



中国-联合国信托基金项目
“利用普查数据进行人口预测”国际培训班

The background of the slide is a composite image. On the left, a white rocket is shown launching, with a large plume of white smoke and fire. On the right, a satellite is shown in orbit, emitting a bright light. The Great Wall of China is visible in the foreground, stretching across a landscape of green hills and mountains. The sky is a deep blue with some white clouds.

省际迁移与乡城迁移估计方法 交流

徐熙阳

北京信息控制研究所(航天710所)

2013年1月17日

交流内容



1. 城乡预测模型简介



2. 如何从普查资料分析迁移信息



3. 省际迁移估计过程与计算公式



4. 乡城迁移估计过程与计算公式



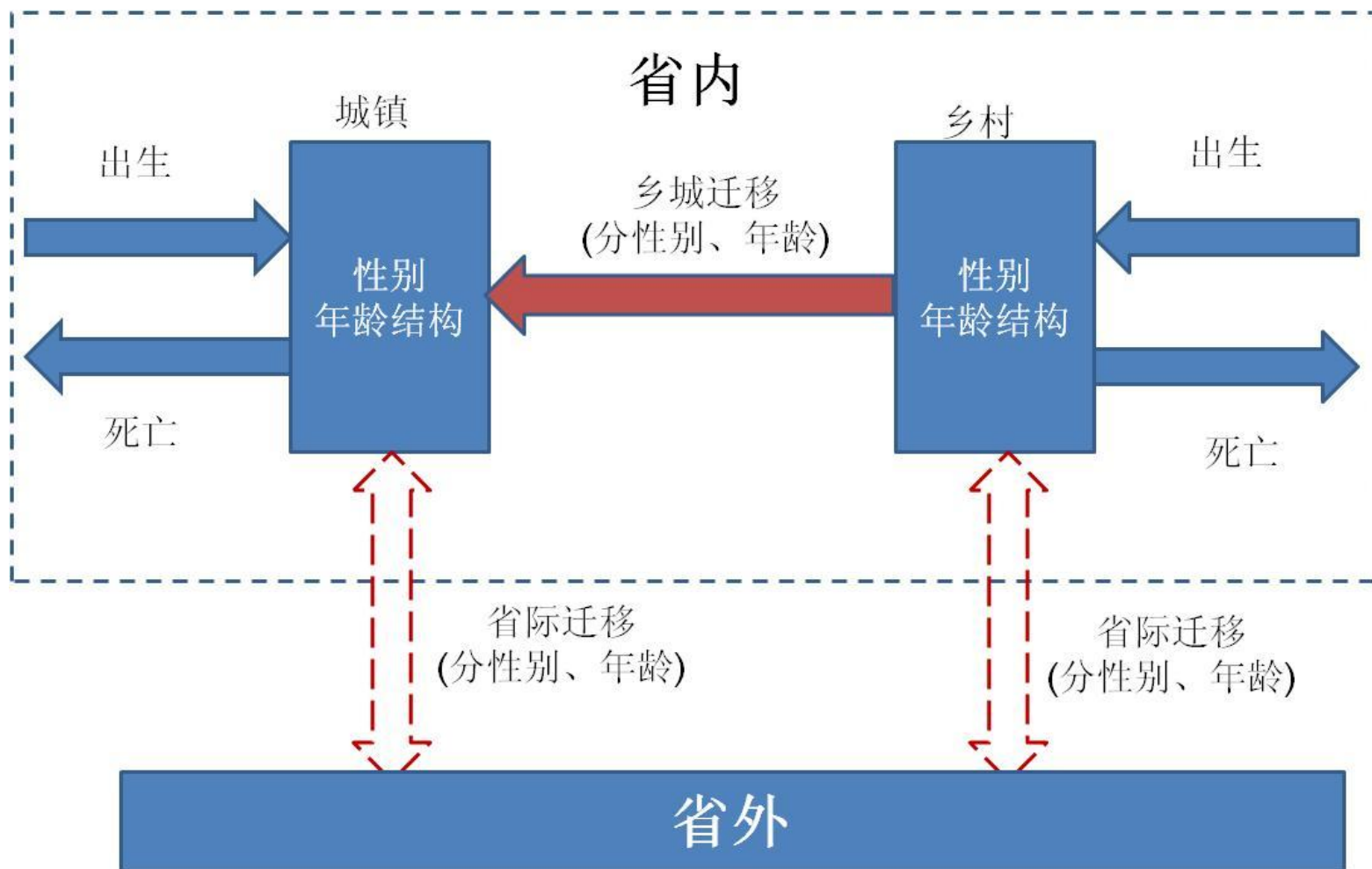
5. 估计结果比较

城乡预测模型简介—城镇化背景

- 李克强副总理指出，协调推进城镇化是实现现代化的重大战略选择，是未来经济增加的最大潜力。
- 目前，全国处于城镇化快速发展的阶段，未来每年将有大量人口由乡村转移到城镇。
- 考虑城镇化因素，开展未来城乡人口发展变化趋势预测，是十分必要并且非常及时。



城乡预测模型简介—省级区域人口变动框架



城乡预测模型简介—迁移的概念

- 迁移：由于某种目的或者动机，人们居住地在一定时期内（半年以上）发生改变。

包含以下意思：

- 地点：常住地改变，临时居住地改变叫做人口流动
- 时间：半年以上
- 特定地域范围：最小的迁移范围是乡、镇、街道
- 原因：具有明确的目的性和动因



城乡预测模型简介—影响人口迁移的因素

- 自然环境因素：由自然环境较差、不适宜居住的地区向自然环境较好、适宜居住地区迁移
- 经济因素：经济开发、经济重心转移或者经济布局改变引起的大规模的人口迁移。例如边疆开发、经济开发区设立等
- 社会因素：战争、婚姻、宗教、教育等



城乡预测模型简介—迁移的分类方法

- 按人口迁移跨越的行政地域划分：国际迁移和国内迁移。国内迁移又分省际迁移和省内迁移。进一步省内迁移还可以分县际迁移和县内迁移。
- 按人口迁移的经济地域划分：城镇与城镇之间的迁移、城镇与乡村之间的迁移、乡村与乡村之间的迁移
- 按人口迁移的原因划分：政治性、经济性、社会性、文化性。
- 按人口迁移居住时间长短划分：永久性人口迁移和非永久性人口迁移

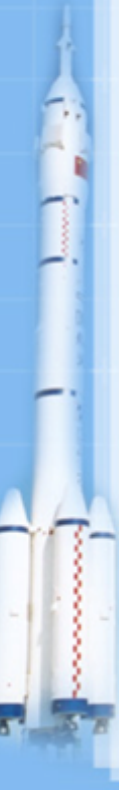
城乡预测模型简介—省级人口预测考虑哪些迁移

- 省级人口迁移类型
 - 跨国迁移:
 - 与港澳台之间的迁移:
 - 省际迁移:
 - 乡城迁移:
 - 现役军人:



城乡预测模型简介—关于迁移的基本假设

- 认为乡村到城镇的迁移者迁移当年的死亡率和出生率与迁出地相同，不受迁移的影响；
- 先考虑迁移行为后考虑死亡和生育行为；
- 在模型中考虑省际之间净迁移时，不论迁入还是迁出，模型预测所需各项参数均以本省参数为准，计算净迁移人数时，都以本省的城乡人口作为分母。



城乡预测模型简介—城乡预测所需数据与参数

- 基年分城乡、分性别、分年龄的人口数
- 基年育龄妇女生育模式、总和生育率
- 基年出生婴儿性别比
- 基年分城乡、分性别、分年龄的生命表

- 基年分性别、年龄的乡城迁移率
- 基年分城乡、分性别、分年龄的省际净迁移率



如何从普查资料分析迁移信息—如何认识迁移人口

- 迁移人口的一种动态指标：
 - 按年度进行统计
 - 明确迁移的地域范围
 - 迁移统计的人口范畴：常住人口或者户籍人口
- 由于我国近年来大量人口移动并未涉及户口的移动，所以从2000年五普开始对迁移人口采取统一定义：**居住在本乡镇街道半年以上、户口在外乡镇街道，以及在本乡镇街道居住不满半年、但是离开户口登记地半年以上的人口。**

如何从普查资料分析迁移信息

—常见人口迁移统计指标

- **迁入人口数(Immigration)**: 一定时期内迁入某地区的人口数量
- **迁出人口数(Emigration)**: 一定时期内迁出某地区的人口数量
- **净迁移人口数(Net-Migration)**: 一定时期内迁入某地区的人口数量减去同期迁出该地区的人口数量, 反映迁移变动的结果对本地区人口数量将产生的影响
- **总迁移人口数(Total-Migration)**: 实际参与迁移的人口总量

如何从普查资料分析迁移信息—短表问卷

• 普查问卷的分析—短表

- 本户地址
- H3本户应登记人数
- R6普查时居住地
- R7户口登记地
- R8离开户口登记地时间
- R9离开户口登记地原因
- R10户口性质

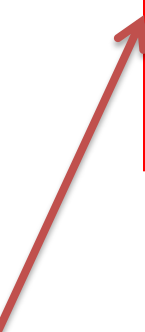
2010年10月31日晚居住本户的人数；
户口在本户，2010年10月31日晚未
居住本户的人数


- 1.本普查小区
- 2.本村（居）委会其他普查小区
- 3.本乡（镇、街道）其他村（居）委会
- 4.本县（市、区）其他乡（镇、街道）
- 5.其他县（市、区），请填写下面地址
_____省_____市_____县
- 6.港澳台或国外

如何从普查资料分析迁移信息—短表问卷

• 普查问卷的分析—短表

- 本户地址
- H3本户应登记人数
- R6普查时居住地
- R7户口登记地
- R8离开户口登记地时间
- R9离开户口登记地原因
- R10户口性质

- 
- 1.没有离开户口登记地→R11
 - 2.半年以下
 - 3.半年至一年
 - 4.一至二年
 - 5.二至三年
 - 6.三至四年
 - 7.四至五年
 - 8.五至六年
 - 9.六年以上

