 932.0  | 2671.4 | 1655.6 | 1842.0 | 444.0 |
| 宁夏 | 16715.1 | 4234.1 | 1388.2 | 3014.3 | 1067.1 | 2724.4 | 2139.5 | 1727.1 | 420.4 |
| 新疆 | 16189.1 | 4691.6 | 1456.0 | 2894.3 | 1082.8 | 2274.4 | 1762.5 | 1592.6 | 434.9 |

## 4.2 因子分析

### 4.2.1 相关矩阵

相关性矩阵

|             |         | 食品烟酒  | 衣着    | 居住    | 生活用品及服务 | 交通通信  | 教育文化娱乐 | 医疗保健  | 其他用品及服务 |
|-------------|---------|-------|-------|-------|---------|-------|--------|-------|---------|
| 相关性         | 食品烟酒    | 1     | 0.58  | 0.834 | 0.814   | 0.85  | 0.82   | 0.582 | 0.865   |
|             | 衣着      | 0.58  | 1     | 0.667 | 0.78    | 0.801 | 0.678  | 0.798 | 0.808   |
|             | 居住      | 0.834 | 0.667 | 1     | 0.902   | 0.869 | 0.9    | 0.738 | 0.912   |
|             | 生活用品及服务 | 0.814 | 0.78  | 0.902 | 1       | 0.91  | 0.877  | 0.803 | 0.889   |
|             | 交通通信    | 0.85  | 0.801 | 0.869 | 0.91    | 1     | 0.861  | 0.776 | 0.914   |
|             | 教育文化娱乐  | 0.82  | 0.678 | 0.9   | 0.877   | 0.861 | 1      | 0.843 | 0.922   |
|             | 医疗保健    | 0.582 | 0.798 | 0.738 | 0.803   | 0.776 | 0.843  | 1     | 0.857   |
|             | 其他用品及服务 | 0.865 | 0.808 | 0.912 | 0.889   | 0.914 | 0.922  | 0.857 | 1       |
| 显著性<br>(单尾) | 食品烟酒    |       | 0     | 0     | 0       | 0     | 0      | 0     | 0       |
|             | 衣着      | 0     |       | 0     | 0       | 0     | 0      | 0     | 0       |
|             | 居住      | 0     | 0     |       | 0       | 0     | 0      | 0     | 0       |
|             | 生活用品及服务 | 0     | 0     | 0     |         | 0     | 0      | 0     | 0       |
|             | 交通通信    | 0     | 0     | 0     | 0       |       | 0      | 0     | 0       |
|             | 教育文化娱乐  | 0     | 0     | 0     | 0       | 0     |        | 0     | 0       |
|             | 医疗保健    | 0     | 0     | 0     | 0       | 0     | 0      |       | 0       |
|             | 其他用品及服务 | 0     | 0     | 0     | 0       | 0     | 0      | 0     |         |

由相关矩阵表可知,8个指标之间高度相关。其中食品烟酒和衣着的相关系数最小,为0.58,教育文化娱乐和其他用品及服务的相关系数,为0.922。各指标之间的相关程度都很高,可以进行因子分析。

#### KMO 和巴特利特检验

|              |      |         |
|--------------|------|---------|
| KMO 取样适切性量数。 |      | 0.816   |
| 巴特利特球形度检验    | 近似卡方 | 358.227 |
|              | 自由度  | 28      |
|              | 显著性  | .000    |

如图所示,对数据进行了KMO和巴特利特检验,KMO值0.819接近1,这些数据适合因子分析。球形度检验的显著性为0,小于显著水平拒绝原假设。变量存在相关关系,可以进行因子分析。

### 4.2.2 公因子方差

公因子方差

|         | 初始 | 提取    |
|---------|----|-------|
| 食品烟酒    | 1  | 0.752 |
| 衣着      | 1  | 0.689 |
| 居住      | 1  | 0.871 |
| 生活用品及服务 | 1  | 0.907 |
| 交通通信    | 1  | 0.908 |
| 教育文化娱乐  | 1  | 0.889 |
| 医疗保健    | 1  | 0.759 |

表1.2.3可知,第一个因子的特征值大于1,由于用特征

|         |   |       |
|---------|---|-------|
| 其他用品及服务 | 1 | 0.956 |
|---------|---|-------|

