

优化的 Leslie 人口预测模型及其应用

石敏军,张佛德,张晓鹏

(昆明理工大学 理学院,云南 昆明 650500)

摘要:中国是一个人口大国,人口与经济和社会的发展密切相关,因此准确地预测人口信息,是制定未来人口发展目标及生育政策等有关人口政策的基础,对研究制定经济和社会的发展决策具有重要的参考价值。文中由中国人口信息网的一些数据,利用 Excel 进行统计分析,得出了当前我国人口的年龄分布结构及各年龄阶层相关指标(死亡率、生育率、男女性别比例等)的期望值。在这些数据的基础上,运用优化的 Leslie 人口预测方法建立了预测中国人口数量和年龄结构特点的数学模型。并根据中国的实际情况,依算法分别设计了 Mathematical 程序预测长期的人口数量和分布结构。结果表明,优化的 Leslie 模型预测结果准确可靠,为人口控制提供了准确有效的理论依据。

关键词:Leslie 预测法;年龄结构;死亡率;出生率;Mathematical

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2013)09-0258-04

doi:10.3969/j.issn.1673-629X.2013.09.066

Optimized Leslie Population Forecast Model and Application

SHI Min-jun, ZHANG Fo-de, ZHANG Xiao-peng

(Faculty of Science, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: As a populous nation, predicting the increase of population in China is of great importance for the close relationship between population and the development of economy and society. To project the information on population accurately is the basic work for making the population policy, such as the goal of the future population development and reproductive policy, which is also of great reference value in making the economic and social development decision. By the China population information website data, analyze with Excel and conclude current Chinese people's age distribution structure and the expectations of the related indexes of different ages, such as mortality, fertility, gender ratio etc. On the basis of these data, not only establish a mathematical model featuring in predicting Chinese population size and age structure by Leslie's prediction means, but also design the mathematical program on the purpose of predicting long-term of population size and distribution structure in accordance with the actual situation in China. The results show that the optimized Leslie model prediction results are accurate and reliable, for the population control provide the theory accurate and effective basis of theory.

Key words: Leslie forecast method; age structure; mortality; birth rate; Mathematical

0 引言

人口问题是人类社会发展的核心问题,中国是世界上人口最多的发展中国家。人口多,底子薄,人均耕地少,人均占有资源相对不足,是我国的基本国情,人口问题一直是制约中国经济发展的首要因素^[1]。

基于此,必须建立一个有效的人口预测数学模型^[2],显然不同的方法得出的结果会各有差异。所熟悉的对人口估计的模型有马尔萨斯和 Logistic 模型,但这两种模型只能在比较理想的情况下进行预测。一旦考虑到出生率、死亡率、性别比例及年龄结构等问题时便不凑效,尤其是在进行长期预测时更不理想。

基于这一事实,文中从中国的实际情况和人口增长的特点出发,参考相关数据(也可以搜索相关文献和补充新的数据),利用优化的 Leslie^[3] 预测法建立一个考虑人口增长和年龄结构的离散模型,并由此对中国人口增长的长期趋势做出预测。

1 基本假设

(1) 计划生育政策落到实处,且将长期执行下去,在此假设的前提下。育龄妇女的生育率将处于一个相对稳定的状态。

(2) 在一个相当长的时期,中国人口的迁移主要

收稿日期:2012-11-13

修回日期:2013-02-23

网络出版时间:2013-04-22

基金项目:国家自然科学基金资助项目(11101193)

作者简介:石敏军(1987-),男,硕士研究生,研究方向为组合最优化及其应用。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20130422.1553.010.html>

为城乡之间的迁移。

(3)假设人为控制新生儿性别的行为被打击,新生儿的性别出身比率为自然比率。

(4)所有的妇女均能婚配生育,且每胎只生一个。

(5)各年龄阶段(以年为单位)的男女比率保持相对稳定。

(6)各年龄阶段(以年为单位)的死亡在一个不太长的时期内相对稳定。

2 符号说明

1) $X_i = (x_i^0, x_i^1, \dots, x_i^{90})^T, i = 0, 1, \dots, k, x_i^j (j = 0, 1, \dots, 90)$ 为第 i 年 j 年龄段女性数量。