- 
- 1、务工经商
 - 2、工作调动
 - 3、学习培训
 - 4、随迁家属
 - 5、投亲靠友
 - 6、拆迁搬家
 - 7、寄挂户口
 - 8、婚姻嫁娶
 - 9、其他

如何从普查资料分析迁移信息—普查资料短表


— 第七卷 户口登记状况


- 7-1 全国按现住地、户口登记地、性别分的户口登记地在外乡镇街道的人口(城镇、乡村)
- 7-2 全国按户口登记地、年龄、性别分的户口登记地在外乡镇街道的人口(城镇、乡村)
- 7-3 全国按现住地、性别分的户口登记地在外省的人口(城镇、乡村)
- 7-4 全国按现住地、离开户口登记地时间分的户口登记地在外乡镇街道的人口(城镇、乡村)
- 7-5 全国按户口登记地、性别、受教育程度分的户口登记地在外乡镇街道的人口(城镇、乡村)
- 7-6 全国按现住地、受教育程度、性别分的户口登记地在本省其他乡镇街道人口(城镇、乡村)
- 7-7 全国按现住地、受教育程度、性别分的户口登记地在外省人口(城镇、乡村)
- 7-8 全国分年龄、性别、迁移原因的户口登记地在外乡镇街道的人口(城镇、乡村)
- 7-9 全国按现住地、性别、迁移原因分的户口登记地在本省其他乡镇街道人口(城镇、乡村)
- 7-10 全国按现住地、性别、迁移原因分的户口登记地在外省人口(城镇、乡村)

如何从普查资料分析迁移信息—长表问卷

• 普查问卷的分析—短表

- 本户地址
- H3本户应登记人数
- R6普查时居住地
- R7户口登记地
- R8离开户口登记地时间
- R9离开户口登记地原因
- R10户口登记地类型
- R11户口性质
- R12出生地
- R13五年前常住地（五周岁及以上）

- 
- 1.乡
 - 2.镇的居委会
 - 3.镇的村委会
 - 4.街道

- 
- 1.省内
 - 2.省外
——省

如何从普查资料分析迁移信息—普查资料长表

— 第七卷 迁移和户口登记地

- 7-1 全国按现住地、户口登记地类型分的户口登记地在外乡镇街道人口
- 7-2 全国按现住地、职业和性别分的户口登记地在本省其他乡镇街道人口
- 7-3 全国按现住地、职业和性别分的户口登记地在外省人口
- 7-4 全国按现住地、户口登记地、孩次分的2009.11.1-2010.10.31有过生育的妇女人数
- 7-5 全国按现住地、户口登记地类型、受教育程度分的户口登记地在本省其他乡镇街道人口
- 7-6 全国按现住地、户口登记地类型、受教育程度分的户口登记地在外省人口
- 7-7 全国按现住地和出生地分的人口
- 7-8 全国按现住地和五年前常住地分的人口

如何从普查资料分析迁移信息

一与省际迁移有关的信息分析

- 普查问卷与资料中与省际迁移有关的信息
 - 现住地：
 - 户口登记地：
 - 出生地：
 - 五年前常住地：



如何从普查资料分析迁移信息

一与乡城迁移有关的信息分析

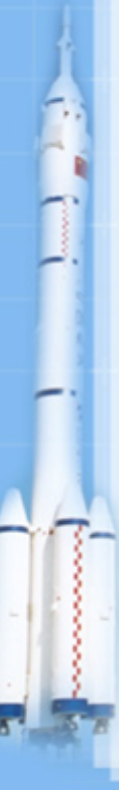
- 普查问卷与资料中与乡城迁移有关的信息
 - 现住地：
 - 户口登记地：
 - 户口登记地类型：
 - 户口性质：
 - 出生地：
 - 五年前常住地：



如何从普查资料分析迁移信息

一与乡城迁移有关的信息

- 其他可以估计乡城迁移有关的信息
 - 2000年五普现住地：
 - 2005年小普查现住地：



省际迁移估计过程与计算公式

1

使用数据与变量

2

有关假设

3

估计省际迁移概率

4

举例



省际迁移估计过程与计算公式—使用数据与变量

主要数据与变量

- 普查时分城乡、性别、年龄人口数 (0-100岁)
- 普查时分城乡、分性别、分年龄的居住在本省、五年前居住在省外的人口数 (5岁以上)
- 普查时居住在省外、五年前居住在本省的人数 (5岁以上)



