

# 2020 年第五届“数维杯”大学生 数学建模竞赛论文

## 题 目 全面建成小康社会进程研究

### 摘 要

问题一总结了我国在过去 20 年间的全面发展阶段和创新发展阶段的社会经济发展和可持续发展所取得的成功经验，为此建立了皮尔森系数模型，通过对“人民美好生活需要”和“不平衡不充分发展”的主要矛盾进行分析得到了影响全面建成小康社会进程发展的五个指标。提出在我国未来 20 年在经济活力、创新效率、绿色发展、人民生活、幸福和谐 5 个部分等共计 27 项指标来探讨我国可持续发展的思路。

问题二，根据国家统计年鉴，地方统计年鉴等官方数据，运用综合指数模型，构建了从经济发展、文化建设、民主制度、人民生活、资源环境这样五个维度的社会统计监测体系指标。再结合各省市实际发展需要的基础上，通过 GIS 技术方法及数学统计模型对小康社会统计监测指标进行定量化评价，从时空分布层面剖析区域之间的小康社会五个维度的耦合协调关系，给出了我国各省市全面建成小康社会总体统计情况。

问题三，通过对我国 1980~2016 年间的人口总数、出生率、死亡率、自然增长率等数据的分析建立金字塔评价模型分析人口结构、人口变化以及老龄化趋势同时建立 Logistic 数学模型，根据模型可以推导出人口增长的公式表达式为

$$P(t) = \frac{P_m}{1 + \lambda e^{-rt}}。然后根据查找到的 1980~2016 年间的人口增长数据，采用回归$$

分析法可以计算出 Logistic 模型推导公式的具体参数，可以得出公式为

$$P(t) = \frac{14.5833}{1 + 0.522e^{-0.0582t}}。$$

第四问根据前面三问的模型和所确立的全面小康的评价指标给出了供参考的各省的合作计划。

**关键词** 金字塔评价模型；皮尔森系数模型；Logistic 模型；综合指数模型

## 目录

一、问题重述.....	3
二、问题分析.....	3
2.1 问题 1 的分析 .....	4
2.2 问题 2 的分析 .....	4
2.3 问题 3、4 的分析.....	4
三、模型假设.....	4
四、定义与符号说明 .....	5
五、模型的建立与求解.....	5
5.1 问题 1 的模型建立与求解 .....	6
5.1.1 全面建成小康社会过去 20 年成就 .....	6
5.1.2 皮尔森相关系数模型的建立.....	6
5.1.3 模型的求解.....	7
5.1.4 结果 .....	8
5.2 问题 2 的模型建立与求解 .....	9
5.2.1 层次分析法模型的建立.....	9
5.2.2 层次分析模型的求解 .....	10
5.2.3 结果 .....	12
5.3 问题 3 的模型建立与求解 .....	13
5.3.1 金字塔模型和 Logistic 模型的建立 .....	13
5.3.2 Logistic 模型的求解 .....	17
5.4 问题 4 各省的合作计划 .....	20
六、 模型的评价 .....	20
模型的评价本文采用模型误差灵敏度指标来分析: .....	20
参考文献.....	24
附 录 .....	25

## 一、问题重述

2020 年是我党提出的“两个百年奋斗目标”中的第一个百年奋斗目标之中国共产党成立 100 年时（2021 年）全面建成小康社会的决战之年。党和国家通过又一个二十多年的艰苦奋斗，即将由 20 世纪末的我国人民生活水平总体达到小康水平迈向全面建成小康社会的胜利道路。这即将成为中华民族发展史上的又一项新的里程碑。然而在这关键性的一年的初期新型冠状病毒疫情首先在我国武汉开始大规模爆发，之后向全国蔓延传播。虽然目前我们已经很好地控制住了疫情，但对我国的经济发展必然造成了重要的影响。另外，尚未得到有效控制的全球性新型冠状病毒疫情正在更大规模蔓延以及世界处于百年未有的大变局的总体形式使得我国经济内外部环境变得更加恶劣。为此请您结合数学建模方法对如下问题进行合理的优化建模与规划。

**问题 1：**总结归纳过去 20 年间我国全面建成小康社会历程中的成功经验，并结合全球气候变暖、生态环境与人口压力倍增等现状提出未来 20 年的我国可持续发展思路。

**问题 2：**结合国家统计局年鉴、地方统计年鉴等官方数据，总结归纳诸如我国各省全面建成小康社会完成难度系数、完成比率及未来可持续发展能力等在内的多重指标对各省全面建成小康社会总体情况进行合理的评价与排序。

**问题 3：**针对世界处于百年未有的大变局及新型冠状病毒疫情全球性蔓延的大环境，建立以某省为代表的短期及长远的经济社会可持续发展模型。这一模型需要充分考虑诸如人口变化、老龄化趋势、城乡居民收入、资源禀赋及生态环境承载力等关键性因素。

**问题 4：**您能否提供一个各省之间的合作计划。这一计划在充分考虑各省份经济社会发展状况及资源禀赋，同时也能够对合作双方产生积极的促进作用和总体效率的提升。

## 二、问题分析

## 2.1 问题 1 的分析

问题一总结了我国在过去 20 年间的全面发展阶段和创新发展阶段的社会经济发展和可持续发展所取得的成功经验，为此建立了皮尔森系数模型，通过对“人民美好生活需要”和“不平衡不充分发展”的主要矛盾进行分析得到了影响全面建成小康社会进程发展的五个指标。提出在我国未来 20 年在经济活力、创新效率、绿色发展、人民生活、幸福和谐 5 个部分等共计 27 项指标来探讨我国可持续发展的思路。

## 2.2 问题 2 的分析

问题二，根据国家统计年鉴，地方统计年鉴等官方数据，运用综合指数模型，构建了从经济发展、文化建设、民主制度、人民生活、资源环境这样五个维度的社会统计监测体系指标。再结合各省市实际发展需要的基础上，通过 GIS 技术方法及数学统计模型对小康社会统计监测指标进行定量化评价，从时空分布层面剖析区域之间的小康社会五个维度的耦合协调关系，给出了我国各省市全面建成小康社会总体统计情况。

## 2.3 问题 3、4 的分析

问题三，通过对我国 1980~2016 年间的人口总数、出生率、死亡率、自然增长率等数据建立金字塔评价模型分析人口结构、人口变化以及老龄化趋势同时建立 Logistic 数学模型，根据模型可以推导出人口增长的公式表达式为

$P(t) = \frac{P_m}{1 + \lambda e^{-rt}}$ 。然后根据查找到的 1980~2016 年间的人口增长数据，采用回归

分析法可以计算出 Logistic 模型推导公式的具体参数，可以得出公式为

$$P(t) = \frac{14.5833}{1 + 0.522e^{-0.0582t}}。$$

第四问根据前面三问的模型和所确立的全面小康的评价指标给出了供参考的各省的合作计划。

## 三、模型假设

- 1、假设本文数据真实可靠，符合大数定理。
- 2、第三问模型的估计只是针对国家统计局的数据并且数据具有合理性。
- 3、假设近几年我国人口的迁入迁出比率保持基本平衡。
- 4、假设所有表征和影响人口变化的因素都是在整个社会人口平均意义下确定的。

## 四、定义与符号说明

表4-1 定义符号与说明

符号	含义
$A$	判断矩阵
$M_i$	矩阵 $A$ 中各行元素乘积
$w_i$	$M_i$ 的 $n$ 次方根
$W_i$	一级指标权重
$\lambda_{\max}$	矩阵 $A$ 的最大特征根
CI	一致性指标
CR	随机一致性比率
$W_{ij}$	二级指标权重
$W$	组合权重
$\sigma$	标准差
$B_1, B_2, \dots, B_{14}$	二级组合权重
$F_1, F_2, \dots, F_{14}$	根据评价指数表中打出的对应分数
$\bar{X}$	B1 指标调查数据均值
$\sigma$	B1 指标调查数据标准差
$\Sigma$	线上教学效果评价得分