提取方法:主成分分析法。

由表中数据可知,8个指标对公因子全部大于0.7,所有指标的依赖程度都很高。

### 4.2.3 解释的总方差

总方差解释

| 成分 | 初始特征值 |        | 提取载荷平方和 |       | 方差百分比  |
|----|-------|--------|---------|-------|--------|
|    | 总计    | 方差百分比  | 累积%     | 总计    |        |
| 1  | 6.731 | 84.132 | 84.132  | 6.731 | 84.132 |
| 2  | 0.569 | 7.118  | 91.251  | 0.569 | 7.118  |
| 3  | 0.293 | 3.661  | 94.911  |       |        |
| 4  | 0.154 | 1.928  | 96.839  |       |        |
| 5  | 0.103 | 1.283  | 98.122  |       |        |
| 6  | 0.075 | 0.936  | 99.057  |       |        |
| 7  | 0.058 | 0.724  | 99.781  |       |        |
| 8  | 0.017 | 0.219  | 100     |       |        |

提取方法:主成分分析法。

值法只能得到一个因子,所以固定因子数量,提取前两个因子。第一个因子的贡献度为84.132%,第二个因子的贡献度为7.118%。两因子累计贡献度为91.251%,而其他因子所占贡献度不足10%,所以选取这两个就足够了。

成分矩阵 a

|         | 成分    |        |
|---------|-------|--------|
|         | 1     | 2      |
| 食品烟酒    | 0.867 | -0.417 |
| 衣着      | 0.83  | 0.454  |
| 居住      | 0.933 | -0.201 |
| 生活用品及服务 | 0.952 | -0.022 |
| 交通通信    | 0.953 | -0.037 |
| 教育文化娱乐  | 0.943 | -0.093 |
| 医疗保健    | 0.871 | 0.372  |
| 其他用品及服务 | 0.978 | -0.009 |

提取方法:主成分分析法。

a. 提取了 2 个成分。

由成分矩阵表可得,八个指标对于第一个公因子的相关程度很高,第一个因子很重要。而这些指标和第二个因子的相关性较小,不利于分析,因此进行旋转。

**旋转后的成分矩阵 a**

|         | 成分    |       |
|---------|-------|-------|
|         | 1     | 2     |
| 食品烟酒    | 0.929 | 0.251 |
| 衣着      | 0.332 | 0.886 |
| 居住      | 0.838 | 0.457 |
| 生活用品及服务 | 0.735 | 0.606 |
| 交通通信    | 0.746 | 0.594 |
| 教育文化娱乐  | 0.775 | 0.545 |
| 医疗保健    | 0.417 | 0.851 |
| 其他用品及服务 | 0.747 | 0.632 |

提取方法:主成分分析法。

旋转方法:凯撒正态化最大方差法。

a. 旋转在 3 次迭代后已收敛。

在对表进行了旋转以后,出现了明显的变化。第一个因子主要和食品烟酒,居住,生活用品及服务,交通通信,教育文化娱乐,其他用品和服务有关,这些消费领域和我们的生活息息相关。第二个因子主要和衣着,医疗保健有关,是每个人都必须要接触的消费领域。

**成分得分系数矩阵**

|         | 成分    |       |
|---------|-------|-------|
|         | 1     | 2     |
| 食品烟酒    | .575  | -.470 |
| 衣着      | -.428 | .685  |
| 居住      | .335  | -.176 |
| 生活用品及服务 | .132  | .064  |
| 交通通信    | .150  | .043  |
| 教育文化娱乐  | .213  | -.033 |
| 医疗保健    | -.329 | .579  |
| 其他用品及服务 | .121  | .082  |