省际迁移估计过程与计算公式—使用数据与变量

辅助数据与变量

- 加权调整倍数
- 分城乡、性别、年龄死亡率
- 城乡比例调整系数



省际迁移估计过程与计算公式—使用数据与变量

- 加权调整倍数
 - 五年前常住地只对长表进行统计，需要加权得到全省的数据



省际迁移估计过程与计算公式—使用数据与变量

- 三种加权方式：

- 方式1：加权倍数为10，对应10%的抽样比

- 方式2：短表与长表普查总人口数的倍数

- 方式3：

- 短表与长表居住在其他省、户籍在A省人数的倍数

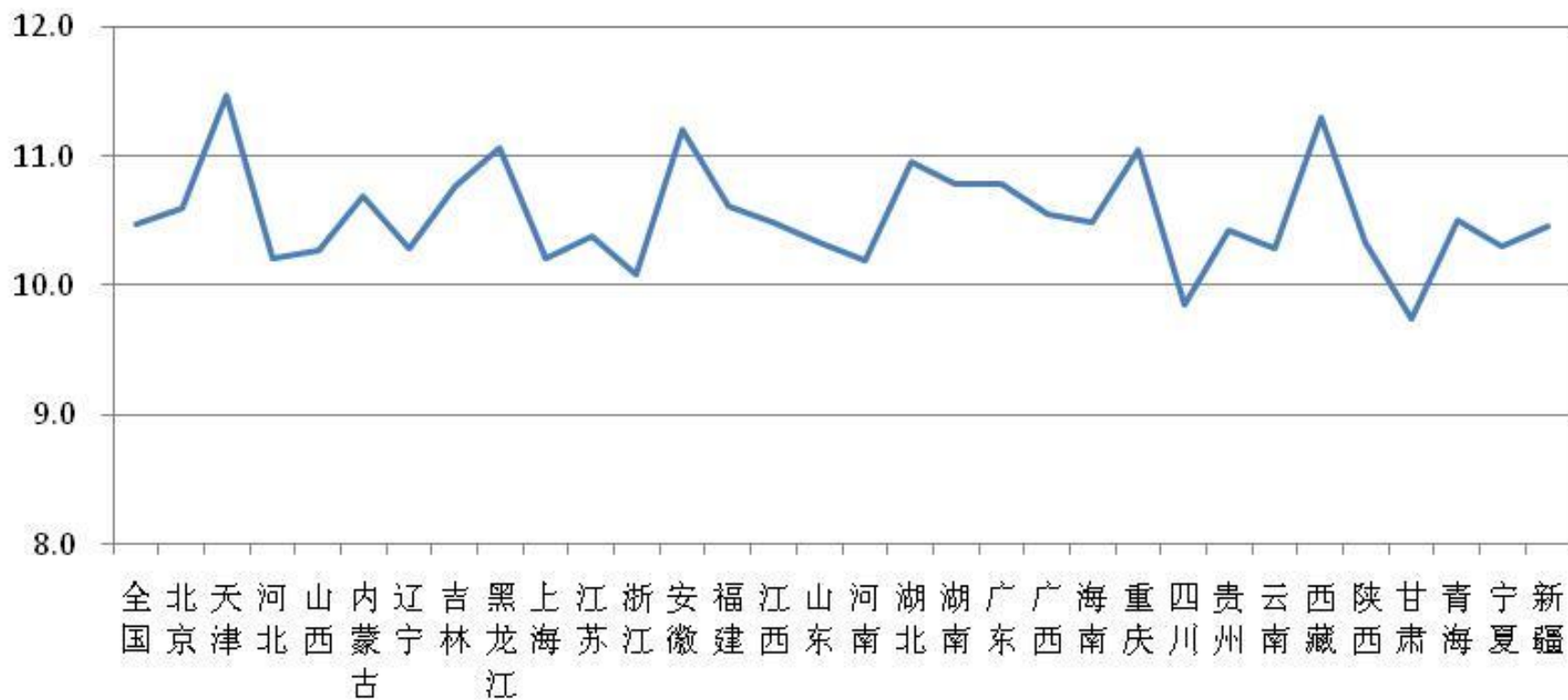
- 短表与长表居住在A省、户籍在外省人数的倍数



省际迁移估计过程与计算公式—使用数据与变量

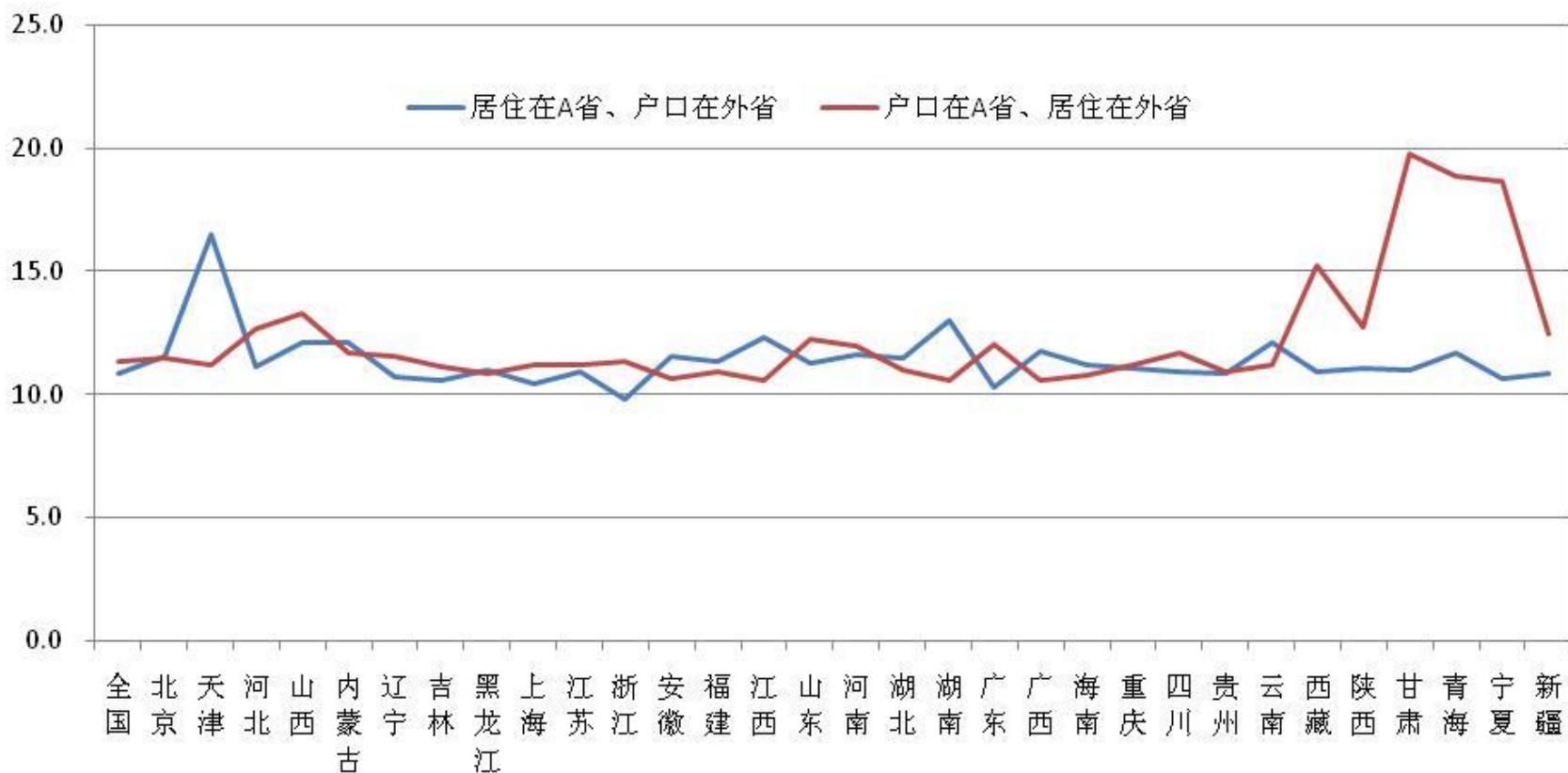
— 方式2:

总人口加权倍数



省际迁移估计过程与计算公式—使用数据与变量

— 方式3:



省际迁移估计过程与计算公式—使用数据与变量

- **分城乡、性别、年龄死亡率**

- 自2005年11月1日至2006年10月31日的前向死亡率
- 自2006年11月1日至2007年10月31日的前向死亡率
- 自2007年11月1日至2008年10月31日的前向死亡率
- 自2008年11月1日至2009年10月31日的前向死亡率
- 自2009年11月1日至2010年4月30日的前向死亡率



省际迁移估计过程与计算公式—使用数据与变量

- 如果没有各年的年龄别前向死亡率
 - Logit变换得到各年的死亡概率
 - 直接用普查数据得到的死亡概率代替



省际迁移估计过程与计算公式—使用数据与变量

- 城乡比例调整系数

通过汇总原始普查问卷，得到普查时居住在省外、五年前居住在A省的人数（5岁以上），以及现在居住地的城乡类型，无法得到迁出地的城乡类型，这里就只能近似处理。



省际迁移估计过程与计算公式—使用数据与变量

- 利用全国普查资料长表7.1的数据，认为乡、镇的村委会为乡村，街道、镇的居委会为城镇，可以得到以下a、b、c、d 4个数。

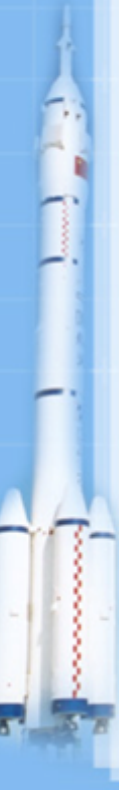
现住地	户口登记地类型	
	城镇（人）	乡村（人）
非A省-城镇	a	b
非A省-乡村	c	d

现住地	户口登记地类型	
	城镇	乡村
非A省-城镇	$a/(a+b)$	$b/(a+b)$
非A省-乡村	$c/(c+d)$	$d/(c+d)$

- 意味着普查时在非A省的其他省城镇居住、户口登记在A省的人中， $a/(a+b)$ 比例的人户口登记在A省城镇， $b/(a+b)$ 比例的人户口登记在A省乡村
- 同理，普查时在非A省的其他省乡村居住、户口登记在A省的人中， $c/(c+d)$ 比例的人户口登记在A省城镇， $d/(c+d)$ 比例的人户口登记在A省乡村

省际迁移估计过程与计算公式—使用数据与变量

- 由此对于直接由原始普查问卷得到的分城乡普查时居住在省外、五年前居住在A省的人数（5岁以上） x_1, x_2 ，使用以下公式进行处理可近似得到迁出地的城乡类型。
 - 来自A省城镇 $y_1 = x_1 * a / (a + b) + x_2 * c / (c + d)$
 - 来自A省乡村 $y_1 = x_1 * b / (a + b) + x_2 * d / (c + d)$



- 以A省为例：通过查阅全国普查资料，计算得到下表比例

现住地	户口登记地类型	
	城镇	乡村
非A省-城镇	0.25	0.75
非A省-乡村	0.12	0.88

同时，对原始普查数据进行分析，得到性别为 k 、年龄为 i 、5年前常住地在A省、目前在其他省城镇居住的人数 $x_{1,k,i}$ ；

性别为 k 、年龄为 i 、5年前常住地在A省、目前在其他省乡村居住的人数 $x_{2,k,i}$ ；

- 则性别为k、年龄为i、5年前常住地在A省城镇、目前在其他省居住的人数

$$Y_{1,k,i} = X_{1,k,i} \times 0.25 + X_{2,k,i} \times 0.12$$

- 则性别为k、年龄为i、5年前常住地在A省乡村、目前在其他省居住的人数

$$Y_{2,k,i} = X_{1,k,i} \times 0.75 + X_{2,k,i} \times 0.88$$

省际迁移估计过程与计算公式—有关假设

- 五年前常住地与现住地不在同一省的人，在普查半年前之前的4年半的时间内只有一次省际迁移
- 假设同批人每年发生省际迁移概率相等
- 假设5年内不同年份相同年龄人的净迁移概率相等



省际迁移估计过程与计算公式—有关假设

- 同批人每年发生省际迁移概率相等

$$m_{cw4}^k(2009) = m_{cw3}^k(2008) = m_{cw2}^k(2007) = m_{cw1}^k(2006) = m_{cw0}^k(2005)$$

$$m_{cw5}^k(2009) = m_{cw4}^k(2008) = m_{cw3}^k(2007) = m_{cw2}^k(2006) = m_{cw1}^k(2005)$$

$$m_{cw6}^k(2009) = m_{cw5}^k(2008) = m_{cw4}^k(2007) = m_{cw3}^k(2006) = m_{cw2}^k(2005)$$

$$m_{cw7}^k(2009) = m_{cw6}^k(2008) = m_{cw5}^k(2007) = m_{cw4}^k(2006) = m_{cw3}^k(2005)$$

$$m_{cw8}^k(2009) = m_{cw7}^k(2008) = m_{cw6}^k(2007) = m_{cw5}^k(2006) = m_{cw4}^k(2005)$$

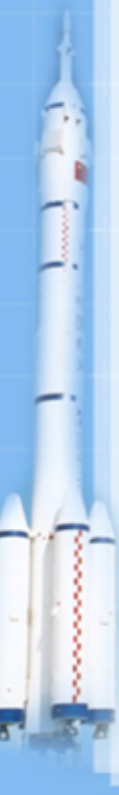
$$m_{cw9}^k(2009) = m_{cw8}^k(2008) = m_{cw7}^k(2007) = m_{cw6}^k(2006) = m_{cw5}^k(2005)$$

$$m_{cwi-1}^k(2009) = m_{cwi-2}^k(2008) = m_{cwi-3}^k(2007) = m_{cwi-4}^k(2006) = m_{cwi-5}^k(2005) \quad i > 10$$

省际迁移估计过程与计算公式—有关假设

- 不同年份同龄人的净迁移概率相等

$$m_{cwi}^k(2009) = m_{cwi}^k(2008) = m_{cwi}^k(2007) = m_{cwi}^k(2006) = m_{cwi}^k(2005)$$



省际迁移估计过程与计算公式—估计省际迁移概率

- 估计5岁以上年度平均省际迁移概率

利用加权调整后的分城乡、性别、年龄的普查时常住本省、五年前常住外省人数与普查时常住外省、五年前常住本省人数之差，除以普查时本省分城乡、性别、年龄的人口数，再除以在4年半内迁移的人口存活到普查时的累积存活概率之和，得到年度平均省际迁移概率。



城镇k性别i岁人口在t年前的一年内发生省际迁移的净迁移概率

城镇k性别i岁人口在过去5年内的净迁移人数，等于城镇k性别i岁人口中普查时居住在本省5年前在省外的人数减去普查时居住在省外5年前居住在本省的人数

$${}^{-k}m_{cwi}(t) \approx \frac{y_{cwi}^k(t)}{x_{cwi}^k(t)} \div \sum_{j=i-4}^{i-1} \eta_{cj}^k(t+j-i) + \eta_{ci}^k(t) / 2 \quad (i \geq 5)$$