## 五、模型的建立与求解

## 5.1 问题 1 的模型建立与求解

### 5.1.1 全面建成小康社会过去 20 年成就

在第十八届中央委员会第五次全体会议提出决胜全面建成小康社会时期——“十三五”时期的到来，根据这一宏伟的奋斗目标，制定了相应的“十三五”规划。根据国家统计局资料，到 2017 年中国小康指数已经达到 97.06%；我国全面小康实现度排名在前一百的地级市主要分布在 26 个省份中，大多数城市位于江苏、浙江、山东、广东和福建等沿海地。分区域看，我国东部省份远比中西部省份完成度更高，仍旧存在发展不均衡、区域差异较大等问题

### 5.1.2 皮尔森相关系数模型的建立

皮尔逊积矩相关系数（或简称为皮尔森相关系数）是两个变量之间线性关联强度的度量，用  $r$  表示。基本上，皮尔逊积矩相关性试图通过两个变量的数据绘制一条最佳拟合线，而皮尔逊相关系数  $r$  表示所有这些数据点与该最佳拟合线相距多远（即数据点非常适合这种最适合的模型）。相关系数公式用于查找数据之间的关系有多强。公式返回的值介于 -1 和 1 之间，其中：1 表示强烈的正向关系。-1 表示强烈的负相关。结果为零表示完全没有关系。

因此，皮尔逊相关系数  $r$  可以采用从 +1 到 -1 的值范围。值为 0 表示两个变量之间没有关联。大于 0 的值表示正关联；也就是说，随着一个变量的值增加，另一个变量的值也增加。小于 0 的值表示负关联；即，随着一个变量的值增加，另一变量的值减小。

两个变量的关联性越强，则分别取决于关系是正还是负，皮尔逊相关系数  $r$  越接近 +1 或 -1。达到 +1 或 -1 的值意味着您的所有数据点都包含在最佳拟合线中 - 没有数据点显示出与该线相距任何偏差。 $r$  的值介于 +1 和 -1 之间（例如， $r = 0.8$  或  $-0.4$ ），表明最佳拟合线周围存在差异。 $r$  的值越接近 0，则最佳拟合线附近的变化越大。下图显示了不同的关系及其相关系数：

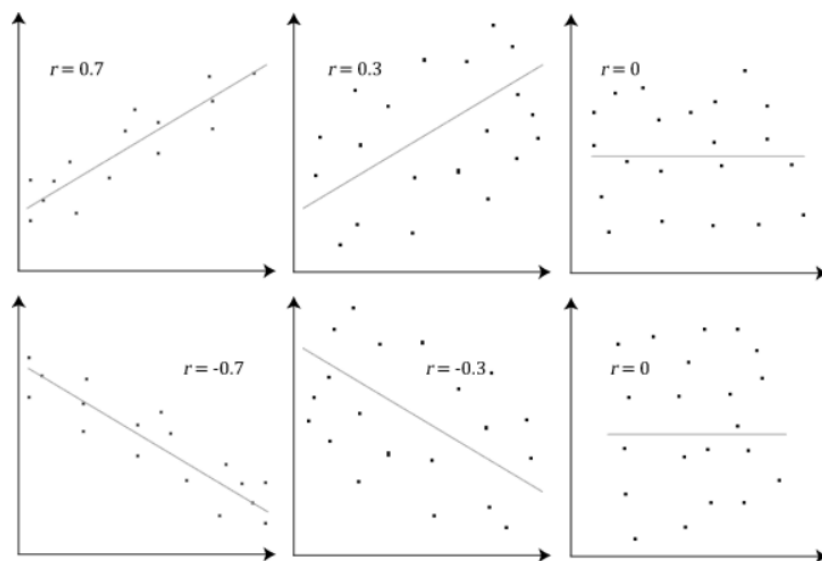


图 5-1 皮尔森相关系数

对于 Pearson  $r$  相关性，两个变量都应呈正态分布。即正态分布描述了变量值如何分布。有时称为“钟形曲线”或“高斯曲线”。一种简单的方法是使用 Shapiro-Wilk 检验分别确定每个变量的正态性。同时，每个变量应该是连续的，即间隔或比率，例如体重，时间，身高，年龄等。如果一个或两个变量在测量中是序数，则可以进行 Spearman 相关。

模型可以表示为：

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}.$$

其中， $x_i, y_i$  为用以索引的各个样本点， $n$  是样本量。

### 5.1.3 模型的求解

表5-1 我国可持续发展的评价指标

评价视角	评价指标	指标解释
经济活力	GDP 增长速度(+)	报告期 GDP/基期 GDP(可比价)
	GDP 占总产出比重(+)	GDP/总产出
	新经济新动能产业增加值占比(+)	新经济新动能产业增加值/GDP
	现代服务业增加值占比(+)	现代服务业增加值/GDP
	居民消费贡献率(+)	居民最终消费/GDP

	综合负债率指数 (—)	根据银行不良贷款率、地方政府负债率和企业资产负债率综合测算
	外贸依存度 (o)	进出口总额/GDP
创新效率	高技术产业增加值占比 (+)	高技术产业增加值/GDP
	全要素生产率 (+)	根据生产函数计算
	亩均 GDP (+)	GDP/占地面积
	万元 GDP 能耗 (—)	综合能耗/GDP
绿色发展	优质可耕地占比 (+)	优质可耕地面积/区域总面积
	空气质量指数 (+)	空气质量优良天数比重或监测的空气质量平均值
	污水处理率 (+)	处理污水数量/总污水数量
	生活垃圾无害化处理率 (+)	无害化处理生活垃圾数量/生活垃圾总量
	地表水质量 (+)	实际监测数据
人民生活	人均可支配收入与人均 GDP 之比 (+)	人均可支配收入/人均 GDP
	社会保障指数 (+)	五险综合保障水平得分
	社会不安定指数 (—)	失业率 +CPI
	高等教育毛入学率 (+)	高等教育在学人数/18—22 岁人口数
	人均文教医卫财政支出强度 (+)	人均文教医卫财政支出/人均 GDP
	居民家庭文体旅游消费支出比重 (+)	居民家庭文体旅游消费支出/总消费支出
	产品质量指数 (+)	各种各类产品抽检合格率的加权平均值
社会和谐	基尼系数 (o)	按照基尼系数公式测算
	城乡居民可支配收入比 (o)	城镇居民人均可支配收入/乡村居民人均可支配收入
	区域人均可支配收入极值比 (o)	最高收入区域人均可支配收入/最低收入区域人均可支配收入
	社会满意度指数 (+)	公民调查数据

### 5.1.4 结果

通过对经济活力、创新效率、绿色发展、人民生活、幸福和谐指标的分析可以看出我国的可持续发展的实现路径有以下两种：

#### (1) 从实现途径上考虑可持续发展

转变经济发展方式。在发展平衡性、协调性、可持续性明显增强的基础上实现倍增。理论创新与实践创新相结合，通过创新变为发展新动力，用科技对经济增长作为评价指标，上涨幅度较大为理想目标。通过继续实施区域总体发展战略，充分发挥各地区比较优势，区域协调发展机制基本形成。通过培育开放型经济发展新优势，使对外开放水平进一步提高，国际竞争力明显增强。

#### (2) 从未来资源考虑可持续发展



在天空蓝计划下我们不断要探索出一条符合我国国情的经济发展模式，以民生工程为前提条件，遵循着“以人为本”的治国理念，关心民生，关注全民社保都是至关重要的。不仅要加快社会转型，转变经济发展模式，扩大内需，更需要提高资金利用率，完善收入分配制度，建立全民社保体系。

## 5.2 问题 2 的模型建立与求解

### 5.2.1 层次分析法模型的建立

#### （1）小康社会的定义

“小康社会”是一个具有中国特色的社会主义概念，是指温饱有余、但还不富裕的社会状态。自 1979 年邓小平的初次描绘，到近几年党的十七大、十八大、十九大等相关会议和资料陆续对其论述补充。本文定义的全面小康社会，正是指经济得到较大发展，生产力比较发达，生活环境明显改善，广大人民群众能够在教育、医疗、社会保障等方面享有更好的服务，居民生活幸福感比较高的一种社会状态。