提取方法:主成分分析法。

旋转方法:凯撒正态化最大方差法。

表 1.2.7,成份得分系数,由此可算出各地区在两因子下的得分。

$$F1=0.575 * X1 - 0.428 * X2 + 0.335 * X3 + 0.132 * X4 + 0.150 * X5 + 0.213 * X6 - 0.329 * X7 + 0.121 * X8;$$

$$F2 = -0.470 * X1 + 0.685 * X2 - 0.176 * X3 + 0.064 * X4 + 0.043 * X5 - 0.033 * X6 + 0.579 * X7 + 0.082 * X8。$$

各地区的因子得分与排名如下表

| 地区  | F1       | F2       | F     |
|-----|----------|----------|-------|
| 上海  | 3.06702  | 1.14055  | 2.66  |
| 广东  | 1.97037  | -1.36495 | 1.56  |
| 北京  | 1.58182  | 2.34715  | 1.5   |
| 浙江  | 1.21294  | 0.54852  | 1.06  |
| 福建  | 1.29321  | -1.32601 | 0.99  |
| 天津  | 0.84718  | 1.55928  | 0.82  |
| 海南  | 0.96542  | -2.09978 | 0.66  |
| 江苏  | 0.52632  | 0.4207   | 0.47  |
| 广西  | 0.11687  | -1.4733  | -0.01 |
| 安徽  | 0.04586  | -0.85576 | -0.02 |
| 湖南  | -0.03834 | -0.19837 | -0.05 |
| 四川  | -0.0196  | -0.48138 | -0.05 |
| 重庆  | -0.07916 | 0.07192  | -0.06 |
| 江西  | -0.00834 | -1.04263 | -0.08 |
| 湖北  | -0.17921 | 0.20959  | -0.14 |
| 山东  | -0.26221 | 0.22106  | -0.2  |
| 辽宁  | -0.37678 | 1.16355  | -0.23 |
| 云南  | -0.25264 | -1.08835 | -0.29 |
| 贵州  | -0.41035 | -0.91323 | -0.41 |
| 内蒙古 | -0.67564 | 1.08377  | -0.49 |
| 河北  | -0.59298 | 0.02029  | -0.5  |
| 陕西  | -0.6127  | 0.02958  | -0.51 |
| 宁夏  | -0.81657 | 0.53612  | -0.65 |
| 河南  | -0.79725 | -0.08078 | -0.68 |
| 甘肃  | -0.78843 | -0.24347 | -0.68 |
| 新疆  | -0.84646 | 0.41169  | -0.68 |
| 青海  | -0.8689  | 0.46756  | -0.7  |
| 吉林  | -0.93805 | 0.76277  | -0.73 |
| 西藏  | -0.83772 | -1.02441 | -0.78 |
| 山西  | -1.08174 | 0.24384  | -0.89 |
| 黑龙江 | -1.14393 | 0.9545   | -0.89 |

通过因子分析,我国居民消费情况可以按照地区大致分为三类。在这 31 个地区中,按照排名顺序进行排列和分类。第一类是上海,广东,北京,浙江这些经济发达地区,这些地区的综合得分都超过了 1 分。第二类包括了福建,天津,海南,江苏这些沿海经济中心,它们的综合得分为 0 分到 1 分之间,是我国经济的中流砥柱。第三类就是剩下的 23 个地区,这些地区大多位于边疆地区,或是经济欠发达的中西部地区,经济较为落后的同时消费水平也不高,这些地区的得分都在 0 分以下。

按照综合得分排名,从得分越高消费水平越高的角度来看,我国得分一分以上的高消费地区位于沿海和首都。由于消费拉动当地的经济增长和促进就业,产业链形成的作用,这些地区往往是当地的经济增长点,高消费水平会促进附近的地区经济进一步发展,从而创造更多的经济价值,进一步

推高消费水平。

与此同时,得分为0到1分的沿海经济中心承担了从内陆和海外市场交换的中介,无论是依托内地劳动力和资源进行出口型生产,还是依赖海外资源加工进行来料加工,它们都是国际市场和国内市场的连接点,天然会吸引大量的资金和人员,技术流入。这些地区的进口和出口都会产生大量的经济价值,推高当地经济水平的同时,也促使了当地消费水平的进一步上涨,而上涨的消费水平会带动周边地区的消费水平变化,所以它们大多在沿海经济带上,这是由于消费的传导作用决定的。

而大部分得分在1分以下的我国大部分地区,则因为多处于边疆或者内地,商品经济和消费市场尚未充分发展,决定消费水平的多种因素都处于较低的程度,因此,大部分地区的消费水平仍然处于一个较低的位置,这也是由我国的生产加工链条决定的,内地省份主要停供资源和劳动力,深加工产业和高科技产业不多,在吸引人才的方面不及沿海发达省份,同时较低的经济发展水平和经济增速也难以吸引足够的资金,大部分省市的产业可替代性不高,在现有的产业结构下难以发展,这就使得当地的消费水平也徘徊在低水平的位置上。

