

## 中国第四次人口普查出生 人口登记质量的初步检验

~~~~~ 涂 平 ~~~~~

1990年7月,中国进行了规模空前的第四次全国人口普查。国家统计局于1990年10月30日起陆续公布了第四次全国人口普查手工汇总数据,并在第四次全国人口普查10%计算机汇总数据公布后不久,于1991年底向国内有关单位提供了1%抽样数据带,及时为有关政府部门和科研单位提供了大量的宝贵资料。

第四次全国人口普查表共设了21个项目(西藏自治区使用的普查表仅19个项目),其中一项是育龄妇女1989年和1990年上半年的生育状况。该项的设立对于准确了解中国人口生育的水平、变化趋势和地区差异,正确评估各地区人口政策的执行情况和人口自然增长的态势有着十分重要的意义。然而,近年来由于众所周知的原因,出生瞒报和漏报现象比较严重,给普查的出生人口登记工作带来了很大的困难。本文根据第四次全国人口普查1%抽样数据的出生人口、死亡人口和尚存人口数据的一致性检验结果,对第四次全国人口普查出生人口登记的质量进行初步的评估,并在此基础上就如何进一步提高普查出生人口登记的质量提出一点粗浅的看法。

### 一、出生、死亡、尚存人口登记一致性检验的基本依据

第四次全国人口普查除调查了育龄妇女1989年和1990年上半年生育的分性别的活产子女数以外,还登记了普查时尚存人口和1989、1990年上半年死亡人口的出生年月。因此,可以根据1989年和1990年上半年出生队列的尚存人数和死亡人数推算出1989年和1990年上半年的预期出生人数,并与实际报告的出生人数相比较,从而对第四次全国人口普查的出生、死亡和尚存人口登记的质量进行初步的评估。在不考虑迁移的前提下,预期出生人数可由下式估算:

$$B'(t) = P(t) + D(t) \quad (1)$$

式中的 $t=1, 2, 3$ ,分别代表1989年上半年、1989年下半年和1990年上半年等3个不同时期; $B'(t)$ 为时期 $t$ 的预期出生人数; $P(t)$ 为时期 $t$ 出生的婴儿在普查时点的尚存人数; $D(t)$ 为该出生队列在普查时点以前死亡的人数。

普查得到的出生、死亡和尚存人口数据的内部一致性可用实际报告的出生人数 $B(t)$ 与根据上式估算的预期出生人数 $B'(t)$ 的比值 $K$ (以 $B'(t)$ 为100)表示, $K$ 接近100表示出生人口登记的结果与死亡人口和尚存人口登记的结果一致。一般说来,尚存人口的登记比较准确,而出生人口和死亡人口比较容易漏报。但是,在中国婴儿死亡率较低,死亡婴儿只占出生总数中极小的一部分(约在5%以内),因此除了个别死亡率极高、死亡漏报严重的地区以外,由于死亡漏报导致的出生人数的估算误差并不是很大。所以,依据上式估算的预期出生人数一般要比实际报告的出生人数更接近真实值。可以根据预期出生人数和报告出生人数的差值估算漏报出生人数,还可以根据报告出生人数和预期出生人数的比值粗略地推算出生人口漏报率 $R$ ,即

$$R \approx 100 - K \quad (2)$$

实际上由于尚存人口和死亡人口都可能存在不同程度的漏报,所以根据式(2)推算的出生人口漏报率一般都低于实际漏报率,其误差当死亡率较低时主要取决于尚存人口的漏报

率。在尚存人口和死亡人口漏报超过了出生人口漏报或离开亲生母亲单独迁出（由于寄养、收养等原因）的人数大于离开亲生母亲单独迁入人数的情况下，K可能大于100。由于0~1.5岁人口离开亲生母亲单独跨省份迁移的可能性很小，且无这方面的数据，故在估算预期出生人数时未予考虑。

因1989年的死亡人口有明显的漏报，而1990年的死亡人口登记比较准确，故在计算预期出生人数之前，按1990年死亡人数与其相应出生队列的出生人数之比，对1989年出生队列死于1989年的人数进行了调整，以减小由于1989年死亡人口严重漏报所带来的误差。调整前全国1989年出生队列的死亡人数为553 100，仅占该队列报告出生人数的22%；调整后的死亡人数为854 200，占该队列报告出生人数的34%。

## 二、出生、死亡和尚存人口数据一致性检验的结果

根据表1所列的报告出生人数推算的1989年和1990年上半年全国出生的总人数为35 405 300人，仅为根据这两个出生队列的尚存人数和死亡人数估算的全国同时期出生总人数（36 810 300人）的96.2%。由此可见，第四次全国人口普查登记的1989年和1990年出生队列的出生、死亡和尚存人数之间还不完全吻合，其中女性报告出生人数与预期出生人数的差异大于男性。全国1989年和1990年上半年出生的人口大约漏报了140.5万人（其中男婴42.8万、女婴97.7万），漏报率为3.8%。总的说来，1990年出生人口的漏报（4.9%）比1989年严重（3.3%）；女性出生人口的漏报（5.6%）比男性出生人口的漏报（2.2%）严重（见表1）。以上结果是在假设尚存人口和调整后的死亡人口准确的前提下得出的；如果考虑到尚存人口和死亡人口也可能存在不同程度的漏报，则出生人口的漏报率将更高。