我们发现经过多年的诠释、完善，党和政府陆续补充了对全面小康的定义，每个时期都有每个时期的建设总目标。

#### （2）评价指标的确定

- 1) 人均 GDP 体现当年经济真实演变过程，将所选取的 2011-2015 年份指标数据进行换算。
- 2) 社会安全指数是一个综合考量复合指标，主要用于衡量区域安全变化程度。
- 3) 环境因素，单 GDP 能耗、水耗及建设用地占用面积计算方法类似，都是指在一定时期内（通常为一年）由于环境统计数据的限制，目前因为数据获取的局限性仅由 PM2.5 达标天数比例。
- 4) 人民幸福指数

#### （3）评价指标权重设定方法选取

目前，常用的权重赋值方法包括专家评分法（Experts Grading Method）、德尔菲法（Delphi Method）和层次分析法（Analytic Hierarchy Process, 简称 AHP）

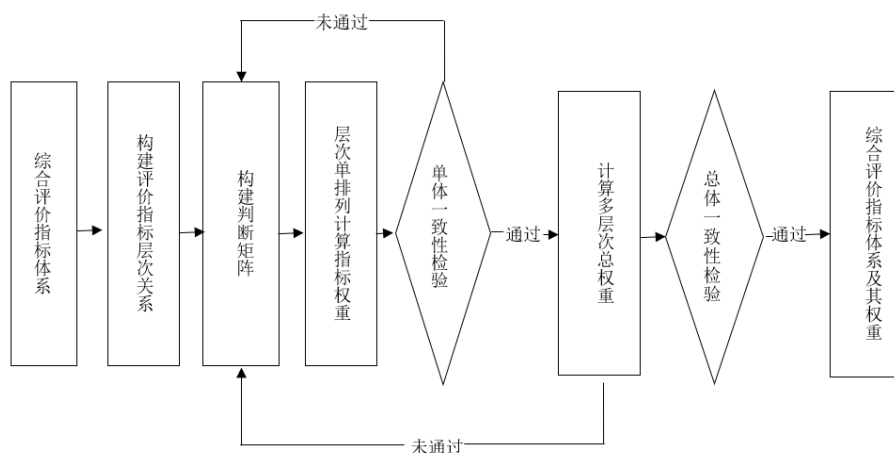


图 5-2 层次分析法流程图

### 5.2.2 层次分析模型的求解

层次分析法（AHP）最早由美国运筹学家匹茨堡大学教授 *T.L.Saaty* 提出，该算法是一种定性与定量相结合的决策分析方法，它可以帮助决策者把一些复杂问题条理化、层次化，对于一些定性的问题从心理学角度去定量化。应用这种方法，可以实现各因素之间的比较和计算，最后可以得出不同方案的权重，为最佳方案的选择提供依据。

#### （1）建立判断矩阵

对在同一层次的两个指标进行对比时，根据以下评分标度表，给出重要性判断值，生成相应的判断矩阵  $A=(a_{ij})_{n \times n}$ 。

表 5-2 判断矩阵评分标度表

标度	定义
1	A 级指标中， $A_i$ 与 $A_j$ 一样重要
3	A 级指标中， $A_i$ 与 $A_j$ 略微重要
5	A 级指标中， $A_i$ 与 $A_j$ 重要
7	A 级指标中， $A_i$ 与 $A_j$ 明显重要
9	A 级指标中， $A_i$ 与 $A_j$ 极为重要
2、4、6、8	A 级指标中，重要性介于相邻分标度之间

1~9 的倒数

A 级指标中,  $A_i$  与  $A_j$  之间的重要  
度衡量

表 5-3 一级指标比较结果

全面小康指标	人均 GDP	社会安全	平均受教	能源消耗	环境质量	污染排放
人均 GDP	1	1/3	2	1/3	2	1/3
社会安全	3	1	3	1	3	1
平均受教	1/2	1/3	1	1/3	1	1/2
能源消耗	3	1	3	1	1/2	1
环境质量	1/2	1/3	1	2	1	1/2
污染排放	3	1	2	1	2	1

表 5-3 的判断矩阵可 W 简化为矩阵 A:

$$A = \begin{bmatrix} 1 & 1/3 & 2 & 1/3 & 2 & 1/3 \\ 3 & 1 & 3 & 1 & 3 & 1 \\ 1/2 & 1/3 & 1 & 1/3 & 1 & 1/2 \\ 3 & 1 & 3 & 1 & 1/2 & 1 \\ 1/2 & 1/3 & 1 & 2 & 1 & 1/2 \\ 3 & 1 & 2 & 1 & 2 & 1 \end{bmatrix}$$

## (2) 进行层次单排序

利用 matlab, 计算矩阵 A 中各行元素乘积  $M_i$ :

$$M_i = \prod_{j=1}^4 b_j \quad (5-1)$$

$$\text{得 } M_1 = \frac{2}{9}; M_2 = 9; M_3 = \frac{1}{12}; M_4 = 6$$

计算  $M_i$  的 n 次方根  $w_i$ :

$$w_i = \sqrt[n]{M_i} \quad (5-2)$$

$$\text{计算得 } w_1 = \sqrt[4]{\frac{2}{9}} = 0.687; w_2 = \sqrt[4]{9} = 1.732; w_3 = \sqrt[4]{\frac{1}{12}} = 0.537; w_4 = \sqrt[4]{6} = 1.565$$

令向量  $w = [0.687 \quad 1.732 \quad 0.537 \quad 1.565]^T$ , 对之进行归一化处理, 计算一级指标权重  $W_i$

$$W_i = \frac{w_i}{\sum_{i=1}^4 w_i} \quad (5-3)$$

$$\sum_{i=1}^4 w_i = 0.687 + 1.732 + 0.537 + 1.565 = 4.521$$

$$W_1 = \frac{0.687}{4.521} = 0.152; W_2 = \frac{1.732}{4.521} = 0.383; W_3 = \frac{0.537}{4.521} = 0.119; W_4 = \frac{1.565}{4.521} = 0.346$$

即一级指标权重  $W = [0.152 \quad 0.383 \quad 0.119 \quad 0.346]^T$

### (3) 层次单排序一致性检验

计算判断矩阵 A 的最大特征根  $\lambda_{\max}$  :

$$\lambda_{\max} = \sum_{i=1}^n \frac{(AW)_i}{nW_i} \quad (5-4)$$

$$\lambda_{\max} = \frac{0.633}{4 \times 0.152} + \frac{1.542}{4 \times 0.383} + \frac{0.496}{4 \times 0.119} + \frac{1.423}{4 \times 0.346} = 4.118$$

$$\text{一致性指标 CI} = \frac{\lambda_{\max} - n}{n - 1} = \frac{4.118 - 4}{4 - 1} = 0.039$$

查 RI 取值表可知, 当  $n=4$  时, RI 取值为 0.90。

表 5-4 RI 取值表

阶数	1	2	3	4	5	6	7	8
RI 取值	0	0	0.5	0.9	1.1	1.2	1.3	1.4
			8	0	2	4	2	1

从而可得随机一致性比率

$$CR = \frac{CI}{RI} = \frac{0.039}{0.90} = 0.043 < 0.1 \quad (5-5)$$

结果显示, 单层次一致性  $CR < 0.1$ , 全部合格。

## 5.2.3 结果

以河南为例, 通过上述计算分析可得, 从时间上看, “十二五”期间河南省

在经济发展、民主法制、文化建设、人民生活、资源环境等方面均有不同程度的增长。民主法制维度实现程度最低且增速缓慢,根据上一节对各个指标项的监测,我们可以发现是由于每万人拥有律师数(仅完成 42.74%)的指标项发展水平太过缓慢而导致的。

就各个维度发展速度而言,文化建设维度无疑是五个维度中增速最多的,从 2011 年的 72.48%到 2015 年的 89.05%,平均每年提升 5.28 个百分点,若保持这个速度发展将有很大的希望在 2020 年完成建设任务;而民主法制维度本身就发展水平落后,建设力度又远低于其他维度,在一定程度上来说,是河南省决胜全面小康的难点重点所在。总的来说,河南省全面建设小康社会五个维度的发展速度还是达到了年增速 3%左右,保持着一种稳中增进的态势发展。

