

文章编号:1005-3085(2007)08-0071-08

中国人口增长的常微分方程模型

任 可, 唐 云

(清华大学数学科学系, 北京 100084)

摘 要: 本文利用已有的数据资料, 从中国人口增长的特点出发, 将城镇与农村中不同性别的人口按不同的年龄段分别处理, 并引进农村人口向城镇迁移的因素, 建立起一个关于中国人口增长的常微分方程组初值问题的数学模型。还利用该模型对中国未来人口的增长变化进行了预测。

关键词: 出生率; 死亡率; 城镇化

分类号: AMS(2000) 91D20

中国分类号: O29

文献标识码: A

1 前言

人口问题的研究一直是人们所关注的一个课题。关于人口增长问题已经有多方面的研究, 如马尔萨斯指数增长模型、logistic 人口阻滞增长模型、偏微分方程模型和 Leslie 模型^[1], 以及 Lotka 积分模型^[2]等。我国是一个人口大国, 在中国的人口增长的数学模型研究方面也有颇多的工作, 如宋键和于景元^[3]、王彦^[4]和霍江明^[5]等。特别在2007年初发布的《国家人口发展战略研究报告》^[6]报告, 对于中国人口的发展现状、未来人口增长的预测、人口素质、人口教育、就业问题以及中国人口现阶段增长等特点进行了总结。本文基于该文章及《中国人口统计年鉴》^[7]等资料所提供的数据信息建立起一个具有中国人口增长特点的常微分方程初值问题的数学模型, 并利用该模型对中国未来人口的增长变化进行预测。

2 基本假设

人口模型的建立必须基于现有人口的增长特点(即人口情况的基本假设)。据《国家人口发展战略研究报告》等资料的分析, 当前中国人口的增长有着如下特点:

- 1) 中国正在进入老龄化社会, 老年人的比重在不断加大;
- 2) 农村与城镇育龄妇女的生育率及出生人口性别比有着较大差异;
- 3) 农村人口不断城镇化(《国家人口发展战略研究报告》估计转化率为百分之一)。

我们知道, 人口的增长率取决于出生率、死亡率和不同人群之间的迁移率。而出生率又取决于育龄妇女的生育率及育龄妇女在总人口中所占的比例。因此我们需要首先对这些相关数据进行分析。

1) 育龄妇女的生育率: 图1是1994—2005年十年间城镇育龄妇女的生育率, 从中可看出虽然成下降趋势, 但是下降已趋于缓和。农村育龄妇女的生育率亦有类似情形, 但总体数据偏高。因此我们有必要将城镇人口和农村人口分离开来讨论。

2) 出生人口性别比: 图2为城镇和农村在1994-2005年期间的出生人口男女性别比例(以每100个女性有多少男性来记录的)。通常的比例应当是105:100左右, 而该图的数据远超过了这个比例。这将导致未来育龄妇女在总人口中的比例减少, 从而造成总出生率下降。因此我们有必要将出生人口的男女性别比因素考虑进去。

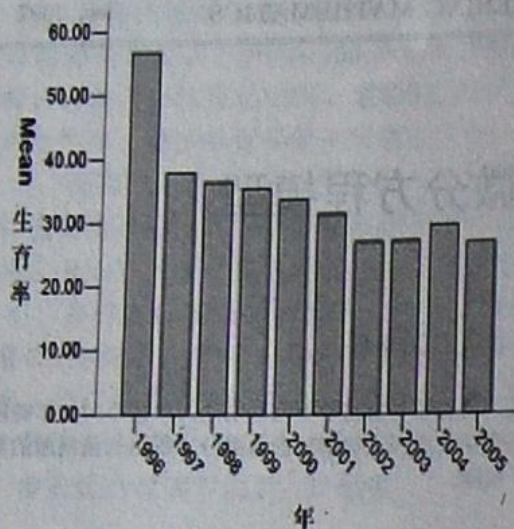


图1: 1994-2005年育龄妇女生育率(千分比)

