

文章编号:1005-3085(2007)08-0031-08

中国人口增长预测模型

郭 媛, 冯俊峰, 杨智超

指导教师: 贺祖国

(北京邮电大学, 北京 100876)

编者按: 该文从不考虑性别、城乡差异的 Leslie 模型入手, 逐步引入性别、城镇乡及城市化几个影响人口增长的重要因素。建模思路从简到繁, 值得借鉴。该文的另一个特点是对预测结果的多方面综合利用加工, 得到许多有用的派生结果, 为决策提供了许多重要信息。

摘 要: 本文通过对我国人口基本国情的研究, 从人口发展方程的离散形式(模型一)出发, 提出基于性别的人口发展方程(模型二), 将男性和女性人口当作两个独立的发展系统。然后, 我们进一步改进模型二并提出模型三—基于城镇化的人口修正预测, 把城市、乡镇和农村作为三个系统, 分别应用模型二, 同时考虑这三个系统间的人口迁移, 即由农村向城市和乡镇迁移, 使城镇化比率符合当前社会规律。最后, 以模型三为基础, 我们分短期和长期研究了中国人口的未来发展走势, 并对预测结果进行探讨得出了一些有意义的结论。

关键词: 人口发展模型; 总和生育率; 老龄化; 性别比; 城镇化

分类号: AMS(2000) 34B60

中图分类号: O241.81

文献标识码: A

1 问题分析

目前中国人口最突出的特点是: 老龄化加速, 出生人口性别比持续增高和乡村人口城镇化。要讨论这些问题, 在人口预测模型中必须将对象的年龄, 性别, 和所在地加以区分。因此可以选用转移矩阵来分析人口变化趋势, 将人口按性别、年龄、所在地分组, 通过转移矩阵迭代出下一年各组的人口数。这种模型能反映各年龄层、性别和地域的区别, 并方便计算各种反映人口结构特点的指标。

按预测期长短, 人口预测可分为短期预测(5年以下), 中期预测(5~20年)和长期预测(20~50年)。中短期预测的参数可以实际调查数据为基础, 根据以往趋势加以估计, 现实意义较大。长期预测中未知因素较多, 不易把握, 受政策的影响大。我们可以选定总和生育率、出生性别比和死亡率作为控制参数, 对这些参数做出假设。基于不同的假设, 我们可以得到不同的预测结果, 进行比较。

2 模型假设

1. 在预测时间内, 不发生大的疫情, 灾难或战争等引起人口重大变化的事件;
2. 中短期内, 总和生育率、死亡率和出生性别比不发生大的波动, 以往年平均值代替预测值;
3. 长期人口预测的参数主要由政策决定;
4. 90岁以后各年龄的人口数较少, 死亡率具有相同的特征, 作为一组数据进行处理;
5. 生育模式在预测时间内保持不变, 并且假设一胎只生一个;
6. 流动人口按城市和乡镇目前的人口总量的比例分别分配;
7. 流入流出人口不改变该地区的人口性别、年龄结构。

3 符号说明

m	最大年龄, 在此视为 90	$f_w(t)$	第 t 年出生的女婴
$x(i, t)$	第 t 年 i 岁的人数	$f_m(t)$	第 t 年出生的男婴
$k(i, t)$	第 t 年 i 岁的人口女性比	$\varphi(t)$	第 t 年的人口出生性别比
$x_m(i, t)$	第 t 年 i 岁的男性人数	$\beta(t)$	第 t 年每个育龄妇女平均生育的婴儿数, 即总和生育率
$x_w(i, t)$	第 t 年 i 岁的女性人数	$h(i, t)$	第 t 年年龄为 i 周岁的女性生育模式, 用以调整育龄妇女不同年龄生育率的高低
$d(i, t)$	第 t 年 i 岁人口的死亡率	$c(t)$	第 t 年的城镇化比率
$d_m(i, t)$	第 t 年 i 岁男性的死亡率	$move(t)$	第 t 年的流动人口数
$d_w(i, t)$	第 t 年 i 岁女性的死亡率		其中 $t = 0, 1, 2, \dots, i = 0, 1, \dots, m$
$b(i, t)$	第 t 年 i 岁女性的生育率		
$f(t)$	第 t 年出生人口数		

4 模型建立

4.1 模型一: 人口发展方程的离散形式即 Leslie 模型, 参见文[1]