## 5.3 问题 3 的模型建立与求解

### 5.3.1 金字塔模型和 Logistic 模型的建立

金字塔模型是表示人口年龄、性别结构的一种特殊的条形图,其画法是将各年龄男性与女性人数或百分比分别在纵轴左右画成并列的横的条形。按年龄增长顺序自下而上排列,人口金字塔数据指标能形象地直观地反映人口年龄、性别结构,便于评价人口现状、类型和未来发展趋势。

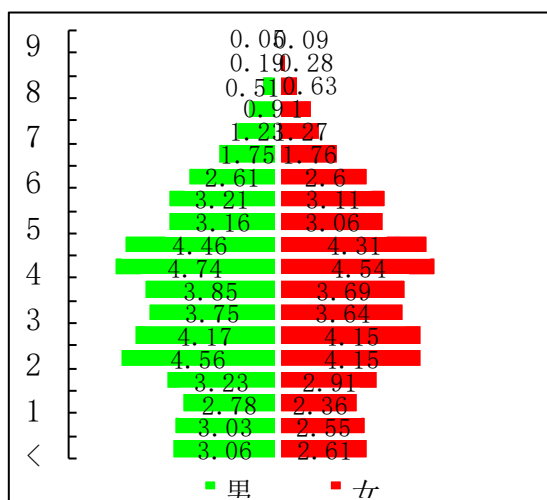


图 5-3 2013 年人口金字塔图

从图 5-3 中可以看出 2013 年的人口金字塔呈现收缩形,这种类型表明少年儿童人口比重缩小,老年人口比重增大,是出生率长期下降的结果。这种类型的

人口由于育龄人群比重低，后备力量更低，如果生育水平不变，它未来的人口再生产趋势呈负增长，人口缩减。

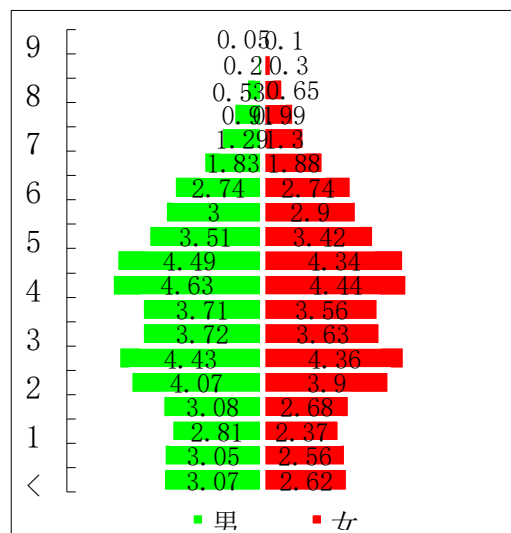


图 5-4 2014 年人口金字塔图

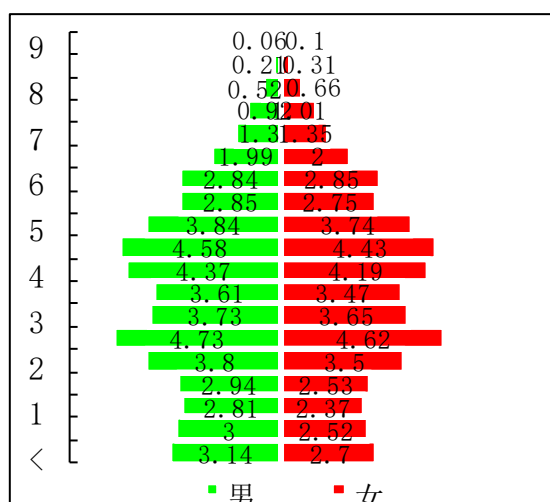


图 5-5 2015 年人口金字塔图

结合 2013、2014、2015 年三年的人口金字塔图可以看出使收缩型人口结构向静止型人口结构转变，但是从图中发现转变的效果并不是很显著，有可能是因为中国的人口基数太大，不是靠短时间能够改变，政策的实施有利于减缓人口的老龄化，但是并不能从根本解决中国人口结构的现状。

而根据中国统计年鉴可以查到每年人口的出生率、死亡率、自然增长率，根据查到的资料可以得出：出生率=死亡率+自然增长率。在一定程度上可以得出出生率和死亡率是影响人口数量的直接因素，下图 5-6 中横坐标代表了时间，纵坐

标代表了比率，图中蓝线代表了出生率，红线代表了死亡率，绿线代表了自然增长率。

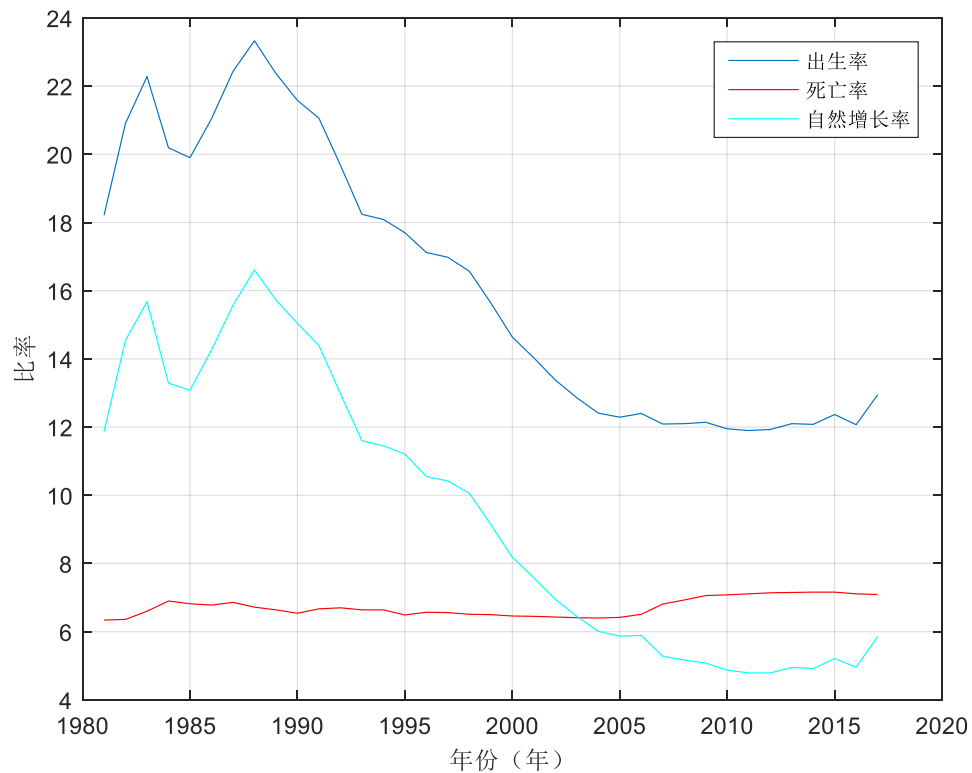


图 5-6 人口指标变化图

从图 5-6 中的蓝线可以看出，出生率在 1987 年达到了峰值，而随后的一直呈现下降趋势，而在 2011 年后又呈现出了上升的趋势，但是这种上升趋势并不是太明显；图中红线代表了死亡率，从图 5-6 中可以看出整体是比较趋于平缓，在 2006 年后略微有上升的趋势，死亡率上升的原因可能与老龄化人口的数量增加有一定的关系；图 5-6 中的自然增长率=出生率-死亡率。

进一步我们发现，关于人口结构变化，通过整理中国统计年鉴上找到的数据，从 0 岁开始，每 5 岁作为一个年龄段，可以采用每个年龄段人数占总数的比例来确定人口结构的变化。其中男性用绿色表示，女性用红色表示。纵轴用来表示年龄段。我们可以以 2013 年、2014 年、2015 年和 2016 年四年的人口金字塔图作为评价人口结构对病床的影响，而影响这个系统人口的因素很多，直接的作用就是出生率和死亡率，而其他的比如环境、经济等因素都是通过对直接因素的影响来达到影响人口数量，框图如图 5-7 所示。根据历年来中国人口数量发展特点，发现其增长趋势与 Logistic 曲线相似，故进一步可利用 Logistic 曲线来模拟中国

