主成分分析理论模型

主成分分析是设法将原来众多具有一定相关性（比如个指标），重新组合成一组新的互相无关的综合指标来代替原来的指标。主成分分析法具体的计算步骤如下所示：

1. 原始指标数据标准化并采集维随机向量， 个样本 ，，，构造样本阵，对样本阵进行如下标准化变换：

其中，得标准化阵。

2、对标准化阵求相关系数矩阵

其中

3、解样本相关矩阵的特征方程得个特征根,确定主成分

　　按确定值，使信息的利用率达85%以上，对每个, , 解方程组得单位特征向量。

4、将标准化后的指标变量转换为主成分

称为第一主成分,称为第二主成分,…,称为第主成分。

5 、对个主成分进行综合评价

对个主成分进行加权求和，即得最终评价值，权数为每个主成分的方差贡献率。

1. 数据来源与数据处理
   1. 数据来源

为了对德州市居民的幸福指数进行综合评价，本文数据完全来源于国泰安CSMAR数据库中经济研究系列的区域经济部分，选取2005至2015年德州市有关衡量幸福感的31个指标，如表3.1所示：

|  |  |  |  |
| --- | --- | --- | --- |
| 符号表示 | 指标 | 符号表示 | 指标 |
|  | 土地面积 |  | 外商实际投资额 |
|  | 医院、卫生院个数 |  | 国内生产总值 |
|  | 医院、卫生院床位数 |  | 第一产业占GDP比重(%) |
|  | 医生数 |  | 第二产业占GDP比重(%) |
|  | 公共图书馆图书总藏量 |  | 第三产业占GDP比重(%) |
|  | 每百人公共图书馆藏书 |  | 地方财政预算内收入 |
|  | 生活垃圾无害化处理率(%) |  | 地方财政预算内支出 |
|  | 年末总人口 |  | 地方财政科学事业费支出 |
|  | 人口自然增长率(‰) |  | 地方财政教育事业费支出 |
|  | 人口密度 |  | 固定资产投资中住宅投资 |
|  | 从业人员 |  | 房地产开发投资额 |
|  | 城镇个体劳动者 |  | 社会消费品零售总额 |
|  | 全部职工年均人数 |  | 公路客运量 |
|  | 职工平均工资 |  | 公路货运量 |
|  | 城乡居民储蓄年末余额 |  | 年末邮电局数量 |
|  | 外商新签协议合同数 |  |  |

表 3.1

本文利用IBM SPSS Statistics 22.0软件对该31个指标进行主成分分析。在使用主成分分析前，先对指标进行相关预处理。

* 1. 数据预处理

利用回归差补法对搜集到的数据的缺失值进行处理。回归插补就是通过建立回归方程的方法，利用该变量与其他变量之间存在的线性关系得出缺失的估计值。本文主要利用影响德州市居民幸福感的各个指标之间可能存在的线性关系运用SPSS 软件进行回归，得出具有显著线性关系的线性方程，最后计算出估计的缺失的数据。将补充好缺失值的数据运用SPSS软件对数据进行标准化处理。

1. 实证分析过程

本文选择31个居民幸福感衡量指标运用SPSS软件进行主成分分析。采取特征值大于1 的方法提取主成分，由总方差解释表可以看出前三个主成分的方差累积率达到86.749%，已经超过经验所要求的85%的水平，因此选择保留三个主成分。

从旋转后的成分矩阵表4.2可以得出，、、、与第二主成分的相关系数较大，故将第二主成分命名为通信和外资发展成分；与第三主成分的相关系数较大，故将第三主成分命名为人口数量成分；其余与第一主成分的相关系数较大，故将第一主成分命名为经济实力成分。

将三个主成分确定后, 还需要根据主成分载荷矩阵计算各个主成分的得分。因此,用主成分旋转后的成分得分系数矩阵对11年德州市居民的幸福感水平进行分析。根据成分得分系数矩阵表4.3，可以进一步计算每年的各主成分的得分，并可以据此对德州市居民幸福感水平进行进一步比较分析。

