

Part I

目录

Part II

摘要

Part III

假设

Part IV

人口增长的简单模型

1 引述

就是抄一抄百科

2 模型概述

我们从人口的产生与消亡这两方面来看待人口的变动。人口的增长，一方面被人口的产生所促进，另一方面被人口的消亡所限制，所以是一个动态平衡的过程。一大波新出生的人口，在达到适育年龄的时候，会反过来对人口的增长起到促进作用；而当他们走向衰老乃至死亡的时候，又会加剧人口的减少。因此我们要使用科学的方法来看待这个问题。下面我们分别着手处理这个两个问题。

2.1 人口的衰老

随着时间的流逝，一部分人口会衰老一岁，而剩下不幸的人则会因为各种各样的原因死去。其中死亡率是一个随着年龄而变化的数据，一般情况是婴儿的死亡率较高，然后随着年龄的增长而下降，到 40、50 岁时开始逐步回升。

2.2 人口的出生

各个年龄层次的适育妇女，都有可能在考察的时间段内生育：对于个人而言，这是一个概率性事件；但是对于群体而言，生育事件的随机性就被大量的人群基数所磨灭，成为一个比率。因此，一段时间内婴儿的出生数目，等于平均生育率乘上适育妇女的人数。

但是，不同地区、学历的妇女在生产意愿上有着不同的态度。为了简化问题，我们认为地区与教育带给人的影响是近似独立的，因此可以直接对平均生育率进行修正。在这种假设下，一段时间内婴儿的出生人数，等于各年龄层的适育妇女乘上修正后的生育率之和。

3 记号表

见 Table1。

Table 1: 人口增长的简单模型			
符号	含义	相关的其他参数	备注
a	年龄		
e	受教育水平（枚举量）		e_0 : 小学及以下 e_1 : 中学 e_2 : 大学及以上
r	地区（枚举量）		r_0 : 城 r_1 : 镇 r_2 : 乡
t	考察的时间段		
I	出生婴儿数	e, r, t	
N	总人口数	a, e, r, t	
W	女性人数	a, e, r, t	
β	生育率	a, e, r	
δ	平均死亡率	a	
λ	修正系数	a, e, r	
μ_p	男女比例		

4 数据引用

Table 2: 来自第六次人口普查的数据

$I(r_k, e_l, 2010)$	e_0	e_1	e_2
r_0			
r_1			
r_2			

5 方程的建立

首先，考虑人口的自然衰老：

$$N(a+1, t+1) = N(a, t) \cdot (1 - \delta(a)) \quad (1)$$

这里 $\delta(a)$ 指的是 a 岁人口的平均死亡率。

再考虑所有地区，由受各种教育水平的妇女生育的婴儿数：

$$N(0, t+1) = \sum_{r_k} \sum_{e_l} I(r_k, e_l, t) \quad (2)$$

这里 $I(r_k, e_l, t+1)$ 表示在 t 这个考察时间段内，在 r_k 地区，受到 e_l 教育水平的妇女生育的婴儿数目。

考察 $I(r_k, e_l, t)$ 是由各个不同年龄阶段的适龄妇女所生育的：

$$I(r_k, e_l, t) = \sum_a \beta(a, r_k, e_l) W(a, r_k, e_l, t) \quad (3)$$

这里 $\beta(a, r_k, e_l)$ 是 r_k 地区，受到 e_l 教育水平的妇女的生育率， $W(a, r_k, e_l, t)$ 是这段时间这类妇女的总人数。

总人口数与总女性数之间存在着简单的比例关系：

$$N(a, r, e, t) = W(a, r, e, t) \cdot (1 + \mu_p) \quad (4)$$

这里 μ_p 表示的是男性与女性人数的比例。

教育水平与地区对于生育率的修正，即 $\beta(a, r_k, e_l)$ ，是这样构成的：

$$\beta(a, r_k, e_l) = \overline{\beta(a)} \cdot \lambda_{r_k} \cdot \lambda_{e_l} \quad (5)$$