人口长期发展的趋势。

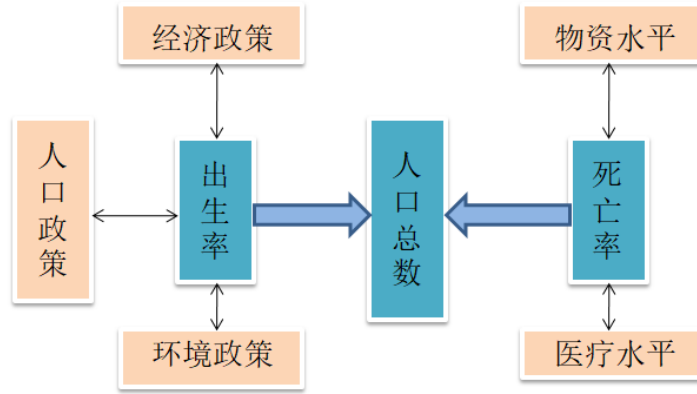


图 5-7 影响人口因素框架图

模型建立：

假定  $t_0 = 0$  年的人口为  $P(0)$ ，人口增长率为  $r$ ，则第  $t$  年的人口为

$$P(t) = P(0)(1+r)^t$$

化成指数形式

$$P(t) = P_0 e^{rt}$$

求导得到齐次方程

$$\frac{dP(t)}{dt} = rP(t)$$

考虑到自然资源、环境条件等因素对人口增长起阻滞作用，并且随着人口增加阻滞作用越大，即人口增长率  $r$  为人口总量  $P(t)$  的减函数。因此对上式做出修改，加上一个表征环境约束因子的二次项  $qP(t)^2$ ，从而得到二阶 Bernoulli 式齐次方程

$$\frac{dP(t)}{dt} = rP(t) - qP(t)^2 = rP(t)\left[1 - \frac{P(t)}{P_m}\right]$$

这里  $q$  为约束参数， $P_m = r/q$  表示区域饱和人口即最大人口容量。该方程的初始条件和饱和条件分别为  $P(t_0) = P_0, P(t) \leq P_m$ 。解之得到 Logistic 预测模型

$$P(t) = \frac{P_m}{1 + \left(\frac{P_m}{P_0} - 1\right)e^{-rt}} = \frac{P_m}{1 + \lambda e^{-rt}}$$



### 5.3.2 Logistic 模型的求解

先通过国家统计年鉴查找到 1980 年到 2016 年的人口数据（见表 5-5），根据这些数据，再结合 Logistic 模型计算出的公式形式，可以采用回归分析计算出公式中的未知变量。

表 5-5 历年人口数量统计表

年份	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
人口	98.71	10.01	10.27	10.30	10.44	10.59	10.75	10.93	11.10	11.27
年份	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
人口	11.43	11.58	11.72	11.85	11.99	12.11	12.24	12.36	12.48	12.58
年份	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
人口	12.67	12.76	12.85	12.92	13.00	13.08	13.14	13.21	13.28	13.35
年份	2010	2011	2012	2013	2014	2015	2016			
人口	13.41	13.47	13.54	13.61	13.67	13.75	13.83			

#### 5.3.2.1 回归分析模型的简介

根据 Logistic 模型建立人口增长的数学模型，根据模型可以推导出人口增长的公式表达式为  $P(t) = \frac{P_m}{1 + \lambda e^{-\alpha t}}$ 。然后根据查找到的 1980~2016 年间的人口增长数据，采用回归分析法可以计算出 Logistic 模型推导公式的具体参数。

#### 5.3.2.2 回归分析模型建立

考虑  $p$  个变量  $y_1, y_2, y_3, \dots, y_p$  与  $m$  个自变量  $x_1, x_2, x_3, \dots, x_m$  的建模问题。偏量小二乘回归的基本作法是首先在自变量集中提出第一部分  $t_1$  ( $t_1$  是  $x_1, x_2, x_3, \dots, x_m$  的线性组合，且尽可能多地提取原自变量集中的变异信息)；同时在因变量集中也提取第一成分  $u_1$ ，并要求  $t_1$  和  $u_1$  相关程度最大。然后建立因变量  $y_1, y_2, y_3, \dots, y_p$  与  $t_1$  的回归，如果回归方程已经达到满意的精度，则算法终止。否则继续第二对成分的提取，直到能达到满意的精度为止。

若最终对自变量集提取  $r$  个成分  $t_1, t_2, t_3, \dots, t_r$ ，偏最小二乘回归将通过建立  $y_1, y_2, y_3, \dots, y_p$  与  $t_1, t_2, t_3, \dots, t_r$  的回归式，然后再表示为  $y_1, y_2, y_3, \dots, y_p$  与原始变量的回归方程式，即偏最小乘回归方程式。

为了方便起见，不妨假定  $p$  个因变量  $y_1, y_2, y_3, \dots, y_p$  与  $m$  个自变量  $x_1, x_2, x_3, \dots, x_m$  均为标准化变量。因变量组和自变量组的  $n$  次标准化观测数据阵分别记为：

$$F_0 = \begin{pmatrix} y_{11} & \dots & y_{1p} \\ \vdots & & \vdots \\ y_{n1} & \dots & y_{np} \end{pmatrix}, \quad E_0 = \begin{pmatrix} x_{11} & \dots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nm} \end{pmatrix}$$