### 4.3 聚类分析

个案处理摘要<sup>a,b</sup>

| 个案  |       |     |     |     |       |
|-----|-------|-----|-----|-----|-------|
| 有效  |       | 缺失  |     | 总计  |       |
| 个案数 | 百分比   | 个案数 | 百分比 | 个案数 | 百分比   |
| 31  | 100.0 | 0   | .0  | 31  | 100.0 |

a. 平方欧氏距离 使用中

b. 平均联接(组间)

由表可知,没有缺失数据,样本容量为31

#### 4.3.1 聚类表

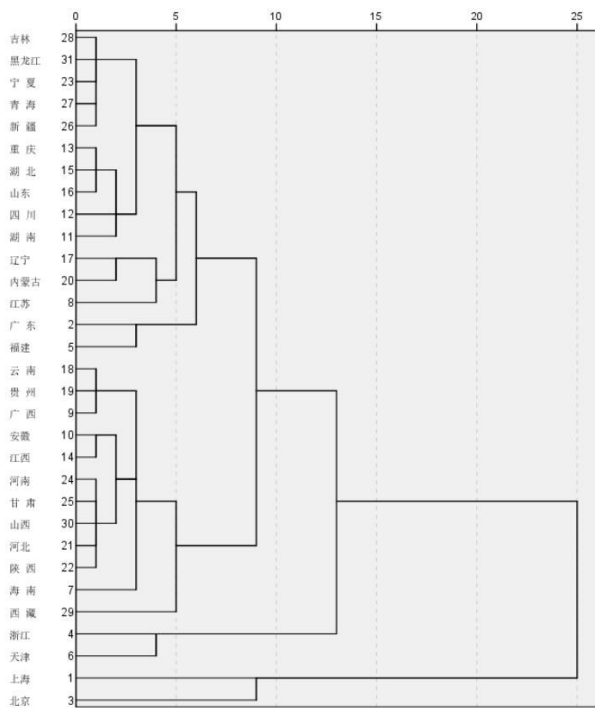
集中计划

| 阶段 | 组合聚类 |      | 系数    | 首次出现聚类的阶段 |      | 下一个阶段 |
|----|------|------|-------|-----------|------|-------|
|    | 聚类 1 | 聚类 2 |       | 聚类 1      | 聚类 2 |       |
| 1  | 28   | 31   | 0.317 | 0         | 0    | 13    |
| 2  | 24   | 25   | 0.319 | 0         | 0    | 5     |
| 3  | 18   | 19   | 0.362 | 0         | 0    | 10    |
| 4  | 10   | 14   | 0.370 | 0         | 0    | 15    |
| 5  | 24   | 30   | 0.408 | 2         | 0    | 12    |
| 6  | 21   | 22   | 0.597 | 0         | 0    | 12    |
| 7  | 23   | 27   | 0.622 | 0         | 0    | 8     |
| 8  | 23   | 26   | 0.648 | 7         | 0    | 13    |
| 9  | 13   | 15   | 0.695 | 0         | 0    | 11    |
| 10 | 9    | 18   | 0.697 | 0         | 3    | 18    |
| 11 | 13   | 16   | 0.715 | 9         | 0    | 14    |
| 12 | 21   | 24   | 0.772 | 6         | 5    | 15    |
| 13 | 23   | 28   | 0.932 | 8         | 1    | 19    |

|    |    |    |        |    |    |    |
|----|----|----|--------|----|----|----|
| 14 | 12 | 13 | 1.085  | 0  | 11 | 16 |
| 15 | 10 | 21 | 1.146  | 4  | 12 | 18 |
| 16 | 11 | 12 | 1.382  | 0  | 14 | 19 |
| 17 | 17 | 20 | 1.462  | 0  | 0  | 22 |
| 18 | 9  | 10 | 1.738  | 10 | 15 | 21 |
| 19 | 11 | 23 | 1.769  | 16 | 13 | 24 |
| 20 | 2  | 5  | 2.184  | 0  | 0  | 26 |
| 21 | 7  | 9  | 2.195  | 0  | 18 | 25 |
| 22 | 8  | 17 | 2.264  | 0  | 17 | 24 |
| 23 | 4  | 6  | 2.688  | 0  | 0  | 29 |
| 24 | 8  | 11 | 2.847  | 22 | 19 | 26 |
| 25 | 7  | 29 | 3.125  | 21 | 0  | 27 |
| 26 | 2  | 8  | 3.942  | 20 | 24 | 27 |
| 27 | 2  | 7  | 5.503  | 26 | 25 | 29 |
| 28 | 1  | 3  | 5.676  | 0  | 0  | 30 |
| 29 | 2  | 4  | 8.196  | 27 | 23 | 30 |
| 30 | 1  | 2  | 16.000 | 28 | 29 | 0  |