其中 $\overline{\beta(a)}$ 是平均生育率, λ_{r_k} 与 λ_{e_l} 分别是地区与教育水平分别对于生育率的修正; 为了记号上的方便, 令 λ_b 为生育修正参数矩阵, 其中

$$\lambda_b(r_k, e_l) = \lambda_{r_k} \cdot \lambda_{e_l} \quad (6)$$

在进一步完善模型之前, 我们给出 $\lambda_b(r_k, e_l)$ 的计算公式与结果。事实上, 根据等式 3 与等式 5, 我们可以立即得到:

$$\lambda_b(r_k, e_l) = \frac{I(r_k, e_l, t)}{\sum_a \overline{\beta(a)} W(a, r_k, e_l, t)}. \quad (7)$$

利用表 2 中的数据, 代入 $t = 2010$, 我们得到了 $\lambda_b(r_k, e_l)$ 的具体结果, 见表 3:

Table 3: 生育修正参数 λ_b

$\lambda_b(r_k, e_l)$	e_0	e_1	e_2
r_0			
r_1			
r_2			

6 小结

把这个 naive 的模型批判一番

Part V

人口增长模型修正与应用

7 假设

明确地列出哪些因素是要考虑的, 而哪些还是不考虑

8 记号表

Table 4: 人口增长简单模型的修正

符号	含义	相关的其他参数
I_i	作为第 $(i + 1)$ 胎出生的婴儿数目	e, r, t
W_i	生育了 i 个孩子的妇女数目	a, e, r, t
β_i	已生育 i 胎妇女的生育率	a, e, r
η_a	a 岁已育一胎妇女的相对生育意愿	
λ_i	已生育 i 胎妇女受到政策阻力修正因素	
ν_{r_k}	从 r_{k+1} 到 r_k 的地区迁移率	
ν_{e_l}	从 e_{l+1} 到 e_l 的受教育水平迁移率	

9 数据引用

- 人口增长率
- 城乡比
- 姿势水平的人数

剩下的就直接甩链接。

10 模型的修正

10.1 依据生育情况区分妇女

首先，适育妇女数，依照她们已生育孩子的数目，可分为 W_0 、 W_1 、 W_2 这三类，分别表示未生孩子、已生一个孩子与生了两个及以上的妇女数目。在这样的分类下，等式 3 将被修正为如下带有下标的形式：

$$I_i(r_k, e_l, t) = \sum_a \beta_i(a, r_k, e_l) W_i(a, r_k, e_l, t) \quad (8)$$

这里 I_i 是作为第 $(i + 1)$ 胎出生的婴儿数目，而 W_i 是生育过 i 胎的妇女数目。

同时，由于计划生育政策，已生育的妇女将会受到较大的政策阻力，因此她们的生育率将会减小。这一点反应在模型上是通过参数对方程进行调节的：

$$\beta_i(a, r_k, e_l) = \overline{\beta(a)} \cdot \lambda_i \cdot \lambda_b(k, l) \quad (9)$$

其中 λ_i 是已生 i 胎妇女受到政策阻力的修正因数。

考虑在放开二胎政策之前，对于已生育的妇女而言，再生育的政策阻力应当是一致的。为了得到政策阻力的量化情况，我们采用数据拟合的手段。考虑 λ_i 具有如下的形式：

$$\lambda_i = \begin{cases} (3 - \lambda_{policy}) & i = 0 \\ \lambda_{policy} & i = 1 \\ \lambda_{policy} & i = 2 \end{cases} \quad (10)$$