根据查找的数据，需要采用 1980~2016 年间的的数据，最后可以得出表达式

$P(t) = \frac{P_m}{1 + \lambda e^{-rt}}$  来做拟合得出的系数，代入表达式最后可以得出公式为

$$P(t) = \frac{14.5833}{1 + 0.522e^{-0.0582t}}。$$

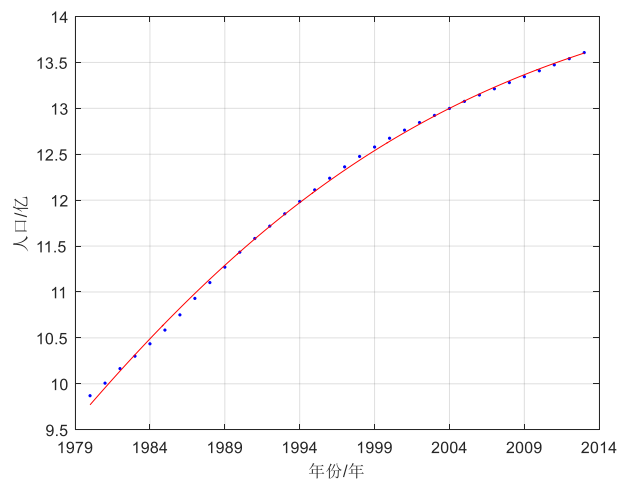


图 5-8 拟合曲线

图 5-8 中以时间作为横坐标，而以人口作为纵坐标，先画出散点图，用蓝色点表示，可以看出拟合图形效果较好。

为了对需要有养老需求的人口数量进行预测，采用附件给定的人口数量与结构进行分析，对每年的老年人人口数量随时间变化进行统计，并整理数据代入线性回归模型中。

经过查阅相关资料，查阅了附件资料中的数据以及数据文献中老年人的实际

数量，将模型代入方程进行求解：

选择 65 岁以上的人口作为急切需要进入养老状态的人群。统计数据如下所示，进行可视化分析：

以河南省为例：

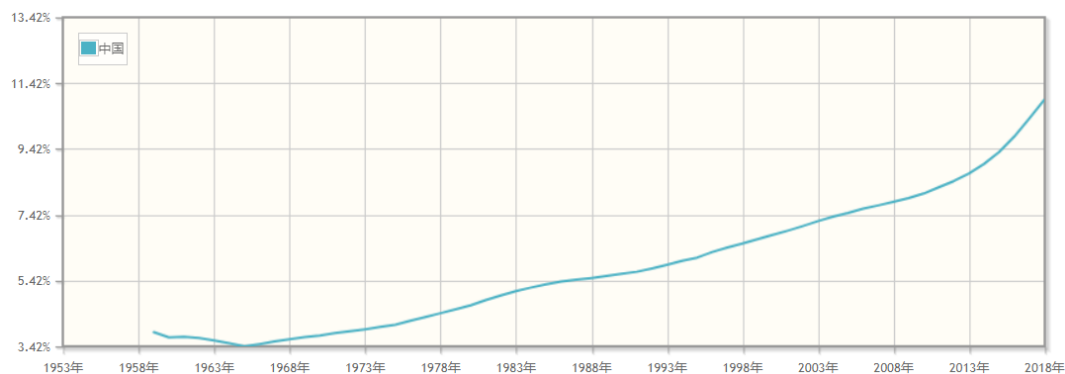
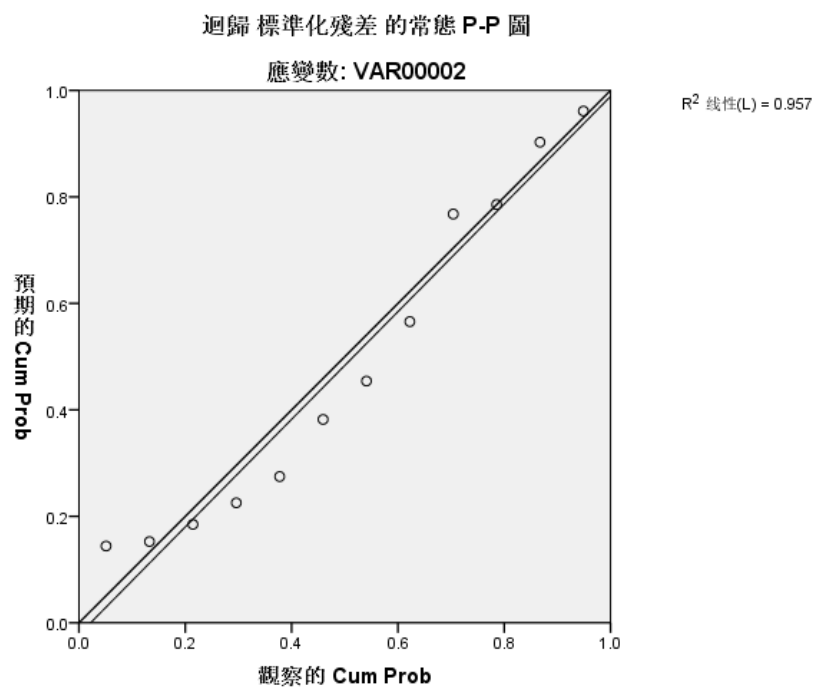


图 5-9 河南省老年人数据随时间变化情况

利用 SPSS 进行编程并求解，在标准化后的误差拟合情况如下所示：



模型的总 R 方为：

模型摘要<sup>b</sup>

模型	R	R 平方	調整後 R 平方	標準偏斜度錯誤	變更統計資料				
					R 平方變更	F 值變更	df1	df2	顯著性 F 值變更
1	.962 <sup>a</sup>	.926	.918	.298750	.926	124.830	1	10	.000

a. 預測值：（常數），VAR00001

b. 應變數: VAR00002

其回归系数为：

模型	非標準化係數		標準化係數	T	顯著性
	B	標準錯誤	Beta		
1 (常數)	-552.883	50.278		-10.997	.000
VAR00001	.279	.025	.962	11.173	.000

a. 應變數: VAR00002

## 5.4 问题 4 各省的合作计划

第四问根据前面三问的模型和所确立的全面小康的评价指标给出了供参考的各省的合作计划：

以一带一路共建小康社会为蓝图将新疆和福建作为两个核心地区，东西部协同发展同时在 18 个重要省份（新疆、陕西、宁夏、甘肃、青海、内蒙、黑龙江、吉林、辽宁、广西、云南、西藏、上海、福建、浙江、广东、海南、重庆）的联合下做好合作发展，将西宁、郑州、武汉、成都、长沙、南昌以及合肥作为 7 大合作高地，为可持续发展贡献力量同时加大 15 个港口（上海、天津、宁波-舟山、广州、深圳、湛江、汕头、青岛、烟台、大连、福州、厦门、泉州、海口、三亚）的布局建设，新兴 2 个枢纽机场（上海、广州）对外合作密切为我国可持续发展助力，把持坚持开放合作、坚持和谐包容、坚持市场运作、坚持互利共赢的共建原则并且要把握政策沟通、设施联通、资金融通、贸易畅通、民心相通的合作重点，各省份之间点对点联合发展。

## 六、模型的评价

模型的评价本文采用模型误差灵敏度指标来分析：

在 EXCEL 中用 VARPA 等函数计算上述误差得到：S1=1.7199e+04；  
S2=1.7270e+04

计算 C 值为 0.9959

得到参数的估计值及统计量如下：

表 6-1

估计值	T 值	P 值
$\theta = 0.78989$	0.988914366	0.333457493

均方误差 240614.847，绝对误差 443.7093059，复相关系数 0.999593415

从计算结果来看， $\theta$ 都通过了显著性检验。数值拟合的误差比较小。说明拟合效果不错，从拟合误差得到的自相关系数图 6-1 及偏相关系数图 6-2 自相关拖尾偏向关 18 阶截尾亦说明该模型是适合的。