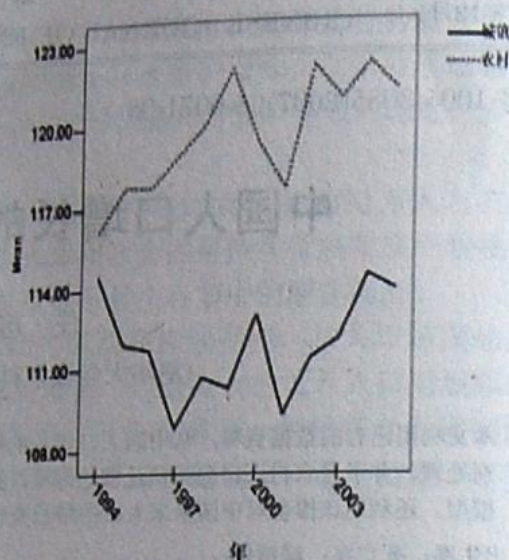


图2: 1994-2005年男女出生比例(以女性为100记)

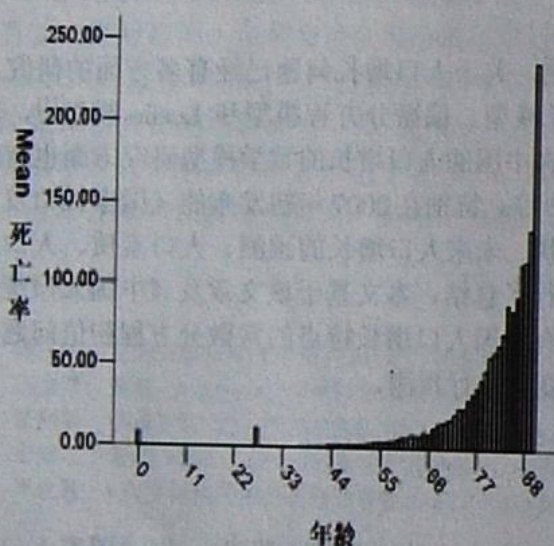


图3: 城镇女性2005年死亡率(千分比)

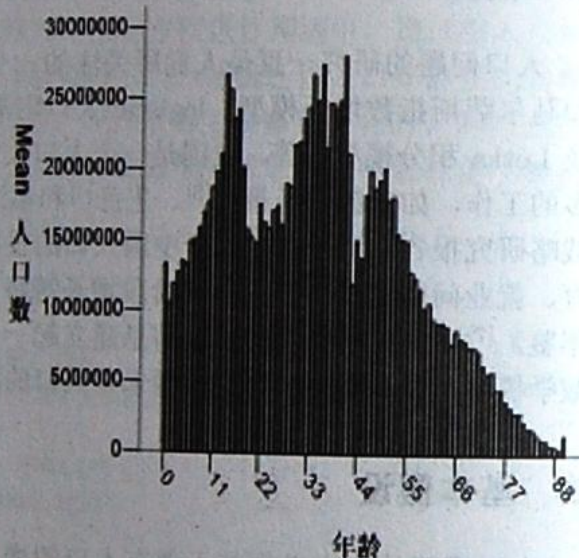


图4: 2005年分年龄段人口数量

3) 死亡率: 以城市妇女2005年各年龄段的死亡率(千分比)为例(城市男及农村男女也类似)。从图3中我们可以清楚地看到不同年龄段的死亡率是不同的: 0-3岁孩子死亡率较高, 3-50岁死亡率较低, 50岁以上人口的死亡率不断提高。这说明有必要将死亡率分年龄段考虑在模型中。

4) 人口的年龄分布: 图4是2005年分年龄段的人口数量(其他年有类似情形), 从中我们可以看出在过去50年的人口出生的变化情况。比如说, 20岁的人口与20年前(即1985年)出生的人口很接近。特别可以看出在15年前(即1990年)的前后和35年前(即1970年)的前后均出现过人口生育高峰。进一步的数据分析表明, 这两次生育高峰主要都出现在农村。也看得出最近15年来人口出生的下降趋势。而50岁以下的人口由于死亡率很低, 这些人成批进入老年阶段将使得老龄人口的比重不断加大。

5) 农村人口城镇化: 由于农村和城镇人口出生率的较大差异, 大量的农村人口转化为城镇人口对中国人口的增长将造成很大影响, 因此必须把这个因素考虑进去。《国家人口发展战略

略研究报告》预测农村人口的城镇化水平是年增长1%。但考虑到在城镇中女性在服务业和婚姻方面较男性的优势, 我们假设农村男性以0.9%的比率转化, 而女性以1.1%转化。

