

et 网上投稿

et 网上订阅

录用查询

汇款查询

杂志栏目

● 经济研究

● 西部大开发

● 改革探索

● 新观察

● 理论经纬

● 三农问题

● 热门话题

● 企业论坛

● 区域经济

● 财经论坛

● 对外开放和贸易

● 综合论坛

● 经济全球化

● 产业集群研究

● 社会主义劳动理论探讨

● 面向21世纪的中国经济学

论文正文

中国人口增长模型的建立与预测

上传日期: 2008年3月7日 编辑: 现代经济编辑部 点击:190次

王喜林, 刘丽红

(吉林省长春市吉林农业大学信息技术学院数学教研室,吉林长春130118)

摘要: 根据中国现有人口情况, 例如出生人口性别比持续升高及现阶段的低生育水平不稳定等系列问题, 建立中国人口预测模型。本模型从多元线性回归函数[1]着手, 最终建立起中国人口总数Y与时间X的函数。过程中应用大量可靠的人口数据[2], (其中人口数量单位均为: 万人)。以及运用TSP软件, MATLAB数学软件和EXCEL等工具, 对已知的大量数据进行整理, 统计, 采用最小二乘法进行数据的拟合, 以误差最小为基准, 择优选取反映出人口数量变化的最佳因子, 建立起影响各因子与人口数量之间的关系子模型, 将人口函数引领到多元函数上来。通过各种因子的综合作用, 最终反映到我们建立的中国人口总数与时间的模型上, 从而能够起到有效预测的目的。

关键词: 中国人口; 模型 ;预测

一、问题分析

人口问题始终是制约我国发展的关键因素之一。预测模型首先应符合人口繁衍变化的自然特征, 人口的变化总是由生育、死亡和人口迁移三大基本要素所决定。在人口生存的社会环境相对稳定(即没有战争及毁灭性的灾难)的情况下, 城乡之间人口将不会出现大量的人口迁入与迁出, 总人口增长是缓慢变化的。其次, 预测模型必须反映人口随时间变动而变动的特征。总人口数量及影响因素众多, 需要进行数据的统计和处理并找出规律, 继而建立合理的预测模型。

二、模型假设

- 1、决定出生率的因素为城市与乡镇的女性人数。
- 2、决定死亡率的因素为城市与乡镇的人口数。
- 3、城镇与乡村人口在预测期内不会发生大规模的迁移。
- 4、城乡女性比例相同。

三、模型的建立及求解

人口的增长有很多影响因素, 比如性别比、国家政策等等。但最直接的影响因素是出生率和死亡率, 因此选取主要影响因素, 建立起在动态中各影响因素与人口增长的预测模型。最终归结到时间与人口数量的关系上, 于是, 要找出各种影响因素与时间之间的关系规律, 从而建立起比较完整的模型网络。

首先, 由总出生率、总死亡率和人口数量建立起多元线性回归函数。人口性别比对人口数有较大影响, 分别建立总出生率与城市女性人数及乡镇女性人数; 总死亡率与城市死亡率及乡镇死亡率的多元线性回归函数。然后, 利用分别建立起城市女性人数、乡镇女性人数、城市死亡率、乡镇死亡率与时间的一元函数关系, 并求出关系式。如下所示, 为初步关系式1-3。

$$1、y = my_1 + ny_2 + c \quad (y_1 = \text{总出生率}, y_2 = \text{总死亡率})$$

$$2、y_1 = ax_1 + bx_2 + c_1 \quad y_2 = dx_3 + ex_4 + c_2$$

(x_1 =城市女性人数, x_2 =乡镇女性人数, x_3 =城市死亡率, x_4 =乡镇死亡率)

[m,n,a,b,d,e,c,c1,c2,全为方程系数]

$$3、x_1=f(x) \quad x_2=g(x) \quad x_3=h(x) \quad x_4=m(x) \quad (x \text{为时间})$$

分别建立城市女性人数与时间的关系；乡镇女性人数与时间的函数关系。均采用《中国人口年鉴》1978-1999年的城市女性人数与乡镇女性人数。拟合得到较为理想的关系式。二次函数的可信度比较高： $R^2=0.9981$ 。故选用 $x_1=-3.0294x^2+12560x-12982901$ 。一次函数与二次函数的可信度差不多；但随着时间的推移，其抛物线上升速率加快过多，可见一次函数与实际更为相符，故选取 $x_2=241.28x-439411$ 。分别计算相应拟合结果，该拟合结果与实际数值误差很小，平均误差约为0.0579%。

建立城市女性人数和乡镇女性人数与总出生率之间的多元函数关系；并计算出拟合的总出生率结果。

用Micro TSP软件包编程来求多元方程，得到以常量序列（C）、城市女子数（X1）、乡镇女子数（X2）为解释变量，总出生率（Y1）为被解释变量执行普通最小二乘法运算。计算完成后得到的C、X1、X2分别为1117.8529、0.0148560、-0.032319，其T-STAT.分别为6.6813280、6.3393649、-6.5290783；同时得到F-statistic为32.76899。由于 $k=2$ ，即为变量数， $n=17$ ，即为17组数据，给定一个显著性水平 $\alpha=0.01$ ，查F分布表，得到一个临界值 $F_{0.01}(2,14)=6.51 < 32.76899$ 。即 $F > F_{\alpha}(k,n-k-1)$ ，通过了方程显著性检验。而查表得 $(n-k-1)=2.977$ ，则有 $|t_0|$ 、 $|t_1|$ 、 $|t_2|$ 均大于其值，说明所有变量显著，该方程有很好的拟合度。拟合方程为： $y_1=1117.8529+0.0148560x_1-0.032319x_2$ 。