成分得分系数矩阵 表4.3

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 |  | 1 | 2 | 3 |
| Zscore(VAR00001) | 0.036 | -0.114 | -0.249 | Zscore(VAR00017) | 0.021 | 0.192 | 0.124 |
| Zscore(VAR00002) | 0.042 | -0.038 | -0.028 | Zscore(VAR00018) | 0.043 | -0.062 | 0.157 |
| Zscore(VAR00003) | 0.043 | 0.045 | -0.075 | Zscore(VAR00019) | -0.04 | 0.058 | -0.172 |
| Zscore(VAR00004) | 0.043 | 0.034 | 0.056 | Zscore(VAR00020) | -0.042 | 0.103 | 0.082 |
| Zscore(VAR00005) | 0.041 | 0.065 | 0.088 | Zscore(VAR00021) | 0.043 | -0.091 | 0.011 |
| Zscore(VAR00006) | 0.041 | 0.066 | 0.096 | Zscore(VAR00022) | 0.046 | -0.038 | 0.053 |
| Zscore(VAR00007) | 0.034 | 0.004 | 0.191 | Zscore(VAR00023) | 0.044 | -0.041 | 0.109 |
| Zscore(VAR00008) | 0.042 | 0.072 | -0.15 | Zscore(VAR00024) | 0.043 | -0.044 | 0.062 |
| Zscore(VAR00009) | 0.029 | -0.091 | -0.132 | Zscore(VAR00025) | 0.044 | -0.011 | 0.186 |
| Zscore(VAR00010) | 0.022 | -0.049 | 0.345 | Zscore(VAR00026) | 0.046 | 0.029 | -0.072 |
| Zscore(VAR00011) | 0.044 | -0.026 | -0.076 | Zscore(VAR00027) | 0.046 | 0.03 | -0.075 |
| Zscore(VAR00012) | 0.033 | -0.015 | 0.218 | Zscore(VAR00028) | 0.045 | 0.041 | -0.124 |
| Zscore(VAR00013) | 0.043 | -0.037 | -0.075 | Zscore(VAR00029) | -0.008 | 0.224 | 0.154 |
| Zscore(VAR00014) | 0.045 | 0.038 | -0.119 | Zscore(VAR00030) | 0.027 | 0.198 | -0.035 |
| Zscore(VAR00015) | 0.042 | -0.04 | -0.054 | Zscore(VAR00031) | -0.03 | -0.154 | 0.153 |
| Zscore(VAR00016) | -0.03 | -0.099 | 0.187 |  |  |  |  |

根据成分得分系数矩阵表可以得到以上三个主成分的得分表达式，可以计算出每年对于三个主成分的得分情况。再根据三个主成分的方差贡献率组成表达式：

以求得11年德州市在幸福感方面的综合得分F。因此，可以建立对幸福指数测算的评价模型为：德州市居民幸福指数=(68.046\*经济实力主成分+13.315\*通信和外资发展主成分+5.388\*人口数量主成分)/86.749

最后，对11年德州市居民幸福感进行计算，求得11年幸福感分数平均值为62.99。

3 实证分析结果

由以上主成分分析可以看出经济实力对德州市居民幸福感的影响最大占68.046%，通信和外资发展对其影响次之占13.315%，最后是人口数量对幸福感水平影响占5.388%。

经过对三个因子的综合计算得出德州市居民幸福感水平分数为62.99分。可以看出通过德州市的客观数据计算出的幸福感水平较低，其主要影响因素为经济水平的发展。解决经济水平问题首先要增加居民收入，这主要依靠政府的力量，加大基础设施的建设，从而可以带动城乡就业水平的提高。结合德州市现在的产业基础和土地资源的利用情况，推动现有产业发展的同时，要鼓励创新创业，扩大就业机会和空间。其次要规范经济发展制度，政府要积极发展宏观调控作用，正确引导经济发展，并采取得当的措施，因地制宜，有针对性的规划。

在通信和外资发展方面，首先政府要完善城市交通网络，加大城市规划力度，加快交通建设进度，致力于为人们提供便捷的生活。其次，政府应该大力发展德州市“走出去”战略，要凭借优厚的地理优势加大对外发展力度，有效落实国家贯彻的开放发展战略。