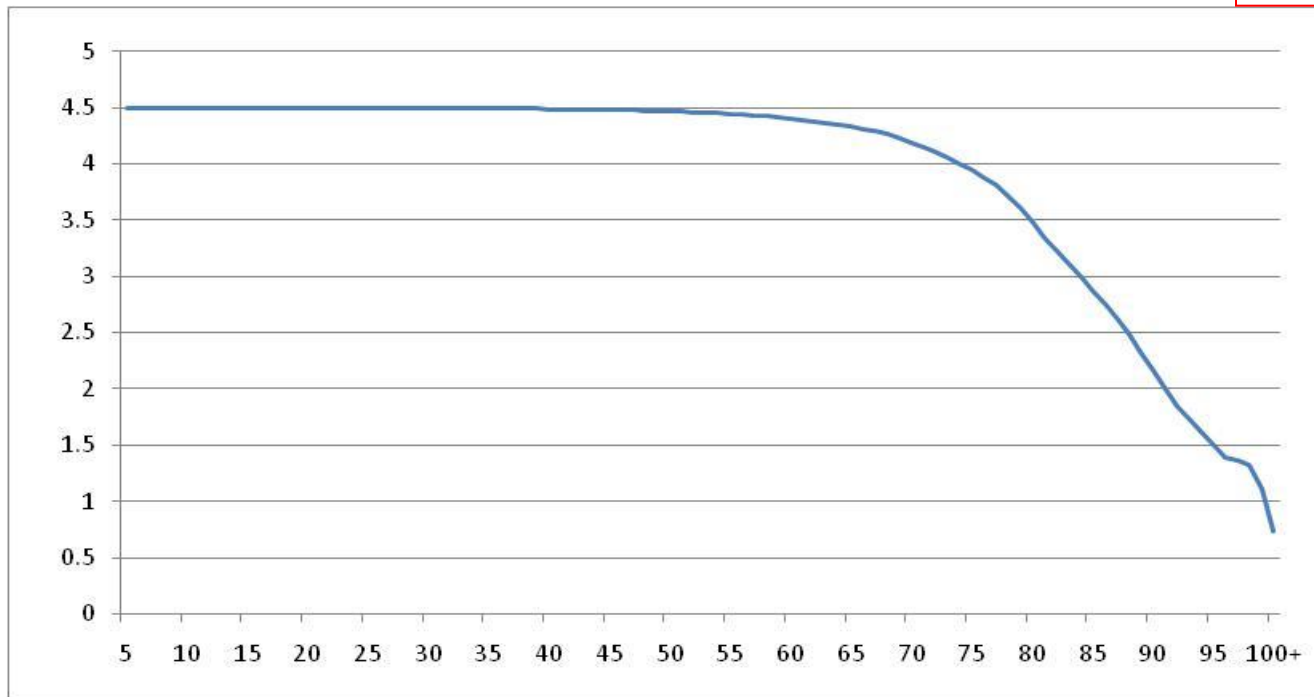
普查时城镇的k性别i岁人口数

在t+j-i年进行了省际迁移的城镇k性别j岁人口存活到普查时i岁的累积留存率

其中， $\eta_{cj}^k(t+j-i) = \prod_{s=j}^{i-1} (1 - \mu_{cs}^k(t+s-i))$

其值小于4.5，并且随着年龄的增加，逐渐减小

t+s-i年城镇的k性别s岁人口的前向死亡率



乡村k性别i岁人口在t年前的一年内发生省际迁移的净迁移概率

乡村k性别i岁人口在过去5年内的净迁移人数，等于乡村k性别i岁人口中普查时居住在本省5年前在省外的人数减去普查时居住在本省5年前居住在本省的人数

$${}^{-k}m_{rwi}(t) \approx \frac{y_{rwi}^k(t)}{x_{rwi}^k(t)} \div \sum_{j=i-4}^{i-1} \eta_{rj}^k(t+j-i) + \eta_{ri}^k(t) / 2 \quad (i \geq 5)$$

普查时乡村的k性别i岁人口数

在t+j-i年进行了省际迁移的乡村k性别j岁人口存活到普查时i岁的累积留存率

省际迁移估计过程与计算公式—估计省际迁移概率

- 估计全年龄省际迁移概率

结合上述假设2和假设3，可以得到全年龄的省际迁移概率。



根据假设2和假设3，则

$$m_{cw5}^{-k}(2010) = m_{cw4}^k(2009) = m_{cw3}^k(2008) = m_{cw2}^k(2007) = m_{cw1}^k(2006) = m_{cw0}^k(2005)$$

$$m_{cw6}^{-k}(2010) = m_{cw5}^k(2009) = m_{cw4}^k(2008) = m_{cw3}^k(2007) = m_{cw2}^k(2006) = m_{cw1}^k(2005)$$

$$m_{cw7}^{-k}(2010) = m_{cw6}^k(2009) = m_{cw5}^k(2008) = m_{cw4}^k(2007) = m_{cw3}^k(2006) = m_{cw2}^k(2005)$$

$$m_{cw8}^{-k}(2010) = m_{cw7}^k(2009) = m_{cw6}^k(2008) = m_{cw5}^k(2007) = m_{cw4}^k(2006) = m_{cw3}^k(2005)$$

$$m_{cw9}^{-k}(2010) = m_{cw8}^k(2009) = m_{cw7}^k(2008) = m_{cw6}^k(2007) = m_{cw5}^k(2006) = m_{cw4}^k(2005)$$

$$m_{cw10}^{-k}(2010) = m_{cw9}^k(2009) = m_{cw8}^k(2008) = m_{cw7}^k(2007) = m_{cw6}^k(2006) = m_{cw5}^k(2005)$$

$$m_{cwi}^{-k}(2010) = m_{cwi-1}^k(2009) = m_{cwi-2}^k(2008) = m_{cwi-3}^k(2007) = m_{cwi-4}^k(2006) = m_{cwi-5}^k(2005) \quad i > 10$$

省际迁移估计过程与计算公式—估计省际迁移概率

- 估计2010年各年龄人口的净迁移概率

$$m_{cw0}^k(2010) \approx m_{cw0}^k(2005) = \bar{m}_{cw5}^{-k}(2010)$$

$$m_{cw1}^k(2010) \approx m_{cw1}^k(2005) \approx m_{cw1}^k(2006) = [\bar{m}_{cw5}^{-k}(2010) + \bar{m}_{cw6}^{-k}(2010)]/2$$

$$\begin{aligned} m_{cw2}^k(2010) &\approx m_{cw2}^k(2005) \approx m_{cw2}^k(2006) \approx m_{cw2}^k(2007) \\ &= [\bar{m}_{cw5}^{-k}(2010) + \bar{m}_{cw6}^{-k}(2010) + \bar{m}_{cw7}^{-k}(2010)]/3 \end{aligned}$$

$$\begin{aligned} m_{cw3}^k(2010) &\approx m_{cw3}^k(2005) \approx m_{cw3}^k(2006) \approx m_{cw3}^k(2007) \approx m_{cw3}^k(2008) \\ &= [\bar{m}_{cw5}^{-k}(2010) + \bar{m}_{cw6}^{-k}(2010) + \bar{m}_{cw7}^{-k}(2010) + \bar{m}_{cw8}^{-k}(2010)]/4 \end{aligned}$$

$$\begin{aligned} m_{cw4}^k(2010) &\approx m_{cw4}^k(2005) \approx m_{cw4}^k(2006) \approx m_{cw4}^k(2007) \approx m_{cw4}^k(2008) \approx m_{cw4}^k(2009) \\ &= [\bar{m}_{cw5}^{-k}(2010) + \bar{m}_{cw6}^{-k}(2010) + \bar{m}_{cw7}^{-k}(2010) + \bar{m}_{cw8}^{-k}(2010) + \bar{m}_{cw9}^{-k}(2010)]/5 \end{aligned}$$

$$m_{cwi}^k(2010) \approx m_{cwi}^k(2005) \approx m_{cwi}^k(2006) \approx m_{cwi}^k(2007) \approx m_{cwi}^k(2008) \approx m_{cwi}^k(2009) \\ = [m_{cwi+1}^{-k}(2010) + m_{cwi+2}^{-k}(2010) + m_{cwi+3}^{-k}(2010) + m_{cwi+4}^{-k}(2010) + m_{cwi+5}^{-k}(2010)]/5 \quad (5 \leq i \leq 95)$$

$$m_{cw96}^k(2010) \approx m_{cw96}^k(2006) \approx m_{cw96}^k(2007) \approx m_{cw96}^k(2008) \approx m_{cw96}^k(2009) \\ = [m_{cw97}^{-k}(2010) + m_{cw98}^{-k}(2010) + m_{cw99}^{-k}(2010) + m_{cw100}^{-k}(2010)]/4$$

$$m_{cw97}^k(2010) \approx m_{cw97}^k(2007) \approx m_{cw97}^k(2008) \approx m_{cw97}^k(2009) \\ = [m_{cw98}^{-k}(2010) + m_{cw99}^{-k}(2010) + m_{cw100}^{-k}(2010)]/3$$

$$m_{cw98}^k(2010) \approx m_{cw98}^k(2008) \approx m_{cw98}^k(2009) = [m_{cw99}^{-k}(2010) + m_{cw100}^{-k}(2010)]/2$$