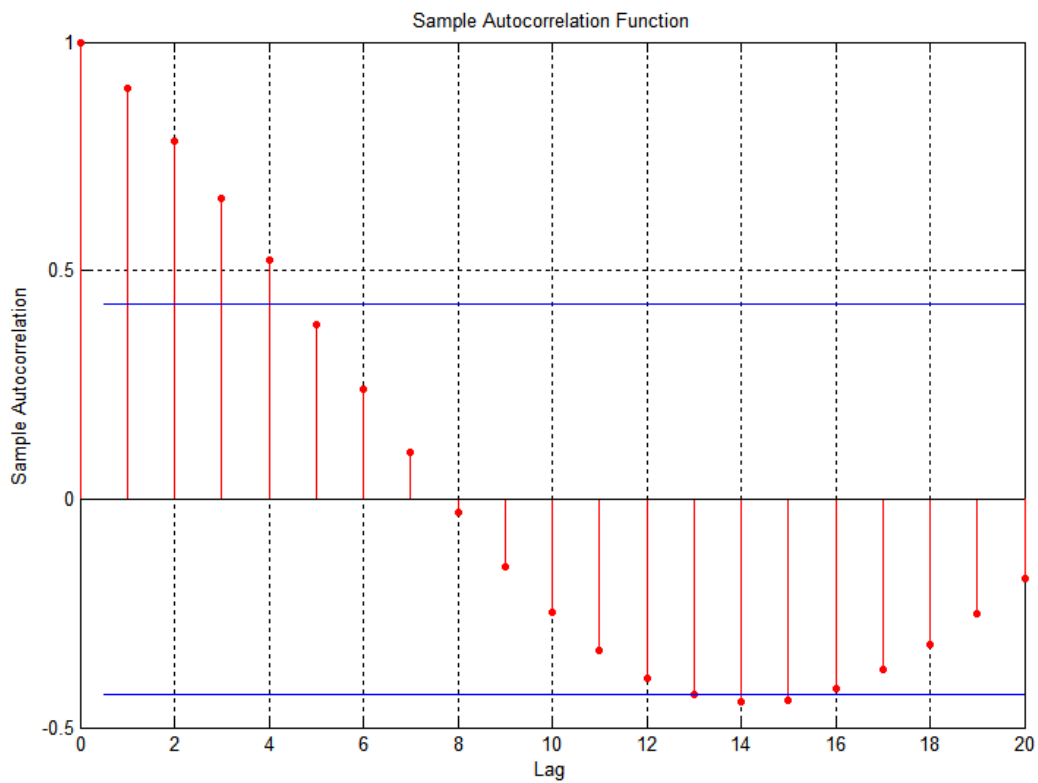


图 6-1 自相关系数图

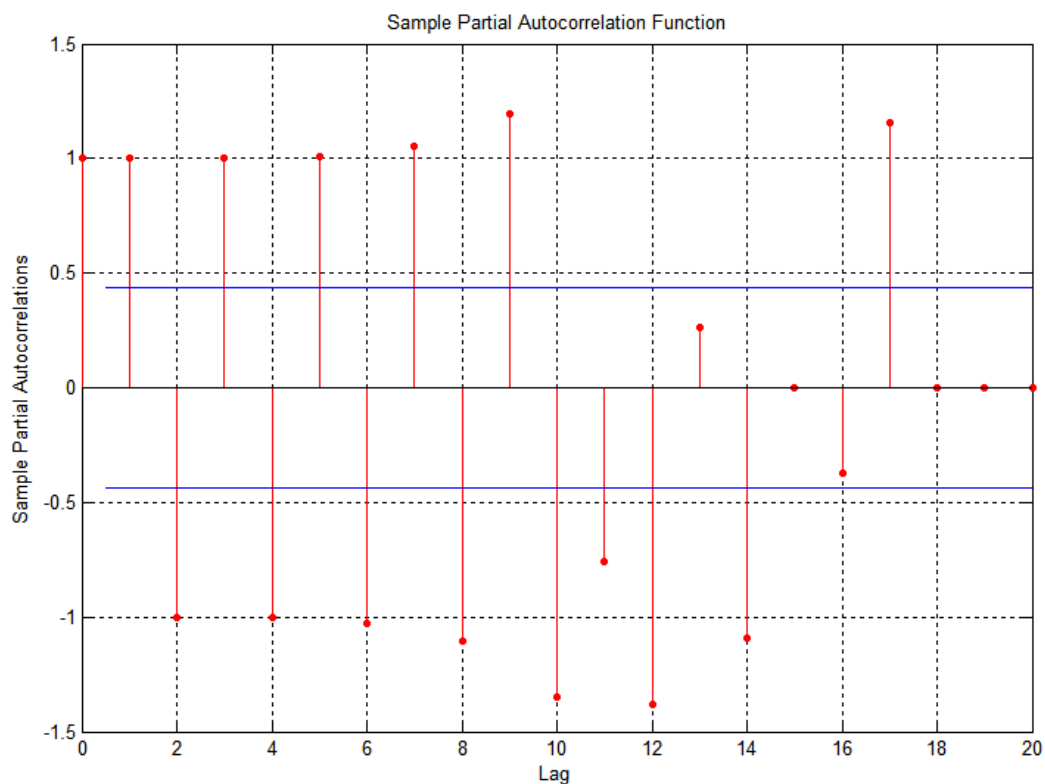


图 6-2 偏相关系数图

又在 MATLAB 工具箱中绘制了互相关图像与绝对误差热力图如下图 6-3 和图 6-4 所示不难看出拟合精度效果良好。

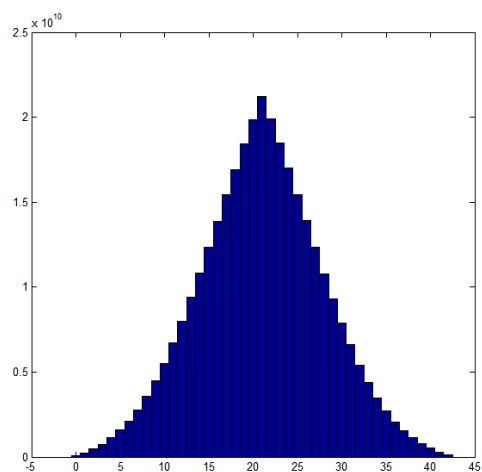


图 6-3 互相关图

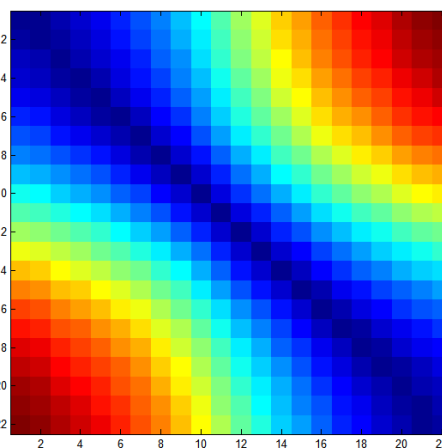


图 6-4 绝对误差热力图

其中计算 F 值为 2.5459, 置信区间为  $[-584.7121, 545.3337]$  和  $[0.9904, 1.0171]$  相关测试数据与实际数据的残差如下图 6-5 所示:

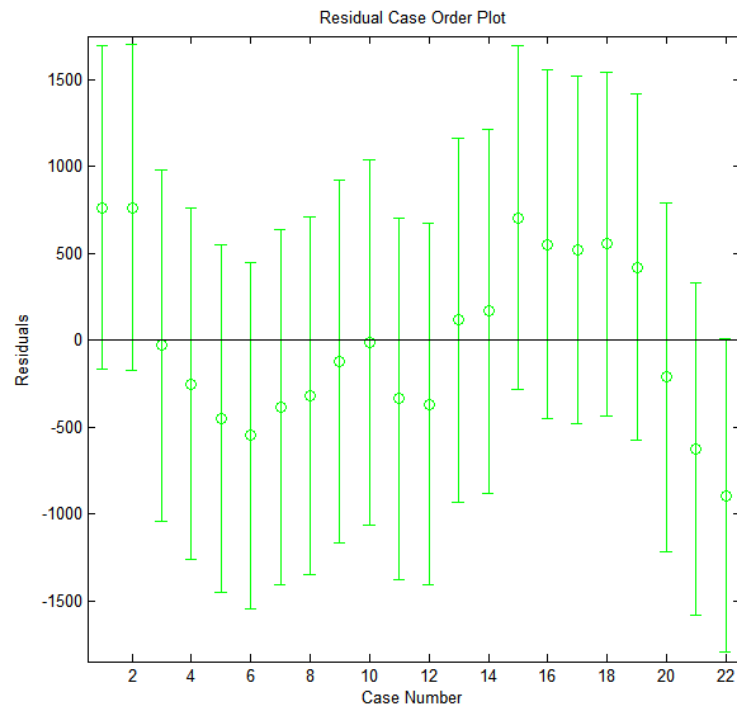


图 6-5 残差分析图

可以看出数据的残差在零点较近的位置，置信区间内均包含零点，误差在可接受范围。

## 参考文献

- [1]朱乐尧. 区域经济发展效果的宏观评价[J]. 数量经济技术经济研究, 1989( 9) : 75—78.
- [2]国务院发展研究中心管理世界杂志社等. 1991 年 188 个地级以上城市经济社会发展水平评价 [J] . 管理世界, 1992( 6) :143—149.
- [3]朱庆芳. 社会指标的应用及评价比较实例———改革开放以来哪些地区经济社会发展速度快、水平高 [J] . 社会学研究, 1993( 2) : 31—35.
- [4]郁磊, 史峰, 王辉, 胡斐. MATLAB 智能算法 30 个案例分析[M]. 北京: 北京航空航天大学出版社, 2015: 27.
- [5]余胜威. MATLAB 数学建模经典案例实践[M]. 北京: 清华大学出版社, 2015: 201-209.
- [6]余胜威. MATLAB 优化算法案例分析与应用[M]. 北京: 清华大学出版社, 2015: 351-359.
- [7]钟壮壮,汪先平.“小康社会”发展历程及其当前问题研究[J].太原理工大学学报(社会科学版),2020,38(01):24-29.
- [8]刘俊艳,张晔,秦善琪.基于层次分析法的新疆 H 县脱贫攻坚成效评价[J].江西农业学报,2020,32(02):145-150.
- [9]易航. 江西省全面建成小康社会监测指标完成度时空分布[D].江西师范大学,2019.
- [10]牛胜强,马秋华.新形势下我国全面建成小康社会进程分析[J].荆楚理工学院学报,2018,33(04):86-91+96.
- [11]高云虹,李帅娜,甘琲.甘肃省全面建成小康社会的进程预测及难点剖析[J].兰州财经大学学报,2018,34(04):12-19+29.
- [12]朱启贵.全面建成小康社会评价指标体系研究[J].人民论坛·学术前沿,2017(04):52-60.