4.2 模型二: 基于性别的人口发展方程

模型一假定男女死亡率相等, 且未考虑新生儿的性别比例, 不利于预测人口性别比例。本模型改进了 Leslie 模型, 男性和女性的人口数作为两个独立的系统进行发展, 总人口趋势由两者共同决定。以下标 w 表示女性人口, 下标 m 表示男性人口。则女性人口发展的差分方程为

$$x_w(i+1, t+1) = [1 - d_w(i, t)]x_w(i, t), \quad i = 0, 1, \dots, m-1, t = 0, 1, 2, \dots \quad (1)$$

第 t 年出生并存活下来的女婴数为

$$x_w(0, t) = \frac{100}{\varphi(t) + 100} [1 - d_{w0}(t)] \beta(t) \sum_{i=i_1}^{i_2} h(i, t) x_w(i, t), \quad (2)$$

其中 $\varphi(t)$ 为第 t 年的出生性别比, $d_{w0}(t)$ 为第 t 年出生但未活到人口统计时刻女婴比例。

记

$$b'_w(i, t) = \frac{100}{\varphi(t) + 100} [1 - d_{w0}(t)] [1 - d_w(0, t)] h(i, t),$$

女性人口数量向量为 $X_w(t)$ 、出生率矩阵为 $B_w(t)$ 和存活率矩阵 $A_w(t)$ 与 Leslie 模型中的定义同形式, 上述关系表述为

$$X_w(t+1) = A_w(t)X_w(t) + \beta(t)B_w(t)X_w(t), \quad (3)$$

同理可推导男性人口发展方程。总人口发展过程模型为

$$X(t+1) = X_w(t+1) + X_m(t+1). \quad (4)$$

4.3 模型三: 基于城镇化的人口修正预测

模型二将男性和女性的人口数当作两个独立的系统进行研究, 由此容易计算出不同性别、年龄的人口分布。但是模型二与模型一都基于一个重要假设: 仅考虑出生率死亡率的影响, 不计迁移等社会因素的影响, 因而只适用于一个孤立的人口系统。

模型三以模型二为基础, 把城市、乡镇和农村作为三个系统, 分别应用模型二; 同时考虑这三个系统间的人口迁移, 即由农村向城、镇迁移, 并保证城镇化比率符合当前社会规律。

记城市、乡镇和农村第 t 年 i 岁的人口数分别为 $xc(i, t)$, $xz(i, t)$, $xx(i, t)$, 其他变量、矩阵都作类似定义, 分别应用模型二。例如城市人口发展方程以矩阵形式表示为

$$\begin{aligned} XC(t+1) &= XC_w(t+1) + XC_m(t+1) \\ &= AC_w(t)XC_w(t) + \beta c(t)BC_w(t)XC_w(t) + AC_m(t)XC_m(t) + \beta c(t)BC_m(t)XC_m(t), \end{aligned} \quad (5)$$

定义城镇化比率

$$\text{城镇化} = \frac{\text{城镇人口}}{\text{全国人口}} = \frac{\text{城市人口} + \text{乡镇人口}}{\text{全国人口}},$$

用数学符号表示为

$$c(t) = \frac{\sum_{i=0}^{90} [xc(i, t) + xz(i, t)]}{\sum_{i=0}^{90} [xc(i, t) + xz(i, t) + xx(i, t)]}, \quad (6)$$

$c(t)$ 表示城镇化比率随时间变化的函数, 是由目前实际情况和今后政策所决定的。

由上式可推算出第 t 到第 $t+1$ 年间走出农村的流动人口数为

$$\text{move}(t) = \sum_{i=0}^{90} xx(i, t+1) - [1 - c(t+1)] \sum_{i=0}^{90} [xc(i, t+1) + xz(i, t+1) + xx(i, t+1)], \quad (7)$$

按照假设6, 流动人口按照城市和乡镇目前的人口总量的比例分配给它们, 即流动人口并不改变城市与乡镇之间的总人口比例。各地区人口总量修正为

$$\begin{cases} \sum_{i=0}^{90} xc'(i, t+1) = \sum_{i=0}^{90} xc(i, t+1) + \text{move}(t) \frac{\sum_{i=0}^{90} xc(i, t+1)}{\sum_{i=0}^{90} [xc(i, t+1) + xz(i, t+1)]}, \\ \sum_{i=0}^{90} xz'(i, t+1) = \sum_{i=0}^{90} xz(i, t+1) + \text{move}(t) \frac{\sum_{i=0}^{90} xz(i, t+1)}{\sum_{i=0}^{90} [xc(i, t+1) + xz(i, t+1)]}, \\ \sum_{i=0}^{90} xx'(i, t+1) = \sum_{i=0}^{90} xx(i, t+1) - \text{move}(t). \end{cases} \quad (8)$$