$$m_{cw99}^k(2010) \approx m_{cw99}^k(2009) = m_{cw100}^{-k}(2010)$$

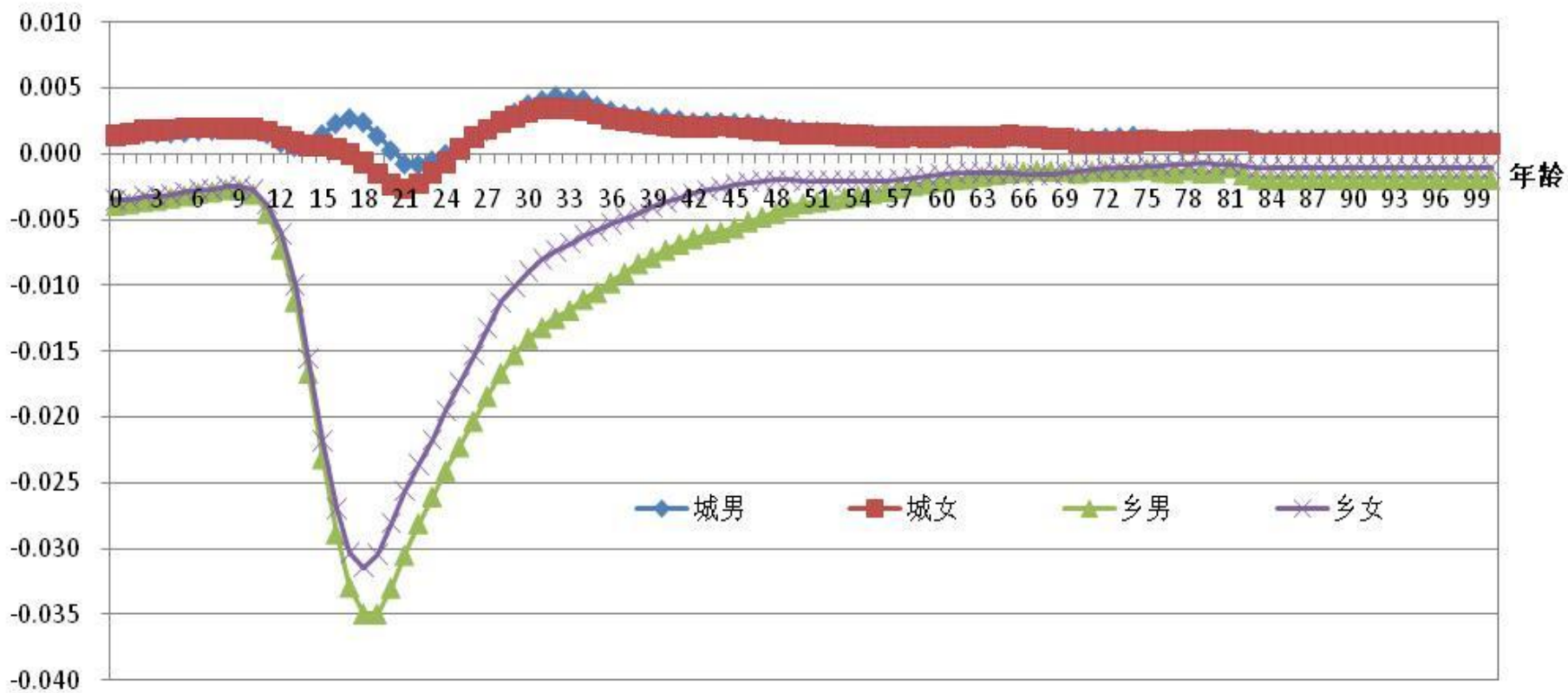
$$m_{cw100}^k(2010) \approx m_{cw100}^{-k}(2010)$$

省际迁移估计过程与计算公式—估计省际迁移概率

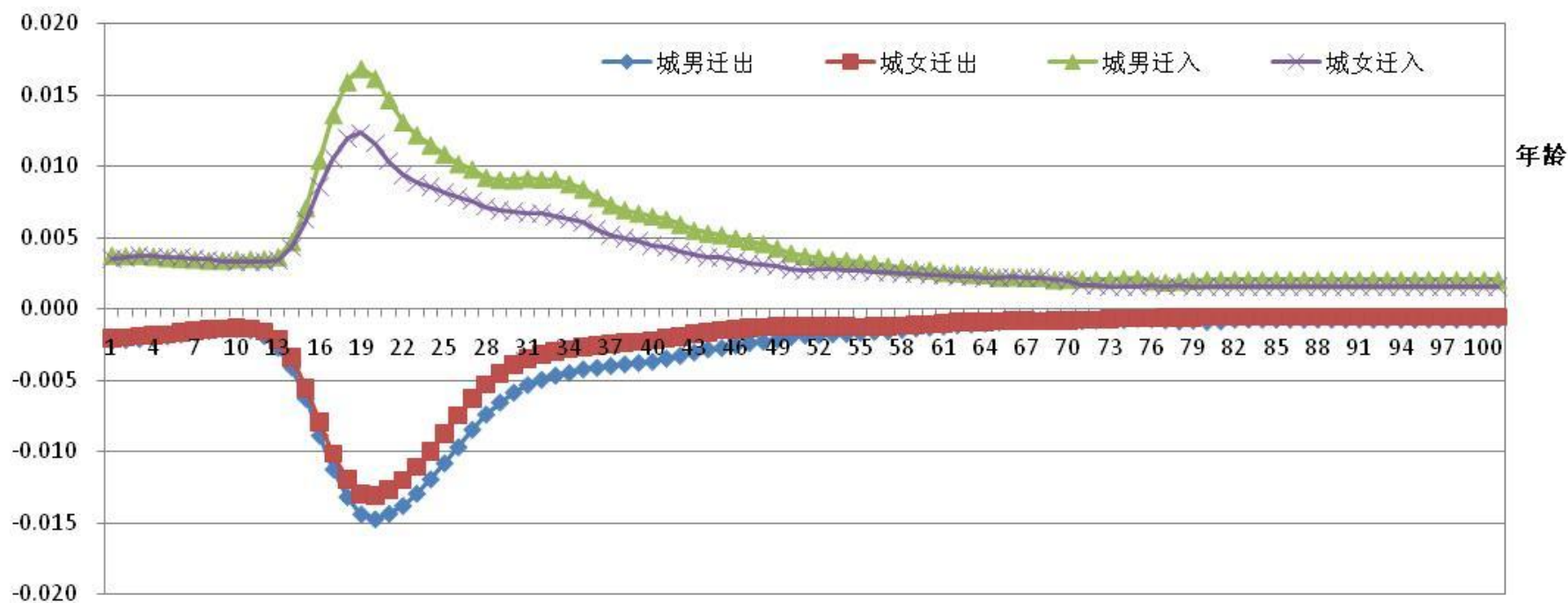
- 根据上述公式可以估计全省全年龄人口的省际净迁移概率
- 相同方法可以估计分城乡、性别的省际净迁移概率



省际迁移估计过程与计算公式—举例



省际迁移估计过程与计算公式—城镇举例



乡城迁移估计过程与计算公式

1

有关假设

2

使用的数据与变量

3

估计乡城迁移率过程与步骤

4

举例

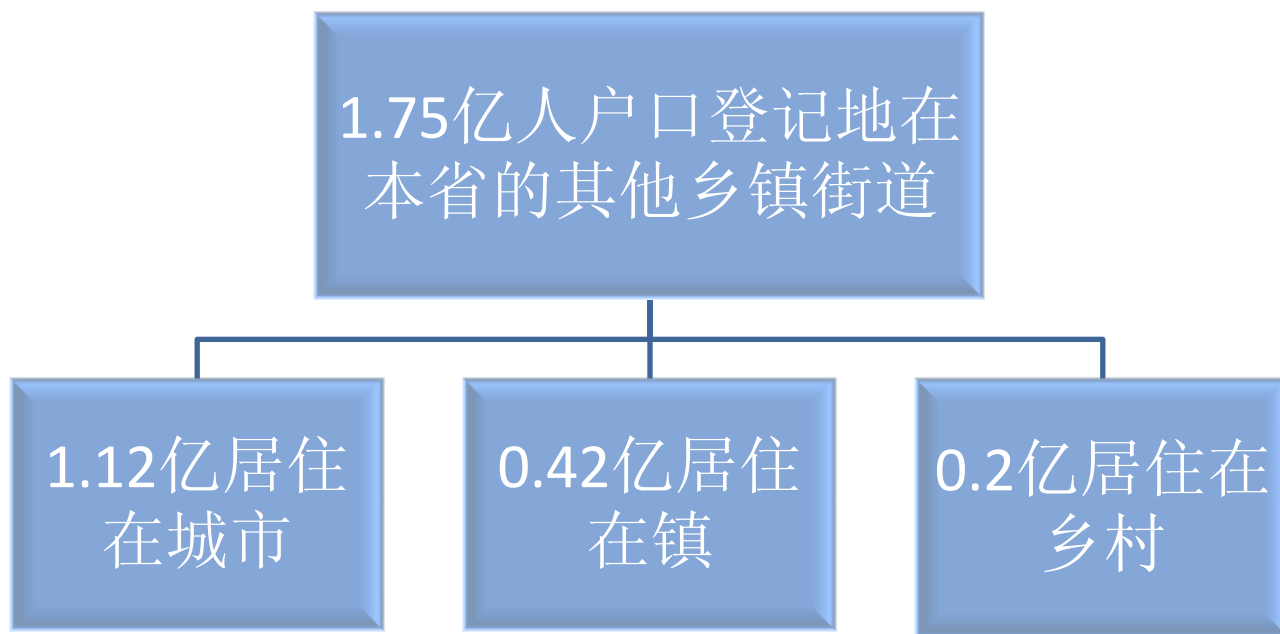
5

简化算法



乡村迁移估计过程与计算公式—有关假设

- 主要考虑乡城之间的净迁移



乡村迁移估计过程与计算公式—有关假设

- 假设只有一次永久性迁移，可以是2000年至2010年之间的任何一年
- 假设乡村人口转移到城镇后第二年开始按照城镇人口的死亡模式死亡



乡村迁移估计过程与计算公式—数据与变量

- 主要数据与变量

- (1) 2000年、2010年普查分城乡、性别、年龄人数
(考虑漏报调整后)
- (2) 2000年-2010年普查分城乡、性别、年龄死亡概率
- (3) 2000 -2010年历年的分城乡出生率、死亡率、自然增长率、出生性别比
- (4) 2000 -2010年历年的人口数及城镇化率



乡村迁移估计过程与计算公式—估计过程与步骤

- (一) • 估计2000年底、2010年底人口结构与2001-2010年出生人数
- (二) • 估计未发生乡城迁移的情况下的2010年底人数
- (三) • 估计10年累计由乡村迁移到城镇并存活到2010年底人数
- (四) • 估计过去10年乡城迁移规模
- (五) • 估计过去10年累计迁移人数与每年的迁移人数
- (六) • 估计 t 年满 i 岁未满 $i+1$ 岁进行迁移的人数
- (七) • 估计2001年底分城乡性别年龄人口数
- (八) • 计算出2001年的各年龄的乡城迁移率
- (九) • 计算出各年龄的乡城迁移率
- (十) • 加权得到2010年乡城迁移率

(一) 估计2000年底、2010年底分城乡、性别、年龄人口数与2001-2010年底0岁人口数

- 根据2000年普查资料和2001年人口统计年鉴，估计出2000年底分城乡、性别、年龄的人口结构，分别记为 $x_{ci}^k(2000)$ 和 $x_{ci}^k(2000)$ 。
- 根据2010年普查资料和2011年人口统计年鉴，估计出2010年底分城乡、性别、年龄的人口结构，分别记为 $x_{ci}^k(2010)$ 和 $x_{ci}^k(2010)$ 。