这里参数 λ_{policy} 是由再分法确定的，使得人口增长率符合当年数据。拟合结果如表 5 所示。

Table 5: λ_i 拟合结果

参数	λ_0	λ_1	λ_2
值			

在第一次修正后，中国人口增长趋势大致如图所示。
图图图

10.2 考虑人口的地域与教育水平的迁移

随着时间的推移，人口逐渐向城市靠拢，同时由于自身的不断学习，受教育水平也在不断提升。这两个方面的迁移由以下的方程组来描述：

$$\begin{cases} N'(a, r_0, e_l, t+1) = N(a, r_0, e_l, t+1) + \nu_{r_0} \cdot N(a, r_1, e_l, t+1) \\ N'(a, r_1, e_l, t+1) = (1 - \nu_{r_0})N(a, r_1, e_l, t+1) + \nu_{r_1} \cdot N(a, r_2, e_l, t+1) \\ N'(a, r_2, e_l, t+1) = (1 - \nu_{r_1})N(a, r_2, e_l, t+1) \\ N''(a, r_k, e_0, t+1) = (1 - \nu_{e_0})N'(a, r_k, e_0, t+1) \\ N''(a, r_k, e_1, t+1) = \nu_{e_0} \cdot N'(a, r_k, e_0, t+1) + (1 - \nu_{e_1})N'(a, r_k, e_1, t+1) \\ N''(a, r_k, e_2, t+1) = \nu_{e_1} \cdot N'(a, r_k, e_1, t+1) + N'(a, r_k, e_2, t+1) \end{cases} \quad (11)$$

这里 $N''(a, r_k, e_l, t+1)$ 是经过两次修正后的不同地区、受教育程度的人口数。

为了确定参数 ν_{r_0} 与 ν_{r_1} 、 ν_{e_0} 与 ν_{e_1} ，从 2010 年的数据出发进行迭代至 2015 年，选取预测结果最为接近的一组参数；其中最为接近是指比例相

等，由如下方程组描述：

$$\begin{cases} \frac{N_P(r_0, t+5)}{N(r_0, t+5)} = \frac{N_P(r_1, t+5)}{N(r_1, t+5)} = \frac{N_P(r_2, t+5)}{N(r_2, t+5)} \\ \frac{N_P(e_0, t+5)}{N(e_0, t+5)} = \frac{N_P(e_1, t+5)}{N(e_1, t+5)} = \frac{N_P(e_2, t+5)}{N(e_2, t+5)} \end{cases} \quad (12)$$

这里 $N_P(t+5)$ 是指从 t 出发迭代预测的数据，而 $N_P(r_k, t+5)$ 应理解为 $\sum_{e_l} \sum_a N_P(a, r_k, e_l, t+5)$ ， $N_P(e_l, t+5)$ 应理解为 $\sum_{r_k} \sum_a N_P(a, r_k, e_l, t+5)$ 。拟合结果如表 6 所示。

Table 6: ν_{r_k} 与 ν_{e_l} 拟合结果

参数	ν_{r_0}	ν_{r_1}	ν_{e_0}	ν_{e_1}
值				

在第二次修正后，中国人口增长趋势大致如图所示。

10.3 考虑计划生育政策的开放带来的影响

考虑到当二胎政策出台后，仅已生一胎的妇女生育意愿会增大，因此等式 5 中 $\beta_1(a, r_k, e_l)$ 将会比原来大，被修正为：

$$\beta(a, r_k, e_l) = \eta_a \cdot \overline{\beta(a)} \cdot \lambda_i \cdot \lambda_b(k, l) \quad (13)$$

其中 η_a 是 a 岁已生育一胎的妇女在政策放开后的相对生育意愿。

11 小结

再批判一番

Part VI

把其他人的论文批判一番

12 你们啊，naive

蛤

13 西方的人口模型我哪个没算过

蛤

14 开放二胎不知道比你们高到哪里去了

蛤

15 你问我支不支持计划生育，那我当然是支持的

蛤

16 我今天算是得罪了你们

蛤

17 小结

蛤

Part VII

北京市的人口增长模型

18 讨论

考虑到北京市城市人口占总人口大多数【来源请求】，所以我们认为北京市的迁入人口转化为了城市人口。为了研究北京市每年净迁入人口的数目，我们假定，外来人口即为无户籍的常住人口。

19 数据的引用

- 流入人口数目

20 模型的再次修正

我们只要对等式 11 进行修正，加入外来人口的因素：

$$\sum_{e_l} N'''(a, r_0, e_l, t+1) = \sum_{e_l} N''(a, r_0, e_l, t+1) + M(a, t) \quad (14)$$

其中 $M(a, t)$ 是 t 时间段内迁入北京市的 a 岁人口数目， $N'''(a, r_0, e_l, t+1)$ 是这一次修正后城市各受教育程度的人口数。