再按照假设7, 流入流出人口不改变该地区人口的性别比例、年龄结构, 则修正后的人口分布与修正前相同, 只是该地区的人口总量变化了。求解时, 先确定函数 $c(t)$, 逐年迭代计算出人口数量及分布, 再进行流动人口修正, 然后将修正后的结果继续迭代。

5 模型求解

以上三个模型中, 模型三考虑最全面, 易于根据模型分析人口趋势和指标, 并且同时适用与中短期预测与长期预测。因此以下均按照模型三进行求解。

5.1 中短期预测与分析

5.1.1 模型参数计算

> 出生率

按照

$$\beta(t) = \sum_{i=1}^{i_2} b(i, t),$$

计算01~05各年总和生育率,列表如下(由于SARS的影响,03年数据出现异常,故求平均值是删除了03年后的数据):

表1: 历年总和生育率

	2001	2002	2003	2004	2005	平均值(删除03)
城市	1.002	0.960	0.095	1.048	0.926	0.984
乡镇	1.189	1.203	0.132	1.347	1.278	1.254
农村	1.604	1.653	0.168	1.687	1.654	1.649

按各地区育龄妇女总数对 β 值加权平均,可以得到全国妇女总和生育率为1.375。根据国家计划,30年内应该把总和生育率控制在1.8以内才能将人口峰值控制在15亿。按照现在的总和生育率,并不需要再降低总和生育率。我们推测存在隐瞒黑户不报等状况,导致普查得到总和生育率低于实际值,中短期预测仍采用1.8作为全国妇女总和生育率。

由 $h(i, t) = \frac{b(i, t)}{\beta(t)}$ 可计算出生育模式 $h(i, t)$,以城市妇女为例,作图如下:

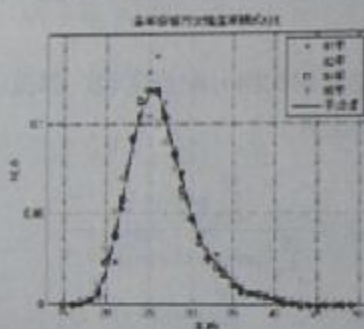


图1: 各年份城市女性生育模式图

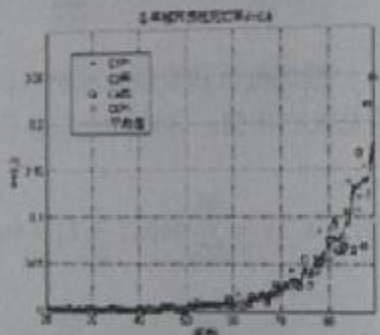


图2: 各年份城市男性死亡率

由上图可见生育模式很稳定,随年份波动很小。因此中短期预测中可以取其平均值作为生育模式。

> 死亡率

各地区人口死亡率随年龄变化的趋势是一致的,随着年龄的增大死亡率明显升高,年与年之间波动明显。资料显示随着医疗水平的提高,死亡率将按年份呈递减趋势。但是整体趋势变化非常缓慢,因此中短期预测取平均值代替。

> 城镇化比率

下表列出了1995年至2005年间的城镇化比例(数据来自国家统计局人口统计年鉴):

表2: 城镇化比率(%)

年份	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
城镇化	29.4	30.48	31.91	33.35	34.78	36.22	37.66	39.09	40.53	41.76	42.99

拟合结果为

$$c(t) = 1.41t - 2783.9, \quad (9)$$

可以看出城镇化比率随年份的发展呈现很强的线性趋势, 残差水平为0.32, 中短期内可采用线形模型预测城镇化比率。当然, 长期预测使用线性模型不可行, 否则农村人口将趋于0。长期预测中城镇化比率与政策相关, 在以后的预测中将予以重新讨论。

5.1.2 人口增长趋势与城镇化的影响

下图为以2005年为起始年, 根据模型得到的全国总人口和城镇乡各地人口趋势图:

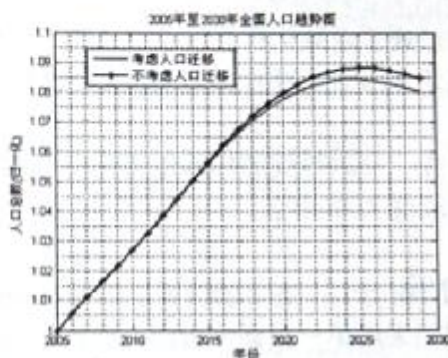


图3: 全国人口预测

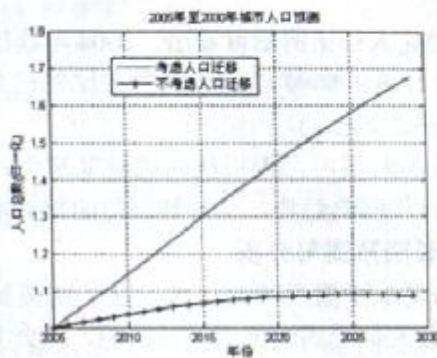


图4: 城市人口预测

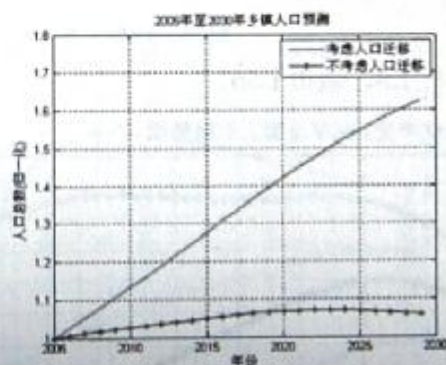


图5: 乡镇人口预测

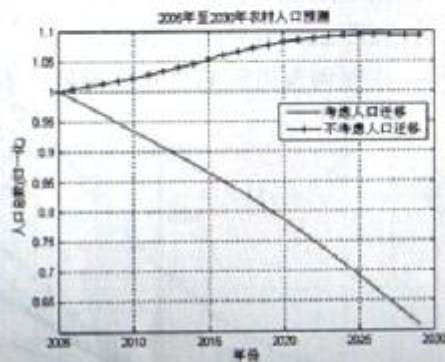


图6: 农村人口预测

按此预测, 总人口在2010年, 2020年分别达到13.43亿人和14.12亿人。在2025年人口达到最大值, 为14.23亿人。与国家人口发展战略研究课题组预测的数据相比, 相同年份人口数偏小, 人口最高峰出现年份提前。另外, 市、镇人口数均呈现增加的趋势, 而乡的人口数则随时间推移逐渐下降。市、镇人口数增加一方面来自于原住人口的自然增长, 另一方面来自于城镇化导致人口数量的增加。农村人口减少则主要是由迁移造成的。

5.1.3 人口老龄化分析

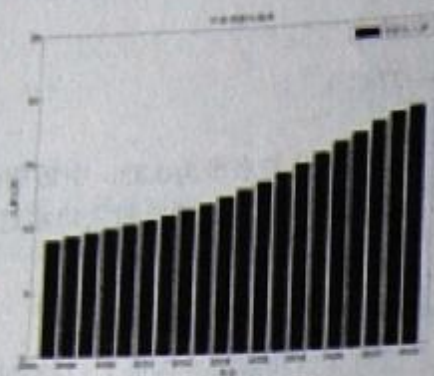


图7: 老龄人口比例随时间变化图

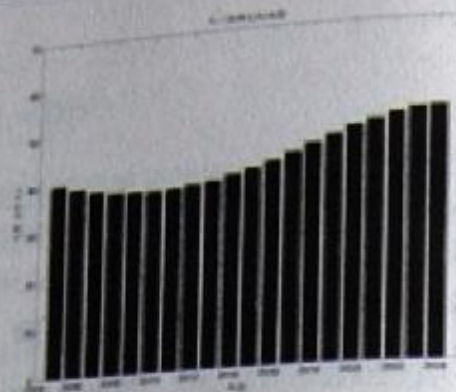


图8: 人口抚养比随时间变化图

由老龄化人口比例图可看出，2004年我国已处于老龄化社会阶段，且这一趋势还在进一步加强，到2020年65岁以上老人所占比例已高达15.1%，超出老年性社会标准的一倍。老年人口总数已接近2亿。

在未来的时期内，人口抚养比也呈增大趋势。预测得到的抚养比变化图与赛题附件1中给出的结果有相同的趋势，且预测所得的抚养率取值大致相同。

5.2 长期预测与分析

与人口增长中短期预测相比，在长期预测的模型中，更多的因素会影响预测的结果，且长期预测参数大都受到政策调控的影响，因此我们采用单一因素分析法，即在保持其它参数不变时，对某一关心的参数进行研究。并在最后综合各个因素，对人口进行长期的预测。