据此, 我们可以把中国人口分成城镇男、城镇女、农村男和农村女四类分别处理, 并引进农村人口向城镇迁移的因素。

3 模型的建立

3.1 人口分年龄段比例的未来走向

由于在讨论到出生率和死亡率时都涉及到不同年龄段的问题, 所以需要对上述每一类人口在各年龄段所占的比例分别做出估计。为方便起见, 我们对每一类人口都按其是否参与生育分为: 幼年(0-14岁)、青壮年(15-49岁)和老年(50岁及以上)三个年龄段人口的比例来处理。当然, 这些量并不是互相独立的, 比如, 城市男性人口应等于城市男的幼、青壮和老年三类人口总和, 余类推。通过对上述数据的分析及相关资料(如《世纪之交的中国人口》^[8])中的一些讨论, 可以看出, 老龄化的结果造成老年人口增加, 而幼年和青壮年的人口在总人口中所占的比例则逐步减少。数据分析还表明, 育龄妇女的生育年龄主要集中在20至35岁, 且其生育率趋于常数。

以城镇女性人口为例, 我们将城镇女性人口按年龄分作81维向量, 按已知各年龄的人口比例、死亡率和育龄妇女的生育率, 利用 Leslie 矩阵^[4]的方法进行迭代。与通常 Leslie 矩阵方法不同的是, 在迭代过程中我们所取的是人口比例而非人口数量。这就可以得到未来60年的人口年龄比例曲线。具体步骤如下:

设向量 $A(n) = [a_0(n), a_1(n), \dots, a_{80}(n)]^T$ 各分量依次为第 n 年的城镇0至80岁女性人口的比例, 并设

$$B(n) = \begin{bmatrix} b_0(n) \\ b_1(n) \\ \vdots \\ b_{80}(n) \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & b & \cdots & b & 0 & \cdots & 0 \\ 1-c_0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 1-c_1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1-c_{79} & 1-c_{80} \end{bmatrix} A(n-1),$$

其中矩阵第一行从第16列到50列为常数 b (城镇育龄妇女的女婴生育率), 而 c_i 为 i 岁城镇女性人口的死亡率,

$$A(n) = B(n) / \sum_{i=0}^{80} b_i(n).$$

取2005年的百分之一人口抽样调查数据为初值, 通过 Excel 进行迭代, 并按上述幼年、青壮年和老年三个年龄段人口的比例来表示, 得到图5。

从图5中我们可以看出, 就城镇妇女的幼年阶段与作为青壮年的育龄阶段所占的比例而言, 其变化均呈指数下降趋势。实际上由于这两个年龄段人口的总体死亡率都很小, 其每次迭代相当于乘以一个小于1的“常数”。

而这一点我们也可以粗略地证明一下: 把人口分成幼年 x_1 、青壮年 x_2 和老年 x_3 三部分, 得到的 $x = (x_1, x_2, x_3)^T$ 满足

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_n = \begin{bmatrix} 1-\alpha & b & 0 \\ \alpha & 1-\beta & 0 \\ 0 & \beta & 1-c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{n-1},$$

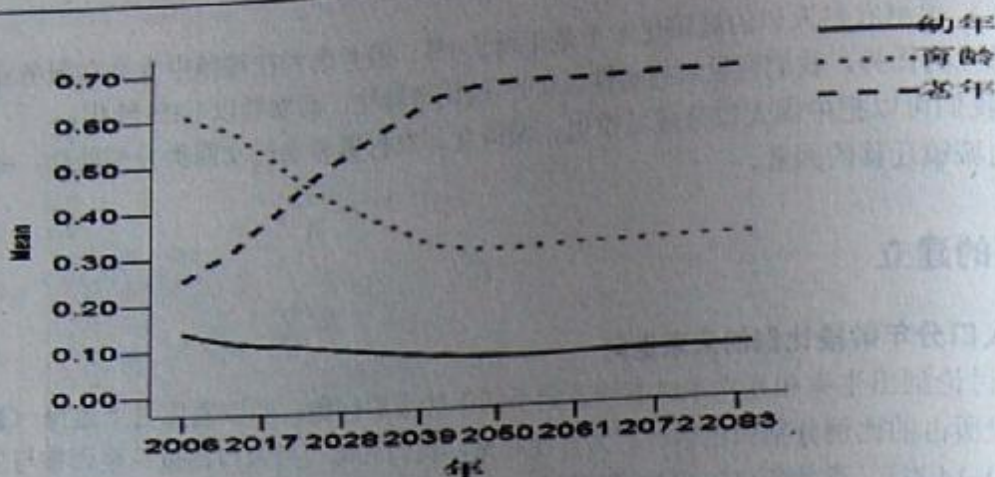


图5: 未来城镇妇女分年龄段走势

其中 α 为每年幼年向青壮年的转化率, β 是青壮年向老年转化率, b 是育龄妇女的女婴生育率, c 是老年的死亡率(幼年和青壮年的死亡率由于很低就忽略不计了)。