2) $r = (r_0, r_1, \dots, r_{90})^T, r$ 为男性所占比例向量, r_i 为第 i 年龄段男性占总人口数的比例。

3) $s = (s_0, s_1, \dots, s_{90})^T, s$ 为女性所占比例向量, s_i 为第 i 年龄段女性占总人口数比例。

4) $q = (q_0, q_1, \dots, q_{90})^T, q$ 表示各年龄段女性与该年龄段人口总数的比值向量, q_i 表示第 i 年龄段的女性与该年龄段人口总数的比例。

5) $p = (p_0, p_1, \dots, p_{90})^T, p$ 表示各年龄段女性的存活率,存活的女性即转入下一年龄段。

6) $u = (u_0, u_1, \dots, u_{90})^T, u$ 为男性死亡率的向量, u_i 为第 i 年龄段对应的男性死亡率。

7) $v = (v_0, v_1, \dots, v_{90})^T, v$ 为女性死亡率的向量, v_i 为第 i 年龄段对应的女性死亡率。

8) $t = (t_0, t_1, \dots, t_{90})^T, t_i$ 表示第 i 年龄段女性的生育率(以‰为单位)。

9) $f = (f_0, f_1, \dots, f_{90})^T, f_i$ 表示第 i 年龄段的女性生育女孩的生育率(以‰为单位)。由中国人口信息网的数据可知: $f_0 = \dots = f_{14} = f_{49} = \dots = f_{90} = 0$ 。

10) $d = (d_0, d_1, \dots, d_{90})^T, d$ 表示各年龄段男性数和女性数的比值向量, d_i 表示第 i 年龄段男性和女性人口数的比值。

11) α ---表示新生幼儿男女性别比($\alpha \geq 100$)。
 β ---表示新生幼儿女孩的死亡率(以‰为单位)。
 m ---表示人口的总数量。
 $G_i = (g_i^0, g_i^1, g_i^2, \dots, g_i^{90})^T$ 表示外来人口数^[4]。

3 模型建立

由模型假设^[5],讨论人口的数量变化和分布结构时,主要针对女性进行分析。将人口分为0~89及90以上91个阶段,各年龄阶段个体数的分布定义为: $X_i = (x_i^0, x_i^1, \dots, x_i^{90})^T, i = 0, 1, \dots, k, x_i^j (j = 0, 1, \dots, 90)$ 为第 i 年 j 年龄段女性数量。

初始值, $X_0 = (x_0^0, x_0^1, \dots, x_0^{90})^T$, 一年后, $X_1 =$

$(x_1^0, x_1^1, \dots, x_1^{90})^T$, 两年后, $X_2 = (x_2^0, x_2^1, \dots, x_2^{90})^T, \dots, k$ 年后, $X_k = (x_k^0, x_k^1, \dots, x_k^{90})^T$ 。

经过一年后, X_0 各年龄段产生的下一代进入 X_1 的第一年龄段,即有 $x_1^1 = f_1 \cdot x_0^0 + f_2 \cdot x_0^1 + \dots + f_{90} \cdot x_0^{90}$, 其中 $f_i (i = 0, 1, \dots, 90)$ 为各年龄段女性生育女孩的生育率(以‰为单位)。而 X_0 的第 $i (i = 0, 1, \dots, 90)$ 年龄段中有存活量 $p_i \cdot x_0^i$ 进入 X_1 的 $i + 1$ 年龄,即 $x_{i+1}^1 = p_i \cdot x_i^0, p_i$ 为第 i 年龄段的存活率。由此,得矩阵表示形式:

$$X_1 = \begin{pmatrix} x_1^0 \\ x_1^1 \\ \vdots \\ x_1^{90} \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & \cdots & f_{89} & f_{90} \\ p_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{89} & 0 \end{pmatrix} \begin{pmatrix} x_0^0 \\ x_0^1 \\ \vdots \\ x_0^{90} \end{pmatrix} + \begin{pmatrix} g_0^0 \\ g_0^1 \\ \vdots \\ g_0^{90} \end{pmatrix}$$

即 $X_1 = MX_0 + G_0$, 其中

$$M = \begin{pmatrix} f_0 & f_1 & \cdots & f_{89} & f_{90} \\ p_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{89} & 0 \end{pmatrix}$$

类推得: $X_{k+1} = MX_k + G_k, M$ 就是著名的 Leslie^[6] 矩阵。但考虑到90岁以上(含90岁)老人的存活率仍大于零($p_{90} \neq 0$)且他们不再进入别的年龄段,故这里有必要将 Leslie 矩阵优化为:

$$B = \begin{pmatrix} f_0 & f_1 & \cdots & f_{89} & f_{90} \\ p_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{89} & p_{90} \end{pmatrix}$$