5.2.1 考虑总和生育率对人口增长的影响

对于总和生育率的取值，我们列出如下几种特殊情况予以讨论：

模式一：保持附件2中的近年总和生育率1.375；

模式二：按照附件1中建议值1.8；

模式三：按照生育更替水平值2.1；

模式四：按照调查得到的生育意愿值：农村1.8，城市1.56。

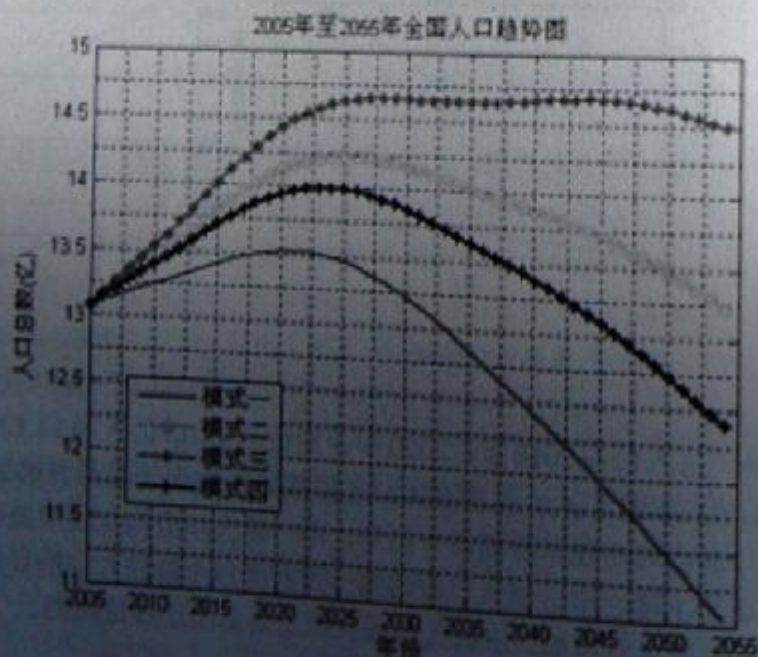


图9: 不同 β 值对全国总人口的影响

四个模式的人口发展趋势表明由于我国计划生育实行效果显著, 在未来50年内人口都将显著减少。模式二是国家提倡的总和生育率。在经过缓慢上升后开始缓慢下降, 并表现出一种稳定在某值的态势。这一趋势最有利于人口的平稳过渡, 既平衡了劳动力突然下降所造成的不良后果, 又不会再产生太大的人口总数。模式三是一个生育更替稳定值, 体现出本模型的合理性。

5.2.2 考虑城镇化对人口增长的影响

很明显我国城镇化进程不可能按照线性增长长期发展下去。据文献, 到2020年, 中国城镇化率将达到最大值, 中国城镇化将逐渐变缓, 最终进入城镇化比率为0.7~0.8的平稳区。鉴于上述特点我们采用 Logistic 曲线进行数据拟合, 得到如下城镇化趋势的预测式