根据统计数据，在一个固定研究时间段内， $M(a, t)$ 随着 a 的分布如表 7 所示。如果假定在各年龄层迁入人口保持相同比例，那么， $M(a, t)$ 与总迁入人数 $M(t)$ 的比例 $\phi_a = \frac{M(a, t)}{M(t)}$ 满足如下关系：

Table 7: 北京迁入人口随年龄的分布

年龄 a 分布	$\frac{\sum_a M(a, t)}{M(t)}$	ϕ_a
$[0, 15)$	7.00%	0.467%
$[15, 40)$	82.4%	3.296%
$[40, 60)$	9.8%	0.490%
$[60, \infty)$	0.8%	0.020%

北京市总迁入人口 $M(t)$ ，是基于历年的迁入人口拟合得出的：

$$M(t) = M_0 + M_1 \cdot e^{-\kappa(t-t_0)} \quad (15)$$

其中参数拟合结果如表 8 所示。

Table 8: 北京迁入人口参数拟合结果 $t_0=2010$

参数	M_0	M_1	κ
值			

21 模型的预测

这大清吃枣药丸

22 小结

唉

Part VIII

政策对社会造成的影响

苟利国家生死以，岂因祸福避趋之

Part IX

养老金模型

下面不加特殊说明时，研究对象均为北京市与北京市人口。

23 概述

抄抄抄

24 符号表

又可以拖页数啦

25 数据的引用

- 2010 年人均收入与养老金水平
- 养老金补贴指数与通货膨胀率保险金征收比例
- 北京养老基金还有多少钱

26 模型的建立

我们从养老金的支出与收入两方面来看待养老金的流动。对于 t 时间国家养老基金的已有积累，采用记号 $F(t)$ 。

26.1 支出

养老基金的支出，即为给符合条件的老龄人口的发放。为了简化问题，我们并不考虑养老金与退休人员的年龄之间的关系；亦即，退休人员在同一

年中获得相同的养老金。考虑 t 时间内养老基金的减少量 F_o :

$$F_o(t) = N_o(t) \cdot P(t) \quad (16)$$

其中 $N_o(t) = \sum_{A_R(t) \leq a} N(a, t)$ 是所有符合退休条件的人数, $P(t)$ 是养老金发放的金额。

而养老金的发放金额, 根据【来源请求】, 是由一个基准量加上修正而得到的; 我们可以依此列出方程:

$$P(t) = P_0 \cdot \gamma_p^{t-t_0} \quad (17)$$

其中 P_0 是 t_0 年养老金发放的平均水平, γ_p 是养老金上调幅度。

26.2 收入

养老基金的收入, 主要依赖于企事业员工定期缴纳的养老保险金。类似的, 我们并不考虑缴纳人员的收入差异, 而使用他们的平均工资代替。那么,

$$F_i(t) = N_i(t) \cdot C(t) \quad (18)$$

其中 $N_i(t) = \sum_{A_W(t) \leq a < A_R(t)} N(a, t)$ 是所有符合缴纳保险金的人数, $C(t)$ 是单个参保人员需要缴纳的保险金。

而单个参保人员需要缴纳的保险金, 根据【来源请求】, 是根据其收入与生活物价水平综合计算而得的; 于是

$$C(t) = S(t) \cdot \gamma_i^{t-t_0} \cdot \theta_p \quad (19)$$

其中 S_0 是 t_0 年人均收入, γ_i 是物价水平上涨指数, θ_p 是养老金征收比例。

小结

因此, 综合以上两方面, 我们写出养老基金积累的递推关系:

$$F(t+1) = F(t) - F_o(t) + F_i(t) \quad (20)$$

$$\begin{aligned} &= \sum_{A_R(t) \leq a} N(a, t) \cdot P_0 \cdot \gamma_p^{t-t_0} \\ &+ \sum_{A_W(t) \leq a < A_R(t)} N(a, t) S(t) \cdot \gamma_i^{t-t_0} \cdot \theta_p \end{aligned} \quad (21)$$

27 数据的预测

养老金肯定是发不出来的，嗯

28 二胎政策的影响

虽然我们还是很 naive，但是 zf 也不能把我们当傻瓜看啊，根本就没有什么卵用

29 小结

你问我吼不吼，那当然吼啊。

Part X

总结

蛤

参考文献

蛤