分别建立城市人口死亡率与时间函数关系；乡镇人口死亡率与时间的函数关系。采用《中国人口年鉴》1989-1999年的城市人口死亡率及乡镇人口死亡率。

城市人口死亡率与时间关系。二次函数与实际明显不符，一次函数更接近实际，则选用 $x_3=-0.031x+67.438$ 。乡镇人口死亡率与时间关系，一元函数更接近实际。故采用线性函数 $x_4=-0.0016x+10.197$ 。

用拟合函数计算城市人口死亡率、乡镇人口死亡率、并分别计算两者与实际数据的误差。平均误差约为0.1663%。

用曲线拟合发现城市死亡率与乡镇死亡率具有线性相关性，则建立城市人口死亡率与总人口死亡率之间的函数关系。采用《中国人口年鉴》1989-1999年的总死亡率和1989-1999年的城市死亡率拟合值。得函数关系 $y_2=0.5513x_3+3.4703$ 。

建立总出生率和总死亡率与总人口的多元函数模型

采用1989-1999年拟合的人口总出生率、总死亡率和1989-1999年实际总人口。数据

见表5。以常量序列（C）、总出生率（Y1）、总死亡率（Y2）为解释变量，总人口（Y）为被解释变量执行普通最小二乘法运算。计算完成后得到的C、Y1、Y2分别为738645.39、466.09508、-95499.129，其T-STAT.分别为80.7478、12.739362、-64.001005；同时得到F-statistic为40814.00。给定一个显著性水平 $\alpha=0.01$ ，查F分布表，得到一个临界值 $F_{0.01}(2,8)=8.65 < 40814.00$

即 $F > F_{\alpha}(k,n-k-1)$ ，通过了方程显著性检验。而查表得 $(n-k-1)=3.355$ ，则有 $|t_0|$ 、 $|t_1|$ 、 $|t_2|$ 均大于其值，说明所有变量显著，该方程有很好的拟合度。

方程为： $Y=738645.39+466.09508y_1-95499.129y_2$ 。

最后的时间与人口数的总函数

对以上各函数进行归纳汇总，归结到时间与总人口的函数表达式上：即 $Y=738645.39+466.09508y_1-95499.129y_2$ 中 y_1 用 $y_1=1117.8529+0.0148560x_1-0.032319x_2$ 代替； y_2 用 $y_2=0.5513x_3+3.4703$ 代替；然后 x_1 用 $x_1=-3.0294x^2+12560x-12982901$ 代替， x_2 用 $x_2=241.28x-439411$ 代替。 x_3 用 $x_3=-0.031x+67.438$ 代替。最后，得到函数关系式： $Y=-20.9765x^2+84967.105x-85901226$ 。

四、模型检验及预测

分析上述已建模型，用MATLAB软件可得：2000年126743万人，2010年135500万人，2020年139820，2030年139940万人。（见下表）1995年到2006年预测人口与实

际人口的平均误差为0.38%。所建人口预测模型在近几年里与真实数据间的误差变化在误差准许范围之内。由上述模型的检验，可以看出以上建立的人口预测模型基本上可以反映出中国人口总数的变化趋势。该模型反映的人口变化过程是平稳的，并且误差很小，很好的预测出近年的人口数量，起到了有效的预测目的，则可以认为，该模型可以作为未来的人口预测模型。

1995-2030年预测人口与实际人口及误差

年份

1995

1996

1997

1998

1999

2000

实际人口

121121

122389

123626

124761

125786

126743

预测人口

121160

122410

123620

124780

125910

126990

误差 (%)

0.0322

0.0172

-0.0049

0.0152

0.0985

0.1949

年份

2001

2002

2003

2004

2005

2006

实际人口

127627

128453

129227

130000

130631

130756

预测人口

128030

129030

129980

130900

131770

132600

误差 (%)
0.3158
0.4492
0.5827
0.6923
0.8719
1.4103

年份

2007
2008
2009
2010
2011
2012

预测人口

133390
134130
134840
135500
136120
136700

年份

2013
2014
2015
2016
2017
2018

预测人口

137240
137730
138180
138590
138960
139290

年份

2019
2020
2021
2022
2023
2024

预测人口

139570
139820
140020
140180
140290
140370

年份

2025
2026
2027
2028
2029
2030

预测人口

140400

140390
140340
140250
140120
139940

参考文献:

[1] 李子奈.[ISBN 7-04-008346-9].计量经济学.北京:高等教育出版社, 2003.44—298

[2] 1982年以前数据为户籍统计数; 1982—1989年数据根据1990年人口普查数据有所调整; 1990-2000年数据根据2000年人口普查数据进行了调整; 2001-2004年数据为人口变动情况抽样调查推算数; 2005年数据根据全国1%人口抽样调查数据推算。

作者简介: 王喜林, 吉林农业大学信息技术学院, 讲师。

版权所有: 《现代经济》编辑部
E-MAIL: mejc@vip.sohu.com 电话: 0898—68928581 传真: 0898—68919810
地址: 海口市龙昆北路24号龙园别墅D1栋 邮编: 570105