$$c(t) = \frac{0.75}{1 + 1.0443e^{-0.0501t}}, \quad (10)$$

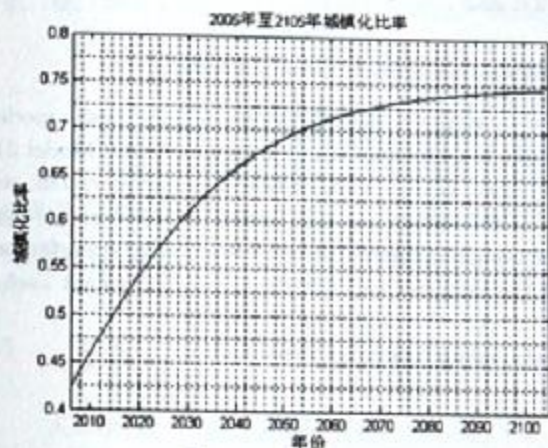


图10: 城镇化比率

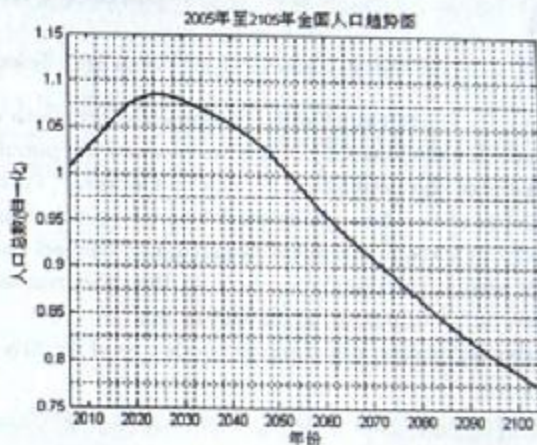


图11: 全国人口预测

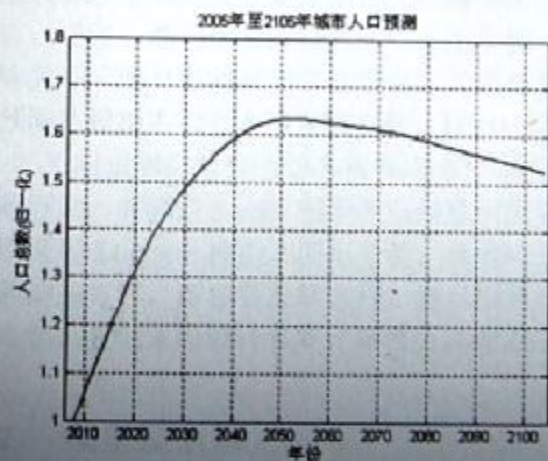


图12: 城市人口预测

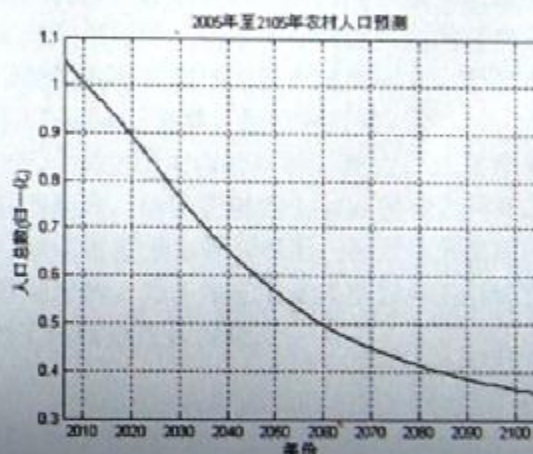


图13: 农村人口预测

从上图我们可以看到城镇化率不断提高到2100年左右时接近0.75, 该数据是根据国家发展规划得到的。城镇人口则先增后减, 体现出两个阶段的人口发展: 第一阶段是2050年之前, 城镇化率的影响占据主要地位, 大量的农村人口涌入到城市或由于许多农村发展而变成城镇人口。因此城市人口不断增加, 乡镇人口净增长也成上升趋势, 而农村人口则不断下降。第二阶段是2050年后一方面全国总人口数不断下降, 城镇人口发展增速也不断减小, 因此均呈现出下降趋势。

参考文献:

- [1] 姜启源, 谢金星, 叶俊. 数学模型(第三版)[M]. 北京: 高等教育出版社, 2003
- [2] 国家统计局人口和就业统计司. 中国人口统计年鉴—2006[M]. 北京: 中国统计出版社, 2006
- [3] 博客中国, 以人为本, 用科学的发展观看待人口问题[OL]
<http://beijingdaxue.bokee.com/668260.html>, 2005-02-03
- [4] 人口发展“十一五”和2020年规划简介[OL]
<http://www.hebrkjs.gov.cn/xinxi/news/2007124/2007124170309.htm>, 2007-01
- [5] 孙祥等. matlab 7.0基础教程[M]. 北京: 清华大学出版社, 2005

Forecasting Models about the Chinese Population Growth

GUO Yuan, FENG Jun-feng, YANG Zhi-chao
Advisor: HE Zu-guo

(Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract: By investigating some fundamental problems of our country's population, we employ Leslie model as our basic model (Model 1). Then we propose the population development equation based on sex (Model 2), which separates the population of men and women. Furthermore, taking urbanization into consideration, we achieve the modified forecasting model based on urbanization (Model 3). In this model, we treat city, town and country as three independent systems, and connect these systems together through the immigration factor. Lastly, we predict the population both in the short run and the long run and analyze the results to get useful conclusions on urbanization, aging and so on.

Keywords: population development model; total fertility rate; aging; sex ratio; urbanization