通有地说, 上述矩阵是可以对角化的, 即存在可逆矩阵 P , 使得

$$A = \begin{bmatrix} 1-\alpha & b & 0 \\ \alpha & 1-\beta & 0 \\ 0 & \beta & 1-c \end{bmatrix} = P^{-1} \begin{bmatrix} u & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & 1-c \end{bmatrix} P = P^{-1}BP.$$

这样, 我们就得到

$$(x_1, x_2, x_3)_n^T = (P^{-1}B^nP)(x_1, x_2, x_3)_0^T.$$

然后再对 x_n 进行归一化处理, 即命

$$y_n = (y_{1n}, y_{2n}, y_{3n})^T = x_n / \sum_{i=1}^3 x_{in},$$

得到

$$y_{in} = l_i + (1 - m_i(u/v)^n)k_i, \quad i = 1, 2; \quad y_{3n} = 1 - y_{1n} - y_{2n},$$

其中 $l_i, m_i, k_i \in \mathbf{R}$ 由 $x_n = (x_{1n}, x_{2n}, x_{3n})^T$ 确定。

由于只考虑短期 ($n < 60$), 为方便起见, 我们可用以常数 e 为底的指数函数进行拟合。这就得到下面的微分方程组

$$\begin{cases} d\omega_1(t)/dt = -a_1(\omega_1(t) - b_1), \\ d\omega_2(t)/dt = -c_1(\omega_2(t) - d_1), \\ \omega_3(t) = 1 - \omega_1(t) - \omega_2(t), \end{cases} \quad (1)$$

其中 $\omega_1(t)$, $\omega_2(t)$ 和 $\omega_3(t)$ 分别为幼年、育龄和老年城镇女性人口比例函数, a_1, b_1, c_1, d_1 为非负参数, 满足 $b_1 + d_1 < 1$ 。

对于该方程我们可以直接写出解 (r, s 为初值)

$$\begin{cases} \omega_1(t) = (r - b_1)e^{-a_1 t} + b_1, \\ \omega_2(t) = (s - d_1)e^{-c_1 t} + d_1, \\ \omega_3(t) = 1 - \omega_1(t) - \omega_2(t). \end{cases} \quad (2)$$

由于农村男女与城镇男有类似的结果(即其0-14岁和15-49岁两个年龄段的趋势函数都呈指数性下降), 在此不再赘述。所不同的是, 如前面假设中所指出的, 农村分别在1970年和1990年前后经历了两次生育高峰, 而指数拟合只能体现出未来分年龄段走势的主趋势。为了能更好地拟合未来农村人口分年龄段比例的走势, 还应当在指数拟合基础上添加一个不断减弱的周期性震荡。这是因为, 在出现某次生育高峰后还会周期性地出现新的生育高峰, 而文化素养的提高和晚婚晚育的政策会减削弱这种峰值效应。为简便起见, 我们添加一个不断减弱的正弦函数来描述这种周期性的变化。

对于其他三类人口的分年龄段的人口比例变化趋势的讨论类似, 我们只给出结果:

设 $\xi_1(t)$ 为幼年城镇男性比例函数, $\xi_2(t)$ 为壮年城镇男性比例函数, $\xi_3(t)$ 为老年城镇男性比例函数, 则有

$$\begin{cases} \xi_1(t) = (r - b_2)e^{-a_2 t} + b_2, \\ \xi_2(t) = (s - d_2)e^{-c_2 t} + d_2, \\ \xi_3(t) = 1 - \xi_1(t) - \xi_2(t), \end{cases} \quad (3)$$