使用平均联接(组内)的谱系图

重新标度的距离聚类组合



由树状图,把数据分成5个类别进行分析

| 类别  | 地区  |
|-----|---|
| 第一类 | 上海  |
| 第二类 | 北京  |
| 第三类 | 浙江 天津   |
| 第四类 | 广东 福建 江苏 湖南 四川 重庆 湖北 山东<br>辽宁 内蒙古 宁夏 新疆 青海 吉林 黑龙江 |
| 第五类 | 海南 广西 安徽 江西 云南 贵州<br>河北 陕西 河南 甘肃 西藏 山西            |

通过聚类,将我国31个地区,按不同类型分成了五个大的类别。

我国的经济模式是由南北经济中心向周边辐射带动周边地区的经济发展。根据这些省份距离经济中心的地理位



置,可以由此对消费结构做出解释。

#### 第一类,南方经济中心(上海)

上海是我国的南方的经济中心,也是沿海经济中心。这里消费水平高,主要产业是高科技产业和金融产业,这个地区相对发达的经济状况和经济结构决定了上海的消费结构主要以居住消费和教育文化为主。上海的高房价以及丰富的文化产业使得上海的消费结构较全国其他地方而言更加注重于住房消费以及文化教育消费。上海同时还是我国医疗技术较为发达的省份,上海的医疗服务支出也在一个较高的水平。同时,得益于沿海优势,上海的食品烟酒以及生活服务消费占比维持在一个较低的水平。由于上海独特的消费结构,高比重的住房以及文化消费,极高比重的医疗服务消费,较低比重的食品烟酒消费和生活服务消费,就是上海独特的消费特征。

#### 第二类,北方经济中心(北京)

作为我国的北方的经济中心,同时也是政治中心和京津翼地区的经济中心,北京在具有做为首都的特殊地位的同时,北京的经济结构和消费结构不可避免的受到了影响。由于北京是一个超级大城市,北京的房价非常高,就连周边地区的房价也受到了巨大的影响。北京的医疗技术也是全国顶尖级别,大量的高级别医院就坐落在北京。同时,北京的大量人口决定了北京居民的医疗需求极大,因此,北京的医疗消费占比极高。作为我国北部最大的文化中心,北京的文化产业非常繁荣。大量的文化产业由充足的文化需求支撑,北京的文化消费也是北京居民消费中的重要一环,得益于北京的周边的地理环境和经济政治环境,北京的食品研究消费占比处在较低的位置,在高工资和低物价的共同作用下,食品烟酒消费只是北京居民消费支出的一小部分。同时,由于北京大量的外来务工人员,加上北京周边大劳动力密集涌向北京,北京的生活服务成本也随之降低。北京的生活服务消费占比也相对较低。北京的消费特征,就是极高的住房消费和医疗消费,较高的文化消费以及较低的食品烟酒消费和生活服务消费。

#### 第三类,经济中心周边区域(浙江和天津)

浙江和天津有着非常相似的地理位置和经济地位,它们都是经济中心影响圈中最接近经济中心的省份。浙江距离上海仅有178公里,而天津距离北京也只有133公里。这正好是一辆货运卡车一个白天的运输距离。浙江和天津是类似的,它们都处于上海和北京这两个巨型经济中心的周边。由于距离正好可以满足货运卡车一个白天车程的地理条件,浙江和天津的经济结构和消费结构也就受到了它们最近的经济中心,也就是北京和上海的影响。如果距离过近,在聚集效应的作用下,大量的产业将转移至附近的经济中心而不

是留在本地,而距离过远则使得经济中心的影响无法影响和覆盖到当地。因此,这恰到好处的距离保证了浙江和天津可以作为上海和北京的粗加工产业基地和市场,或是市场的中转站。毕竟这只需要货车跑一天就够了,交易完全来得及。因此,这里的房价较北京和上海低得多,食品烟酒和生活服务也维持在较低的水平。而文化和医疗却受到北京和上海的影响依旧有较高的比重。浙江和天津的消费特征是中等的住房消费,中等的教育文化消费和医疗消费,较低的食品烟酒消费和生活服务消费。