## 附 录

### 附录一

```
import pandas as pd
import numpy as np
from pandas import Series, DataFrame
import matplotlib.pyplot as plt
import numpy as np
import math

def calcMean(x,y):
    sum_x = sum(x)
    sum_y = sum(y)
    n = len(x)
    x_mean = float(sum_x+0.0)/n
    y_mean = float(sum_y+0.0)/n
    return x_mean,y_mean

def calcPearson(x,y):
    #计算Pearson系数
    x_mean,y_mean = calcMean(x,y) #计算x,y向量平均值
    n = len(x)
    sumTop = 0.0
    sumBottom = 0.0
    x_pow = 0.0
    y_pow = 0.0
    for i in range(n):
        sumTop += (x[i]-x_mean)*(y[i]-y_mean)
    for i in range(n):
        x_pow += math.pow(x[i]-x_mean,2)
    for i in range(n):
        y_pow += math.pow(y[i]-y_mean,2)
    sumBottom = math.sqrt(x_pow*y_pow)
    p = sumTop/sumBottom
```

```

return abs(p)
def batchCalcPearson(item,items):
    """
    批量计算Pearson
    """
    res = { }
    for k,v in items.items():
        res[k] = calcPearson(item,v)
    return res
t=1:34;
    y=[98705;100072;101654;103008;104357;105851;107507;109300;111026;1127
04;114333;115823;117171;118517;119850;121121;122389;123626;124761;125786;1
26743;127627;128453;129227;129988;130756;131448;132129;132802;133450;1340
91;134735;135404;136072];
    y=y'/10000;
    %%拟合分析
    myfun=inline('beta(1)./(1+beta(2)*exp(-beta(3)*t))','beta','t');
    beta=nlinfit(t,y,myfun,[15 0.5 0.05])
    a=beta(1);
    b=beta(2);
    c=beta(3);
    %%绘制散点图
    plot(t,y,'b.')
    hold on
    t=1:0.1:40;
    plot(t,a./(1+b*exp(-c*t)))
    %%验证结果
    d=35:37;
    f=[13.6782 13.7462 13.8271];
    plot(d,f,'r*')
    %%加入坐标轴
    [num,txt] = xlsread('1.xlsx');

```

```

xx=0:5:length(num);
set(gca,'xtick',xx);
set(gca,'xticklabel',txt(xx+1,1));
xlabel('年份/年');
ylabel('人口/亿');

grid on

x=[1:1:22];
y=[1975 2744 4515 5974 7711 9692 11791 14380 17205 20438 24324 28018
31161 34562 37198 40171 42638 44653 46473 48467 49970 51091 ];
beta0=[50000,0.8]';
[beta,r,J] = nlinfit(x,y,@myfun,beta0);
beta(1)
beta(2)
figure
plot(x,y,'r+')
hold on;
x_new = [1:1:60];
y_new =
(beta(1))./(1+((beta(1))/1975-1).*exp(-(beta(2)).*x_new)));
plot(x_new,y_new,'k')
x_newqianfu = [15:1:60];
plot(x_newqianfu,myfunnew(49170,0.2864,x_newqianfu),'b')
syms x_new1
y_new1 =
(beta(1))./(1+((beta(1))/1975-1).*exp(-(beta(2)).*x_new1)));
Feng_candidate =
[x_new',((beta(1))./(1+((beta(1))/1975-1).*exp(-(beta(2)).*x_new))))
']

Y_erjie=diff(y_new1,x_new1,2);
x_guai = solve(Y_erjie,x_new1)
y_guai =
(beta(1))./(1+((beta(1))/1975-1).*exp(-(beta(2)).*x_guai)))
DATE={'1/24' '2/3' '2/13' '2/23' '3/4' '3/14' '3/24'};
grid on;
set(gca,'XTickLabel',DATE)
hold on;
plot(11.0815,24585,'ko');hold on;
plot(31,49007,'ko');hold on
plot(51,49170,'ko');hold on

```

```

plot(21,46458, '.', 'Markersize',15);hold on
plot(27,48661, '.', 'Markersize',15);hold on
plot(14,34301, '.', 'Markersize',15);hold on
hold on;
ytr=y';
c=xcorr(y,ytrnew);
figure;
bar(0:length(c)-1,c,1)
for i=1:22
    for j=1:22
        Aire(i,j)=(abs(ytrnew(i)-y(j)));
    end
end
figure
imagesc(Aire)
xx=[ones(22,1),y']
[b1,b2,b3,b4,b5]=regress(ytrnew,xx)

```

## 附录二

人口结构								
2016 年			2015 年		2014 年		2013 年	
年龄	男	女	男	女	男	女	男	女
0-4	3.17	2.74	3.14	2.7	3.07	2.62	3.06	2.61
5-9	2.99	2.52	3	2.52	3.05	2.56	3.03	2.55
10-14	2.83	2.39	2.81	2.37	2.81	2.37	2.78	2.36
15-19	2.87	2.45	2.94	2.53	3.08	2.68	3.23	2.91
20-24	3.57	3.26	3.8	3.5	4.07	3.9	4.56	4.15
25-29	4.68	4.53	4.73	4.62	4.43	4.36	4.17	4.15
30-34	3.81	3.76	3.73	3.65	3.72	3.63	3.75	3.64
35-39	3.54	3.41	3.61	3.47	3.71	3.56	3.85	3.69
40-44	4.17	4.01	4.37	4.19	4.63	4.44	4.74	4.54
45-49	4.59	4.44	4.58	4.43	4.49	4.34	4.46	4.31
50-54	4.27	4.16	3.84	3.74	3.51	3.42	3.16	3.06
55-59	2.61	2.54	2.85	2.75	3	2.9	3.21	3.11
60-64	2.92	2.93	2.84	2.85	2.74	2.74	2.61	2.6
65-69	2.06	2.12	1.99	2	1.83	1.88	1.75	1.76
70-74	1.34	1.39	1.3	1.35	1.29	1.3	1.23	1.27
75-79	0.93	1.01	0.92	1.01	0.91	0.99	0.9	1
80-84	0.56	0.68	0.52	0.66	0.53	0.65	0.51	0.63
85-89	0.23	0.33	0.21	0.31	0.2	0.3	0.19	0.28
90-94	0.05	0.11	0.06	0.1	0.05	0.1	0.05	0.09
95+	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02

## 附录三

影响人口变化因素数据				
年份	出生率	死亡率	自然增长率	人口总数
1978	18.25	6.25	12	
1980	18.21	6.34	11.87	98705
1981	20.91	6.36	14.55	100072
1982	22.28	6.6	15.68	101654
1983	20.19	6.9	13.29	103008
1984	19.9	6.82	13.08	104357
1985	21.04	6.78	14.26	105851
1986	22.43	6.86	15.57	107507
1987	23.33	6.72	16.61	109300
1988	22.37	6.64	15.73	111026
1989	21.58	6.54	15.04	112704
1990	21.06	6.67	14.39	114333
1991	19.68	6.7	12.98	115823
1992	18.24	6.64	11.6	117171
1993	18.09	6.64	11.45	118517
1994	17.7	6.49	11.21	119850
1995	17.12	6.57	10.55	121121
1996	16.98	6.56	10.42	122389
1997	16.57	6.51	10.06	123626
1998	15.64	6.5	9.14	124761
1999	14.64	6.46	8.18	125786
2000	14.03	6.45	7.58	126743
2001	13.38	6.43	6.95	127627
2002	12.86	6.41	6.45	128453
2003	12.41	6.4	6.01	129227
2004	12.29	6.42	5.87	129988
2005	12.4	6.51	5.89	130756
2006	12.09	6.81	5.28	131448
2007	12.1	6.93	5.17	132129
2008	12.14	7.06	5.08	132802
2009	11.95	7.08	4.87	133450
2010	11.9	7.11	4.79	134091
2011	11.93	7.14	4.79	134735
2012	12.1	7.15	4.95	135404
2013	12.08	7.16	4.92	136072
2014	12.37	7.16	5.21	136782
2015	12.07	7.11	4.96	137462
2016	12.95	7.09	5.86	138271