其中 a_2, b_2, c_2, d_2 为非负参数, 而且满足 $b_2 + d_2 < 1$,

设 $\psi_1(t)$ 为幼年农村女性比例函数, $\psi_2(t)$ 为育龄农村女性比例函数, $\psi_3(t)$ 为老年农村女性比例函数, 则有

$$\begin{cases} \psi_1(t) = (r - b_3)e^{-a_3 t} + b_3 + g_1 e^{f_1 t} \sin \pi(t - \delta)/k_1, \\ \psi_2(t) = (s - d_3)e^{-c_3 t} + d_3 + g_2 e^{f_2 t} \sin \pi(t - \delta)/k_2, \\ \psi_3(t) = 1 - \psi_1(t) - \psi_2(t), \end{cases} \quad (4)$$

其中 $f_1, f_2, g_1, g_2, k_1, k_2, \delta$ 为普通参数, a_3, b_3, c_3, d_3 为非负参数, 而且满足 $b_3 + d_3 < 1$,

设 $\zeta_1(t)$ 为幼年农村男性比例函数, $\zeta_2(t)$ 为壮年农村男性比例函数, $\zeta_3(t)$ 为老年农村男性比例函数, 则有

$$\begin{cases} \zeta_1(t) = (r - b_4)e^{-a_4 t} + b_4 + g_3 e^{f_3 t} \sin \pi(t - \delta)/k_3, \\ \zeta_2(t) = (s - d_4)e^{-c_4 t} + d_4 + g_4 e^{f_4 t} \sin \pi(t - \delta)/k_4, \\ \zeta_3(t) = 1 - \zeta_1(t) - \zeta_2(t), \end{cases} \quad (5)$$

其中 $f_3, f_4, g_3, g_4, k_3, k_4, \delta$ 为普通参数, a_4, b_4, c_4, d_4 为非负参数, 而且满足 $b_4 + d_4 < 1$.

3.2 模型的最终确立

由于上面已确定了城镇和农村的男女各年龄段的比例函数, 现在我们可以按前面所作的基本假设, 即生育率是由育龄妇女的数量决定的、农村和城镇育龄妇女的生育率均不变、出生人口性别比均不变、各年龄段的人口的死亡也均不变, 以及农村人口按照一定的常数比例向城镇迁移这几个要点, 按照各类人口的出生率、死亡率和迁移率这三个方面来建立下面的常微分方程模型(6): 设

λ_1 -城镇育龄女性生育率, α_1 -城镇男性出生比例, χ_1 -农村男转化率;

λ_2 -农村育龄女性生育率, α_2 -农村男性出生比例, χ_1 -农村女转化率;

$\varepsilon_1, \varepsilon_2, \varepsilon_3$ -城镇幼年、育龄、老年女性的死亡率;

ϕ_1, ϕ_2, ϕ_3 -城镇幼年、壮年、老年男性的死亡率;

$\varphi_1, \varphi_2, \varphi_3$ -农村幼年、育龄、老年女性的死亡率;

$\gamma_1, \gamma_2, \gamma_3$ -农村幼年、壮年、老年男性的死亡率。

x -城镇男人口, y -城镇女人口, u -农村男人口, v -农村女人口

$$\begin{cases} dx/dt = \alpha_1 \lambda_1 \omega_2(t)y - (\phi_1 \xi_1(t) + \phi_2 \xi_2(t) + \phi_3 \xi_3(t))x + \chi_1 u, \\ dy/dt = (1 - \alpha_1) \lambda_1 \omega_2(t)y - (\varepsilon_1 \omega_1(t) + \varepsilon_2 \omega_2(t) + \varepsilon_3 \omega_3(t))y + \chi_2 v, \\ du/dt = \alpha_2 \lambda_2 \psi_2(t)v - (\gamma_1 \zeta_1(t) + \gamma_2 \zeta_2(t) + \gamma_3 \zeta_3(t))u - \chi_1 u, \\ dv/dt = (1 - \alpha_2) \lambda_2 \psi_2(t)v - (\lambda_1 \psi_1(t) + \lambda_2 \psi_2(t) + \lambda_3 \psi_3(t))u - \chi_2 v. \end{cases} \quad (6)$$

该模型表明, 各类人口的出生率取决于相应育龄妇女的数量; 而死亡率的变化是因为中国进入老龄化; 农村与城镇的男女出生性别比还在着较大的差别; 农村人口的城镇化正在稳步向前。这均反映出中国人口增长的各个特点。

4 未来人口预测

4.1 参数确定

我们可以确定上述参数, 这些参数数据主要是根据《中国人口统计年鉴》2001-2005年卷中数据整合得到的:

生育率: $\lambda_1 = 0.03026$, $\lambda_2 = 0.04356$,

男性出生比例: $\alpha_1 = 0.535$, $\alpha_2 = 0.547$,

转化率: $\chi_1 = 0.009$, $\chi_2 = 0.011$,

城镇女性死亡率: $\varepsilon_1 = 0.00073$, $\varepsilon_2 = 0.00068$, $\varepsilon_3 = 0.01045$,

城镇男性死亡率: $\phi_1 = 0.00075$, $\phi_2 = 0.00131$, $\phi_3 = 0.01187$,

农村女性死亡率: $\varphi_1 = 0.00217$, $\varphi_2 = 0.00146$, $\varphi_3 = 0.01367$,

农村男性死亡率: $\gamma_1 = 0.00207$, $\gamma_2 = 0.00246$, $\gamma_3 = 0.01542$,

城镇女性走势相关: $a_1 = 0.05$, $b_1 = 0.067099$, $c_1 = 0.05$, $d_1 = 0.294517$,

城镇男性走势相关: $a_2 = 0.05$, $b_2 = 0.073994$, $c_2 = 0.045$, $d_2 = 0.321786$,

农村女性走势相关: $a_3 = 0.065$, $b_3 = 0.11221$, $c_3 = 0.045$, $d_3 = 0.380758$, $g_1 = 0.06$, $g_2 = -0.02$, $f_1 = -0.03$, $f_2 = -0.045$, $k_1 = 14$, $k_2 = 14$, $\delta = 4$,

农村男性走势相关: $a_4 = 0.065$, $b_4 = 0.13192$, $c_4 = 0.045$, $d_4 = 0.432448$, $g_3 = -0.0076$, $g_4 = 0.05$, $f_3 = -0.01$, $f_4 = -0.035$, $k_3 = 13$, $k_4 = 13$, $\delta = 4$.

4.2 方程离散化求解

现在我们利用得到的人口模型来对人口的增长进行预测。首先是确定初值。对于《中国人口统计年鉴》2005年版中所提供的城镇与农村的抽样调查数据进行改良, 我们可以得到初值

$$x(0) = 287740000, \quad y(0) = 288259000, \quad u(0) = 399428000, \quad v(0) = 357638000.$$

方程(6)是一个非自治的耦合的常微分方程组, 可以用数值方法求得其初值问题的解。我们取一年作为步长, 采用简单的 Euler 法, 因为人口数量的第 n 年增长可以由第 $n-1$ 年的数据经过迭代得到。利用上述初值, 通过 Euler 法, 我们可以得到下面的算图。

图6是从2005年开始, 未来60年的人口走势预测, 其中以“...”表示农村男性人口, 以“-”表示农村女性人口, 以“-”表示城镇男性人口, 而以“-”表示城镇女性人口。

从图6我们不难看出, 在未来的60年内(在现有人口政策和经济发展均不变的情况下)城镇人口会进一步增长, 农村人口会出现下降, 而且农村人口的男女比例会进一步加大。