注意：结合城镇化率

(一) 估计2000年底、2010年底分城乡、性别、年龄人口数与2001-2010年底0岁人口数

- 估计2001年至2010年年底0岁人口数

根据2001年至2010年统计年鉴，可以计算出2001年至2010年的出生人数，根据每年出生人口的城乡比例与出生性别比，分别计算出分城乡、性别的出生人数，结合婴儿当年死亡率，可以计算出得到各年年底的0岁人口数，记为 $x_{c0}^k(t)$ 和 $x_{r0}^k(t)$

如果缺少出生人口的城乡比例和性别比例，则借助两次普查年的比例进行内插。

(二) 估计没有乡城迁移情况下2010年底的分城乡、性别、年龄的人口数

$$z_{ci}^k(2010) = x_{ci-10}^k(2000) \prod_{s=i-10}^{i-1} (1 - \mu_{cs}^k(2011+s-i)) \prod_{s=i-10}^{i-1} (1 - m_{cws}^k(2011+s-i))$$

前向死亡率

省际迁移率

$$z_{ri}^k(2010) = x_{ri-10}^k(2000) \prod_{s=i-10}^{i-1} (1 - \mu_{rs}^k(2011+s-i)) \prod_{s=i-10}^{i-1} (1 - m_{rws}^k(2011+s-i))$$

$i \geq 10$

(二) 估计没有乡城迁移情况下2010年底的分城乡、性别、年龄的人口数

$$z_{ci}^k(2010) = x_{c0}^k(2010-i) \prod_{s=0}^{i-1} (1 - \mu_{cs}^k(2011+s-i)) \prod_{s=0}^{i-1} (1 - m_{cws}^k(2011+s-i))$$

前向死亡率

省际迁移率

$$z_{ri}^k(2010) = x_{r0}^k(2010-i) \prod_{s=0}^{i-1} (1 - \mu_{rs}^k(2011+s-i)) \prod_{s=0}^{i-1} (1 - m_{rws}^k(2011+s-i))$$

$i < 10$

(三) 估计过去10年累计由乡村迁移到城镇
并存活到2010年底人数

$$\Delta_{ci}^k(2010) = x_{ci}^k(2010) - z_{ci}^k(2010)$$

$$\Delta_{ri}^k(2010) = x_{ri}^k(2010) - z_{ri}^k(2010)$$

$$\bar{\Delta}_{rci}^k(t) = (|\Delta_{ci}^k(t)| + |\Delta_{ri}^k(t)|) / 2$$



(四) 估计每年的乡城迁移规模与迁移倍数

$$M(t) = X(t) \times U(t) - X(t-1) \times U(t-1) - \bar{X}_c(t)(B_c(t) - D_c(t))$$

$M(t)$ 为 t 年乡村向城镇转移人数的估计值

$X(t)$ 为 t 年年底总人数

$U(t)$ 为 t 年年底城镇化率

$\bar{X}_c(t)$ 为 t 年城镇的平均人口数

$B_c(t)$ 为 t 年城镇的出生率

$D_c(t)$ 为 t 年城镇的死亡率

(四) 估计每年的乡城迁移规模与迁移倍数

年份	乡城转移人数 $M(t)$	所占比例 $Q(t)$	迁移倍数 $P(t)=$ $Q(t)*10$
2001	$M(2001)$	$M(2001)/\Sigma M(t)$	$P(2001)$
2002	$M(2002)$	$M(2002)/\Sigma M(t)$	$P(2002)$
2003	$M(2003)$	$M(2003)/\Sigma M(t)$	$P(2003)$
2004	$M(2004)$	$M(2004)/\Sigma M(t)$	$P(2004)$
2005	$M(2005)$	$M(2005)/\Sigma M(t)$	$P(2005)$
2006	$M(2006)$	$M(2006)/\Sigma M(t)$	$P(2006)$
2007	$M(2007)$	$M(2007)/\Sigma M(t)$	$P(2007)$
2008	$M(2008)$	$M(2008)/\Sigma M(t)$	$P(2008)$
2009	$M(2009)$	$M(2009)/\Sigma M(t)$	$P(2009)$
2010	$M(2010)$	$M(2010)/\Sigma M(t)$	$P(2010)$

(五) 估计2010年底*i*年龄人口在过去10年内 每年进行迁移的人数

实际每年发生乡城迁移的人数不等，假设与迁移
倍数 $P(t)$ 成正比例关系

$$M_{rci}^k(t) = P(t) \times \overline{M}_{rci}^k(t)$$

$\overline{M}_{rci}^k(t)$ 为平均迁移人数呈现均匀分布，每年相等

(五) 估计2010年底*i*年龄人口在过去10年内 每年进行迁移的人数

- 对2010年底*i*岁人口，如果*i*小于10，则*t*年迁移人口存活到2010年底的存活率为

$$\eta_{rci}^k(t) = \prod_{e=t}^{2010} \left[1 - U_{c(e+i-2011)}^k(e) \right]$$

*e*年城镇*k*性别 (*e+i-2011*)
岁人口的前向死亡率

*t*年迁移倍数

- 累计留存率为 $\sum_{t=2011-i}^{2010} \eta_{rci}^k(t) \times P(t)$

(五) 估计2010年底*i*年龄人口在过去10年内 每年进行迁移的人数

- 对2010年底*i*岁人口，如果*i*大于等于10，则*t*年迁移人口存活到2010年底的存活率为

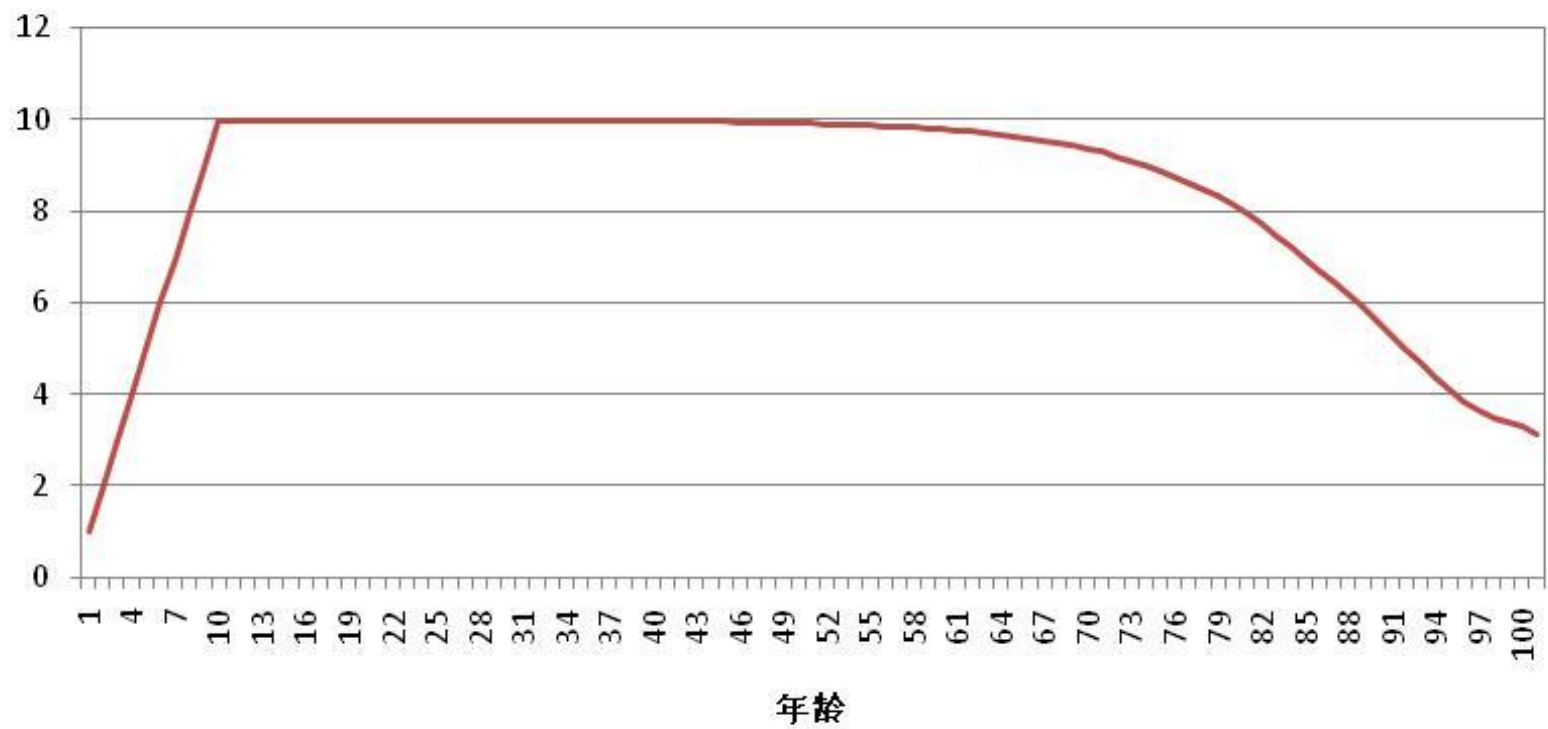
$$\eta_{rci}^k(t) = \prod_{e=t}^{2010} \left[1 - U_{c(e+i-2011)}^k(e) \right]$$

*e*年城镇*k*性别 (*e+i-2011*)
岁人口的前向死亡率

*t*年迁移倍数

- 累计留存率为 $\sum_{t=2001}^{2010} \eta_{rci}^k(t) \times P(t)$

累计留存率曲线



(五) 估计2010年底*i*年龄人口在过去10年内 每年进行迁移的人数

- 过去10年内*k*性别*i*岁平均每年的乡城迁移人数为

$$\bar{M}_{rci}^k(t) = \frac{\bar{\Delta}_{rci}^k(t)}{\sum_t^{2010} \eta_{rci}^k(t) \times P(t)}$$



(五) 估计2010年底*i*年龄人口在过去10年内 每年进行迁移的人数

- 从而得到各年的分年龄的迁移人数:

$$M_{rci}^k(2001) = P(2001) \times \overline{M}_{rci}^k(t)$$

$$M_{rci}^k(2002) = P(2002) \times \overline{M}_{rci}^k(t)$$

.....