故 $X_{k+1} = BX_k + G_k$ 。

设初始时刻各年龄段的人口数向量为: $m = (m_0, m_1, \dots, m_{90})^T$ 。由前面的假设及符号说明可以得到:各年龄段女性存活率 $p = 1 - \frac{v}{1000}$, 新生幼儿女孩所占的比例为 $\frac{100}{100 + \alpha}$ 。

于是在假设各年龄段育龄女性生育男女的自然比例相同的前提下得到:各年龄段女性生育女孩的生育率为 $f = \frac{100}{100 + \alpha} \cdot \frac{t}{1000}$ 。

由假设得初始时刻各年龄段女性人口数量 $X_0 = m_0, s$, 其中 $X_0 = (x_0^0, x_0^1, \dots, x_0^{90})^T, m = (m_0, m_1, \dots, m_{90})^T, s = (s_0, s_1, \dots, s_{90})^T$ 。在这里定义如

下一种运算： $(m_0, m_1, \cdots, m_{90})^T$ 。 $(s_0, s_1, \cdots, s_{90})^T = (m_0 s_0, m_1 s_1, \cdots, m_{90} s_{90})^T$ 。于是便得到： $x_0^i = m_i s_i$ ($i = 0, 1, \cdots, 90$)。

根据优化的 Leslie^[7] 预测模型即可得到第 k 年后各年龄段的女性人口数, $X_1 = BX_0 + G_0$, 这里 $B = \begin{pmatrix} f_0 & f_1 & \cdots & f_{89} & f_{90} \\ p_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{89} & p_{90} \end{pmatrix}$, 得 $X_{k+1} = BX_k + G_k$ 。

4 模型求解

显然, 预测较长时间的人口数量及其年龄结构有很大的困难^[8], 因为中国正处在一个高速发展的时期, 各个方面都有可能在很短的时间内发生巨大的变化^[9]。因而长期的人口出生率和死亡率有很大的不可预测性。

这样就导致精确预测长期的人口数量成为一个困难的问题。但是, 还是可以通过合理的调整优化的 Leslie^[10] 预测模型来进行长期的预报模拟。

- 这里假设:
- 1) 在相当的时间内, 人口的总出生率控制在 1.8 左右。
 - 2) 人口是男女出生比例达自然值 104 左右。
 - 3) 死亡率大幅下降, 保持估计下降 (30%)。
 - 4) 在城乡的男女出生比例相同的条件下, 城乡的人口流动对数据的影响可以忽略, 据此建立预报较长时间中国人口数量和基本年龄结构^[11]。

这里参数均选择各指标数的加权平均值, 依算法编写 Mathematical 程序^[12], 运行程序可得各预测数据如下:

用它来预测连续 40 年 (从 2006 年起) 的人口的总数为:

{ 132 016. , 133 214. , 134 346. , 135 415. , 136 451. ,
137 482. , 138 522. , 139 588. , 140 674. , 141 773. ,
142 854. , 143 889. , 144 863. , 145 748. , 146 543. ,
147 241. , 147 831. , 148 316. , 148 693. , 148 951. ,
149 101. , 149 155. , 149 144. , 149 093. , 149 014. ,
148 916. , 148 801. , 148 672. , 148 518. , 148 351. ,
148 168. , 147 968. , 147 759. , 147 533. , 147 287. ,
147 022. , 146 720. , 146 378. , 145 987. , 145 539. }

用 Mathematical 做出的 2006 年 ~ 2046 年人口总数变化趋势如图 1 所示 (横坐标对应年份, 纵坐标对应人口数, 单位: 万人)。

这里可以看出将在 2030 年左右达到峰值, 那时的人总的数将近达 148 916 万。

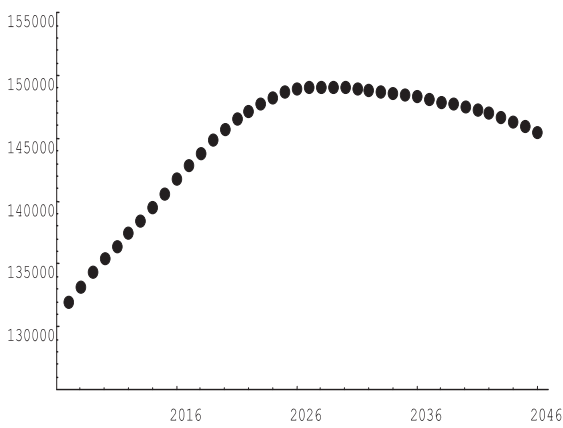


图 1 2006 年 ~ 2046 年人口总数变化趋势

用它预测连续 40 年 (从 2006 年起) 的育龄女性人数为:

{ 10 915.4, 10 545.5, 10 210.1, 10 065.2, 10 051.7,
10 210.7, 10 488.9, 10 793.9, 11 051.1, 11 231.2,
11 350.2, 11 475.8, 11 495.9, 11 476.4, 11 314.4,
11 024.1, 10 590.2, 10 036. , 9 434.97, 8 858.6,
8 322.8, 7 970.13, 7 888.8, 7 878.74, 7 876.89,
7 914.06, 7 992.44, 8 121.78, 8 311.89, 8 565.45,
8 891.88, 9 267.53, 9 551.38, 9 603.41, 9 661.58,
9 713.96, 9 747.53, 9 748.85, 9 705.59, 9 608.29,
9 453.62, 9 245.2, 8 994.54, 8 720.68, 8 444.95,
8 185.3, 7 954.12, 7 758.43, 7 602.17, 7 487.51 }

对应的 2006 年 ~ 2046 年育龄女性人口总数变化趋势如图 2 所示 (横坐标对应年份, 纵坐标对应人口数, 单位: 万人)。

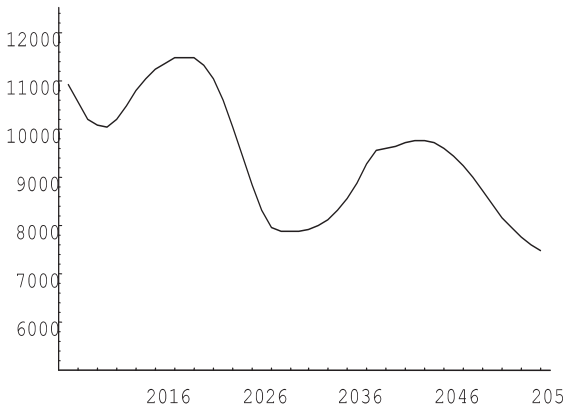


图 2 2006 年 ~ 2046 年育龄女性人口总数变化

根据数据, 可以得出大约经过 10 年 (即 2016 年) 和 35 年 (即 2041 年) 都会出现生育的高峰期。

预测所得的 20 年后女性的年龄结构预测数据^[13] 比较多, 这里不便给出, 它将为预测生育的高峰期作参考, 这里只做出它的折线图供参考, 见图 3 (横坐标对应年龄段, 纵坐标对应该段的人数, 单位: 万人)。

预测连续 40 年 (从 2006 年起) 的老年人所占总人口数的比例数值为 (单位: 1):

{0.085 743 1,0.088 732 9,0.091 677,0.094 521 9,
0.097 821 4,0.100 569,0.103 577,0.106 613,
0.109 995,0.113 906,0.117 927,0.122 398,
0.127 307,0.132 582,0.138 871,0.144 484,
0.150 432,0.156 029,0.160 744,0.164 918,
0.168 046,0.171 46,0.176 085,0.181 764,
0.188 713,0.196 319,0.204 024,0.211 232,
0.218 231,0.225 332,0.232 041,0.238 535,
0.244 427,0.249 251,0.253 575,0.256 478,
0.258 626,0.259 767,0.260 262,0.260 378}

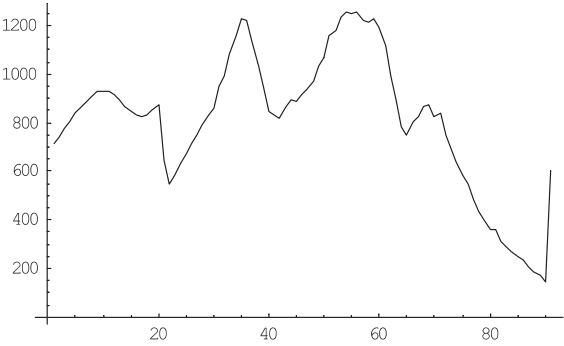


图3 20年后女性的年龄结构

中国人口继续老龄化,到2046年中国的65岁以上的老人占总人数的比例将接近26%,65岁以上的老年人占总人口数的比率的变化曲线图如图4所示(横坐标对应年份,纵坐标对应其比率)。