#### 第四类,经济中心的影响范围内

包括广东,福建,江苏,湖南,四川,重庆,湖北,山东,辽宁,内蒙古,宁夏,新疆,青海,吉林,黑龙江在内的一系列省份。

这些省份的地理位置不尽相同,有些位于沿海,而有些位于内陆。它们的经济发展也各不相同。但是,这些省份有一个非常重要的特征,这些省份大多位于内陆经济带或者沿海经济区内。由于我国的经济结构是由南北两个经济中心带动周边逐步向四周发展,这些省份位于经济中心影响圈的中间位置。广东,福建,江苏,湖南,四川,重庆,湖北位于南方的上海经济中心,江浙经济带的影响范围内,辽宁,内蒙古,宁夏,新疆,吉林,黑龙江位于北方的北京经济中心的影响圈内。青海则是由于政策原因,较为特殊的受到了北方经济中心的影响,例如政策优惠和人才选调。这些省份有大量的粗加工产业和一定量的精密工业,这里的资源和劳动力大部分供给于当地的工厂使用。

这些省份由于受到经济中心的影响,交通通讯消费支出较高。因为这里是连接经济中心和更远处的边缘地区的中间带。无论是向经济中心输送劳动力和资源,又或是由经济中心向这些省份输送产品,都要用到交通运输。同时,经济中心的高房价影响在这里已经荡然无存了,但由于当地较之经济中心更低的收入,住房消费占比仍然处于中等水平的位置。同时,这里的教育文化消费和医疗消费处于中等水平,而食品烟酒的消费由于收入较低而处于更高的水平。由于劳动力外流或者是处于当地的工厂或是矿山,这里的生活服务成本比起经济中心周边并没有显著降低,而由于收入下降显得占比较高。这些地区的消费特征是中等水平的住房消费,中等水平的教育文化消费,医疗消费,中等水平的生活服务消费和食品烟酒消费。

#### 第五类,经济中心的影响范围边缘

包括海南,广西,安徽,江西,云南,贵州,河北,陕西,河南,甘肃,西藏,山西的这些省份。

这些省份一般已经远离了南方经济中心或者是北方经济中心,位于经济中心的边缘地带。由于距离经济中心过

远,这些省份主要向经济中心输送资源和劳动力,并且提供一定的市场。这里是国内市场流通网络的末端和劳动力与资源的主要来源地,即使是粗加工产业都较少,可以说这里的经济发展在全国处于落后水平。由于产业较少,这里的劳动力大量外流,使得这里的房价位于全国最低水平,而缺少先进产业导致高收入群体稀缺,这里的教育文化消费占比较低,医疗消费占比也较低。由于廉价的人力资源成本,这里的生活服务消费占比较低。但是,由于这里极低的收入水平,所以对于食品烟酒消费和衣着消费占总消费中的占比就显得更大。这些省份的消费特征是较低的住房消费,极低的教育文化消费,极低的医疗消费,中等的生活服务消费,较高的食品烟酒消费和较高的衣着消费。

## 5 分析结果与建议

从上述分析中可以看出,从消费水平来看,沿海发达省份的消费水平较高,但是,这大量的需求并没有很好的传导到内陆带动内陆的经济发展。也许应该出台一系列政策,引导资金,技术和人才向内陆转移,毕竟消费已经足够了,市场是有的。

从消费结构来看,我国各个省份的消费结构深受经济中心的影响,而几乎于当地的经济水平无关。无论是经济发达的沿海省份,还是经济中等的内陆省份,只要它们与经济中心的经济影响相同,那么它们的消费结构就会非常类似,和当地的收入水平无关,因为收入水平和消费水平会相互抵消,所以经济发展水平不能够决定消费结构。

但是,这样的以两个经济中心影响全国的模式事实上是把全国变成了网状的产业链和市场分发链条,中心是经济中心,向外是经济中心周边,四周是在经济中心影响中的省份,最外围则是边陲省份或是经济不发达省份。这种基于地理位置分配的消费结构不利于经济的健康发展,越是向外的省份就越不可能得到经济发展,这未免不符合正常的市场规律。

## 参考文献

- [1] 何晓群. 多元统计分析(第五版). 中国人民大学出版社
- [2] 高鸿业. 经济学原理(第二版). 中国人民大学出版社
- [3] 国家统计局. 中国统计年鉴 2019