.....

$$M_{rci}^k(2009) = P(2009) \times \overline{M}_{rci}^k(t)$$

$$M_{rci}^k(2010) = P(2010) \times \overline{M}_{rci}^k(t)$$



(六) 估计在过去第 t 年满 $j-1$ 岁未满足 j 岁进行乡
城迁移的人数

$$M_{rcj}^{1k}(t) = \frac{1}{2} (M_{rcj}^k(t) + M_{rcj-1}^k(t))$$



(七) 估计2001年底的分城乡性别年龄人口数

$$x_{cj}^k(2001) = (x_{cj-1}^k(2000) - x_{cj-1}^k(2000)m_{cwj-1}^k(2001)) \\ \times (1 - \mu_{cj-1}^k(2001)) + \frac{1}{2} (M_{rcj}^k(2001) + M_{rcj-1}^k(2001))$$

$$x_{rj}^k(2001) = (x_{rj-1}^k(2000) - x_{rj-1}^k(2000)m_{rwj-1}^k(2001)) \\ \times (1 - \mu_{rj-1}^k(2001)) - \frac{1}{2} (M_{rcj}^k(2001) + M_{rcj-1}^k(2001))$$

(j ≥ 1)

$$x_{rj}^{-k}(2001) = \frac{[x_{rj}^k(2000) + x_{rj}^k(2001)]}{2}$$

(八) 估计2001年底的分城乡性别年龄人口的乡城迁移率

$$m_{rj}^k(2001) = \frac{1}{2} (M_{rcj}^k(2001) + M_{rcj-1}^k(2001)) / x_{ri}^{-k}(2001)$$



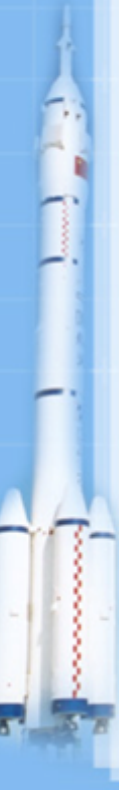
(九) 估计2002-2010年各年龄的乡城迁移率

- 重复上述（七）和（八），逐年计算出2002-2010年的分性别和年龄乡城迁移率



(十) 估计2010年乡城迁移概率

- 对上述计算出的历年的迁移率，根据 $P(t)$ 进行加权平均得到近似的乡城迁移概率
- 平滑处理
- 根据有关数据进行调整



五点三次平滑方法

(Cubical Smoothing algorithm with five-point Approximation)

$$y_{i-2} = (69x_{i-2} + 4x_{i-1} - 6x_i + 4x_{i+1} - x_{i+2}) / 70$$

$$y_{i-1} = (2x_{i-2} + 27x_{i-1} + 12x_i - 8x_{i+1} + 2x_{i+2}) / 35$$

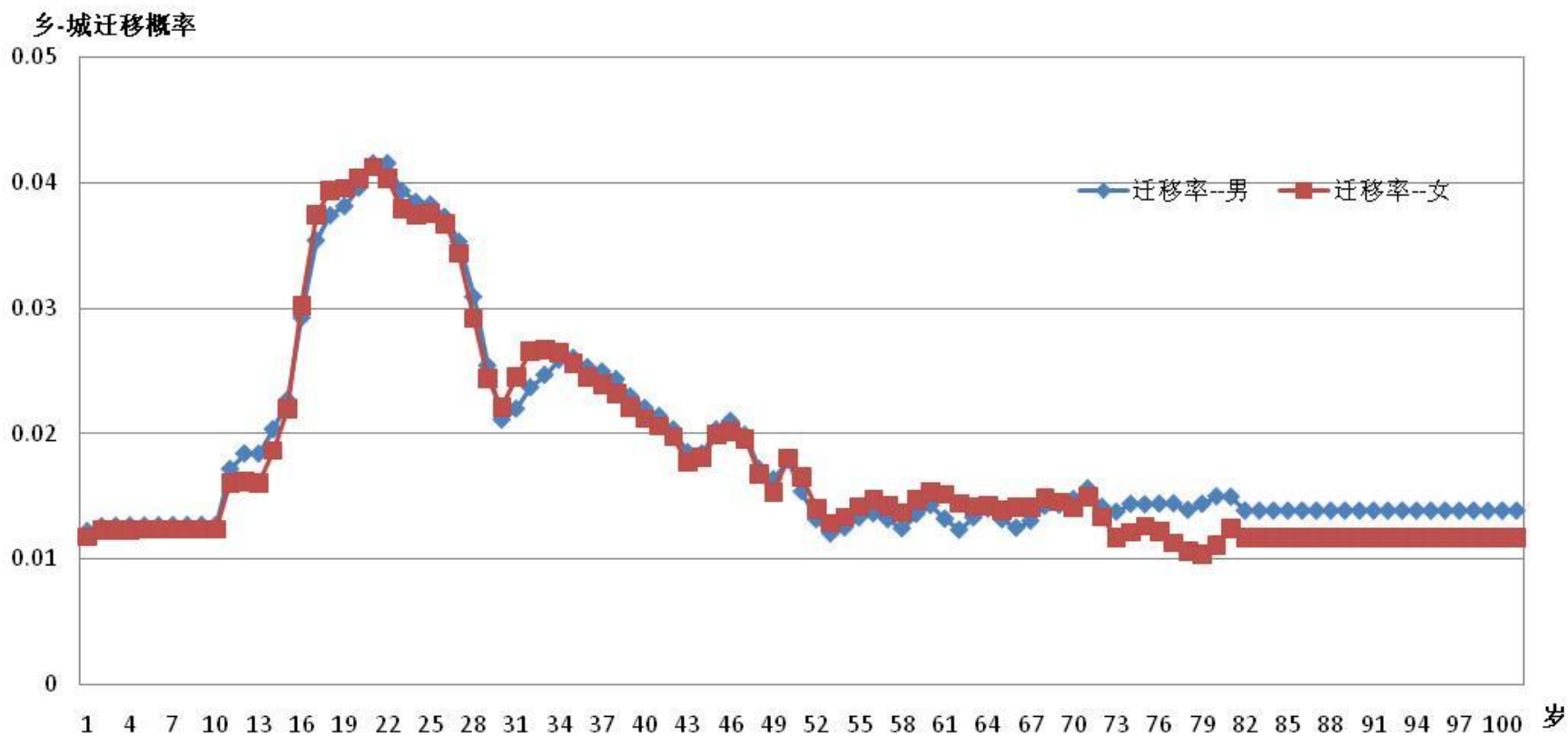
$$y_i = (-3x_{i-2} + 12x_{i-1} + 17x_i + 12x_{i+1} - 3x_{i+2}) / 35$$

$$y_{i+1} = (2x_{i-2} - 8x_{i-1} + 12x_i + 27x_{i+1} + 2x_{i+2}) / 35$$

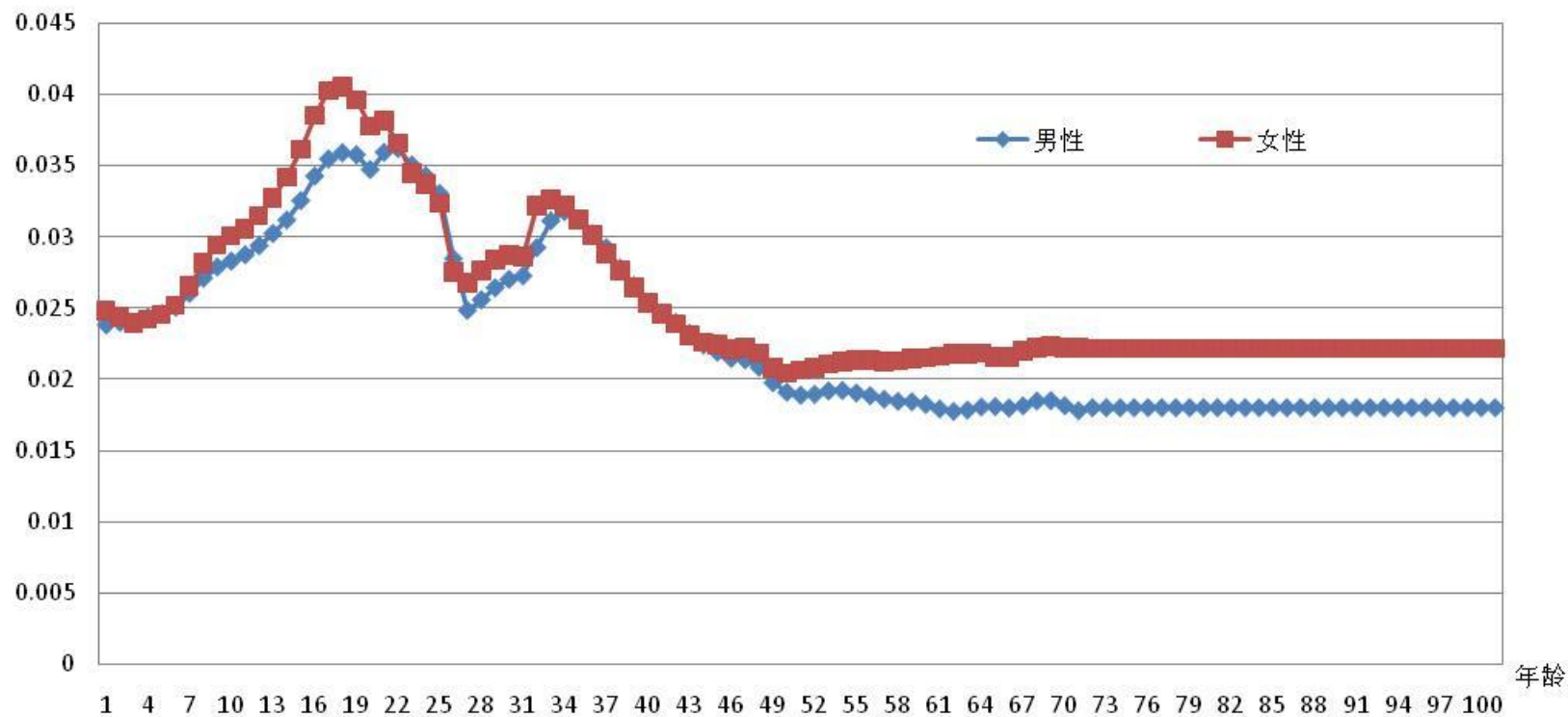
$$y_{i+2} = (-x_{i-2} + 4x_{i-1} - 6x_i + 4x_{i+1} + 69x_{i+2}) / 70$$