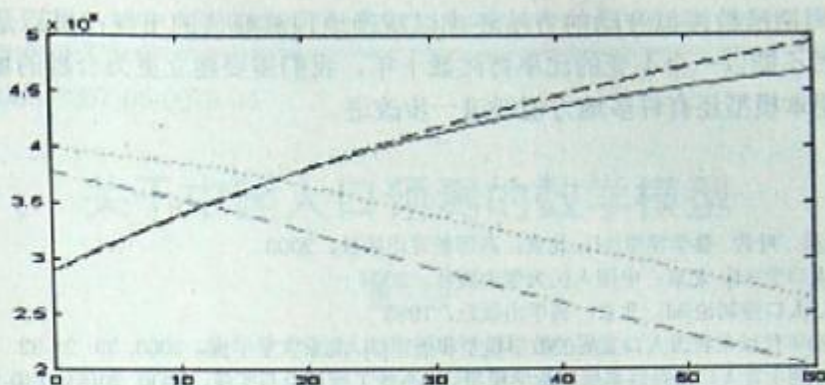


图6: 未来60年人口走势预测

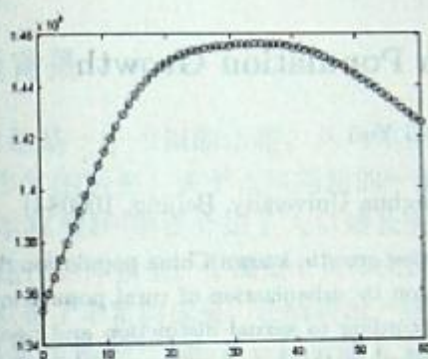


图7: 未来60年人口总量预测

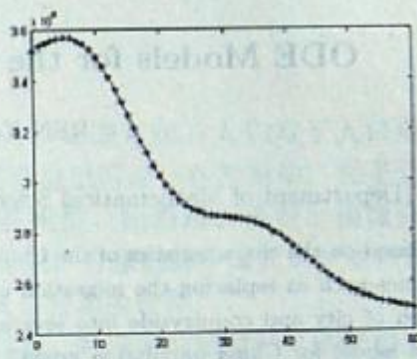


图8: 未来60年育龄妇女总量预测

图7的结果显示, 约30年以后中国人口的总数会达到顶峰, 总人口数接近15亿。而这与《国家人口发展战略研究报告》中的结论有些出入, 但是曲线的走势和该报告中所提供的结果还是比较相近的。

图8为对未来60年育龄妇女人口数的预测。特别从前40年的预测结果可见: 育龄妇女会从2005年的接近3.8亿人减少为约2.7亿人, 其间会出现两个峰值, 间隔约为30年。这个结果与《国家人口发展战略研究报告》的对未来40年育龄妇女的预测结果基本相同。

5 总结

本文在对人口增长的特性进行分析的基础上, 利用多维方法解决中国人口增长的城乡差异和男女差异问题; 利用育龄妇女比例的走势解决人口出生率的变化问题; 又利用将人口按年龄段分类的方法解决人口总死亡率的变化问题; 最后利用农村人口城镇化来代替人口的迁移, 从而得到了最终的常微初值问题数学模型。模型在内容上包含了中国人口增长的各个特征, 从而得到了最终的常微初值问题数学模型。模型在内容上包含了中国人口增长的各个特征, 从而得到了最终的常微初值问题数学模型。但是不足之处依然存在: 首先, 预测结果与《国家人口发展战略研究报告》中的比较一致。但是不足之处依然存在: 首先, 本文中的模型在处理生育率和死亡率的变化时, 由于使用的是年龄段的平均数值, 仍不能很好地模拟这两方面的实际变化; 其次, 在处理未来人口分年龄段比例走势时主要采用指数函数拟合, 尽管在处理农村人口变化时又辅以周期函数, 但由于人口增长的惯性十分复杂, 仅使

用指数函数和周期函数模拟波动的方法还难以精确地预测峰值的出现;最后是对城镇化的处理,城镇化显然不能以一个不变的比率持续数十年,我们需要建立更为合理的城镇化进程预测函数。这些说明本模型还有许多地方值得进一步改进。

参考文献:

- [1] 姜启源, 谢金星, 叶俊. 数学模型[M]. 北京: 高等教育出版社, 2003
- [2] 查瑞传. 数理人口学[M]. 北京: 中国人民大学出版社, 2004
- [3] 宋健, 于景元. 人口控制论[M]. 北京: 科学出版社, 1985
- [4] 王彦. 20世纪80年代以来我国人口发展的数学模型和展望[J]. 北京大学学报, 2003, 39: 28-32
- [5] 霍江明. 实行计划生育人口复杂巨系统的数学模型[J]. 系统工程理论与实践, 2000, 20(8): 140-143
- [6] 国家人口发展战略研究课题组. 国家人口发展战略研究报告[J]. 人口研究, 2007, 31(1): 1-10
- [7] 国家统计局人口统计就业司. 中国人口统计年鉴[M]. 北京: 中国统计出版社, 1995-2006
- [8] 国务院第五次全国人口普查办公室. 世纪之交的中国人口(全国卷)[M]. 北京: 中国统计出版社, 2006

ODE Models for the China Population Growth

REN Ke, TANG Yun

(Department of Mathematical Sciences, Tsinghua University, Beijing, 100084)

Abstract: Based on the characteristics of the China population growth, known China population data and by using techniques such as replacing the migration of population by urbanization of rural population, dividing the population of city and countryside into several parts according to sexual distinction and ages, an ODE mathematical model for China population growth is established in this paper. The model is then used for predicting the future China population growth.

Keywords: birth rate; death rate; urbanization