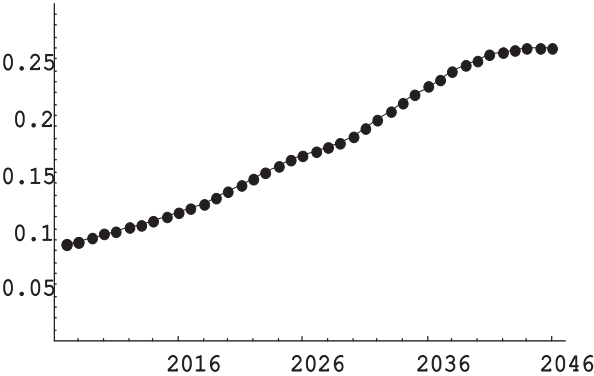


图4 65岁以上的老年人占总人数的比率的变化曲线

5 结束语

文中参考了当前人口的几种预测方法,在 Leslie 模型基础上进行改进,考虑了出生率、死亡率、性别比例及年龄结构等问题,提出了优化的 Leslie 模型。理论与实验分析结果表明,文中提出的优化的 Leslie 模型较之现有人口预测模型模拟的精度要高。应用该模

型对人口进行预测,既能体现出人口增长方式的多样性,又能避免单一模型预测的不足,具有较大的科学性和适用性。

由于人口预测处理方法的复杂性,它不仅依赖于科学的理论、方法和可靠的数据,还依赖于预测者的经验和判断能力。

所以不可能对未来的数据做太多的预测。只能根据现有的数据中的规律(如死亡率)做近似的模拟,所以数据的偏差是允许的。而对于长期的数据根据前面的年份做了修正,但也不认为这是不对的,因为是根据中国人口网中的权威数据预测及有关年龄结构的变化趋势对数据做了函数模拟,最后使数据更接近权威预测值。

参考文献:

[1] Peron G. Studying dispersal at the landscape scale: efficient combination of population surveys and capture-recapture data[J]. Ecology, 2010, 91(11): 3365-3375.

[2] 波拉德. 人口增长的数学模型[M]. 姚志坚, 译. 成都: 四川大学出版社, 1988.

[3] Caceres M O. Random Leslie matrices in population dynamics[J]. Journal of Mathematical Biology, 2011, 63(3): 519-556.

[4] 李勇胜. 人口预测中的模型选择与参数认定[J]. 财经科学, 2004(2): 68-72.

[5] 韩中庚. 数学建模方法及应用[M]. 北京: 高等教育出版社, 2005.

[6] Cushing J M. Three stage semelparous Leslie models[J]. Journal of Mathematical Biology, 2009, 59(1): 75-104.

[7] Du Yihong, Peng Rui, Wang Mingxin. Effect of a protection zone in the diffusive Leslie predator-prey model[J]. Journal of Differential Equations, 2009, 246(10): 3932-3956.

[8] 中国人口网. 国家人口发展战略研究报告[R/OL]. 2007-02-01. http://www.chinapop.gov.cn/fzzlbg/bgyw/t.20070111_172058153.html.

[9] 国家统计局人口和社会科技统计司. 中国人口统计年鉴[M]. 北京: 中国统计出版社, 2006.

[10] Gosselin F, Lebreton Jean-Dominique. Asymptotic properties of infinite Leslie matrices[J]. Journal of Theoretical Biology, 2009, 256(2): 157-163.

[11] 国务院人口普查办公室, 国家统计局. 中国 2000 年人口普查国际研讨会论文集[C]. 北京: 中国统计出版社, 2005.

[12] 胡守信, 李柏年. 基于 MATLAB 的数学试验[M]. 北京: 科学出版社, 2004.

[13] 谢韦克, 黄荣清. 中国妇女生育模型研究[J]. 人口与经济, 1993(1): 35-40.

优化的Leslie人口预测模型及其应用

作者：[石敏军](#)，[张佛德](#)，[张晓鹏](#)，[SHI Min-jun](#)，[ZHANG Fo-de](#)，[ZHANG Xiao-peng](#)

作者单位：[昆明理工大学 理学院, 云南 昆明, 650500](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

ISTIC

年，卷(期)：2013(9)

本文链接：http://d.wanfangdata.com.cn/Periodical_wjfz201309066.aspx