乡村迁移估计过程与计算公式—以全国数据举例



乡村迁移估计过程与计算公式—以某省数据举例



乡村迁移估计过程与计算公式—简化算法

- 进一步假设：
 - 每年迁移人数相等，迁移倍数为1
 - 以10年分性别、年龄的死亡平均值代替每一年的死亡概率
 - 两次普查的乡村人口数的均值为10年平均乡村人口数



乡村迁移估计过程与计算公式—简化算法

(一)

- 估计未发生乡城迁移的情况下的2010年普查时人数

(三)

- 估计10年累计由乡村迁移到城镇并存活到2010年底人数

(三)

- 估计过去10年累计迁移人数与每年的迁移人数

(四)

- 估计过去10年乡村平均人口数

(五)

- 计算出2001年的各年龄的乡城迁移率

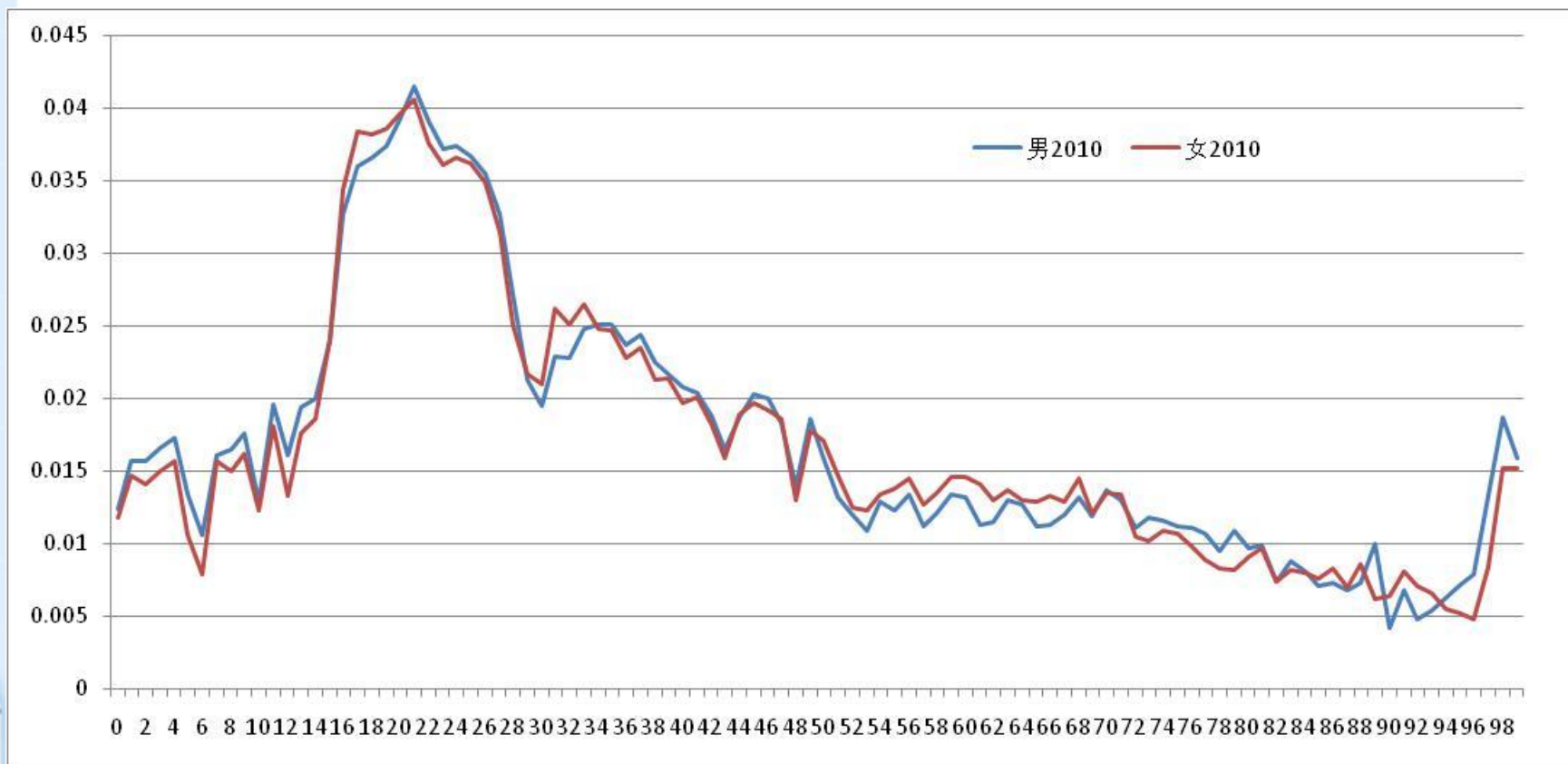
(六)

- 计算出各年龄的乡城迁移率

(七)

- 光滑处理

乡村迁移估计过程与计算公式—简化算法举例



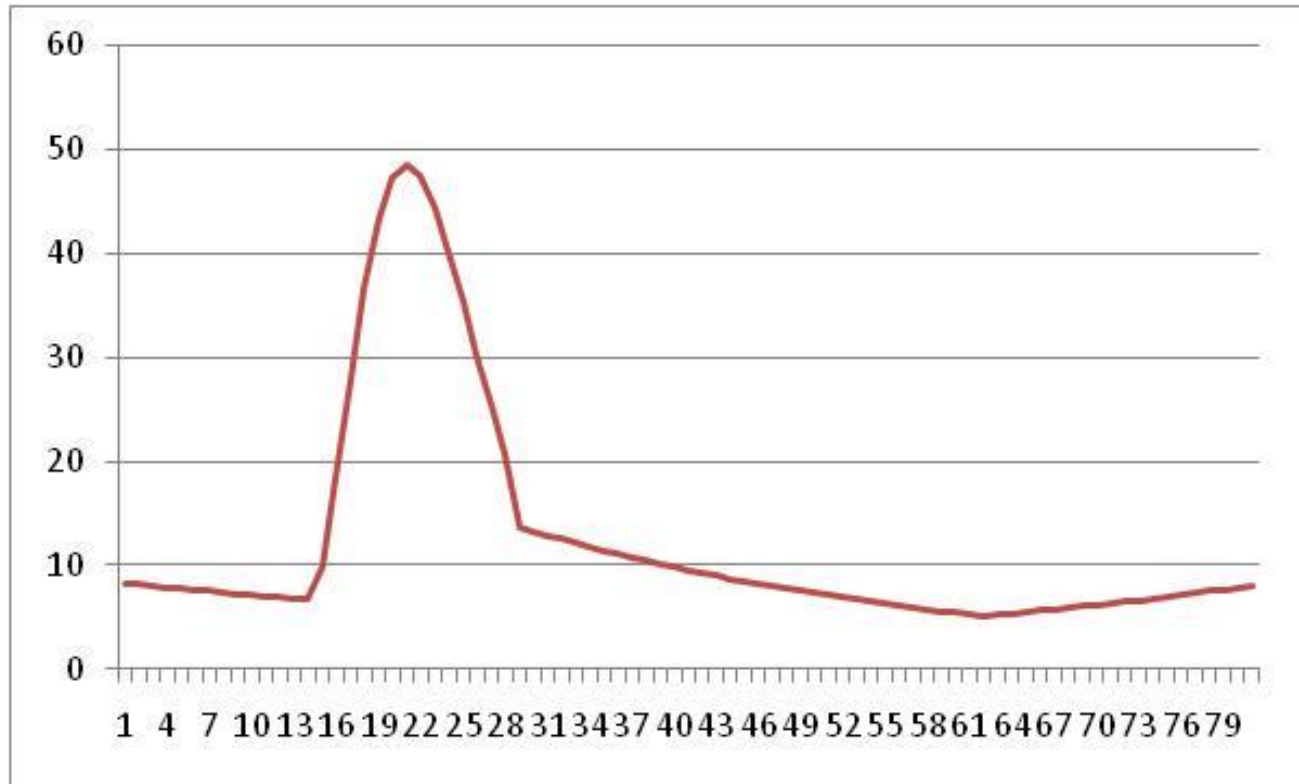
估计结果比较

- 与理论模式图比较
- 与其他来源数据比较
- 局限性分析



估计结果比较一与理论模式图比较

- 杨云彦(1992)提出的规格化年龄别迁移率分段拟合曲线



估计结果比较—与其他来源数据比较

- 省际净迁移人数比较
 - 估计某省年净迁出31.6万，四年半累计迁移142.2万；
 - 居住在外省、户口在A省的人口数350万；
居住在A省、户口在外省的人口数140万；
则净迁出210万。根据迁移时间，最近4年半内迁移人口约占2/3，由此计算4年半内净迁移人口为140万。



估计结果比较—与其他来源数据比较

- 乡城净迁移人数比较
 - 估计全国2011年乡城迁移人口数1520万；
 - 2011年底全国人口13.47亿，城镇人口数69079万，比2010年增加2101万人。
 - 2011年底全国自然增长人口约为644，估计城镇自然增长人口为300万。则乡城转移人口约为1800万，与估计值相比，相差20%。



估计结果比较—与其他来源数据

- A省乡城净迁移人数比较
 - 估计全省2011年乡城迁移人口数100万；
 - 实际2011年全省城镇人口数增加约100万；
 - 实际2011年全省自然增长人数约为46万，城镇约为20万，则乡城转移人口约为80万，与估计值相比，相差25%。



估计结果比较—局限性分析

1、缺少理论证明线性无偏估计

2、没考虑回迁因素

3、假设条件较多

欢迎交流

请批评指正

徐熙阳

北京信息控制研究所

北京市海淀区阜成路16号709室, 100048

010-68371109转610, 13683303865

bitmathxy@126.com