由于1%抽样数据中各省的样本量不是很大，随机波动（主要是死亡人数的波动）较大，

表1 全国1989年和1990年上半年报告出生人数与  
预期出生人数之比和估计漏报出生人数（1%样本）

|                  | 合计      | 男       | 女       | 性别比   |
|------------------|---------|---------|---------|-------|
| 报告出生人数           |         |         |         |       |
| 1989年            | 250 229 | 133 696 | 116 533 | 114.7 |
| 1990上半年          | 103 824 | 55 958  | 47 866  | 116.9 |
| 1989年~1990年上半年   | 354 053 | 189 654 | 164 399 | 115.4 |
| 预期出生人数           |         |         |         |       |
| 1989年            | 258 885 | 136 103 | 122 782 | 110.8 |
| 1990年上半年         | 109 218 | 57 828  | 51 390  | 112.5 |
| 1989年~1990年上半年   | 368 103 | 193 931 | 174 172 | 111.3 |
| 估计漏报出生人数         |         |         |         |       |
| 1989年            | 8 656   | 2 407   | 6 249   | 38.5  |
| 1990年上半年         | 5 394   | 1 870   | 3 524   | 53.1  |
| 1989年~1990年上半年   | 14 050  | 4 277   | 9 773   | 43.8  |
| 报告出生人数/预期出生人数(%) |         |         |         |       |
| 1989年            | 96.7    | 98.2    | 94.9    | —     |
| 1990年上半年         | 95.1    | 96.8    | 93.1    | —     |
| 1989年~1990年上半年   | 96.2    | 97.8    | 94.4    | —     |

资料来源：根据第四次全国人口普查1%抽样数据带汇总计算。

故分省的出生、死亡和尚存人口数据之间的一致性检验结果不如全国的可靠，依此估算的分省的出生人口漏报数也会有较大的误差，仅供参考。图1和图2显示各省、自治区、直辖市1989年和1990年上半年出生队列的报告出生人数普遍低于预期出生人数，说明所有地区出生人口都有一定程度的漏报；但报告出生人数和预期出生人数的差值均未超出10%，且绝大部分地区都在5%以内。这

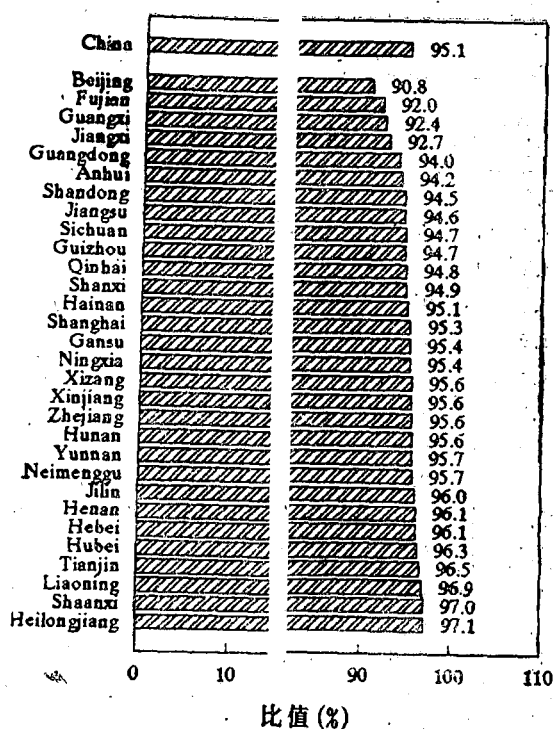


图1 各省、自治区、直辖市1989年报告出生人数与预期出生人数之比

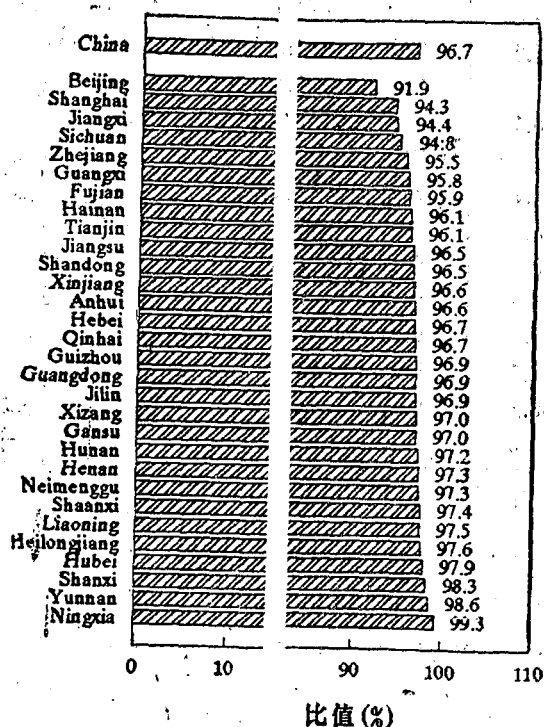


图2 各省、自治区、直辖市1990年上半年报告出生人数与预期出生人数之比

说明相对于尚存人口来说出生人口的漏报幅度并不很大。绝大部分省、自治区、直辖市的女性出生、死亡和尚存人口数据之间的一致性比男性差(见表2)。1990年上半年出生队列出生、死亡和尚存人口数据之间的一致性比1989年出生队列差,即近期出生婴儿较早期出生婴儿容易被漏报和瞒报。

### 三、几点思考

根据笔者比较保守的估算,第四次全国人口普查1989年和1990年上半年出生人口的漏报率约为3.8%。这可使全国的粗出生率比实际水平低0.8个千分点左右,使总和生育率比实际水平低0.1左右(假设母亲年龄不同的出生婴儿漏报无显著的差别),但并不会影响对当前中国人口总的生育水平、变化趋势和人口政策执行情况的基本估计。加拿大1986年人口普查总人口的净漏报率就达3%,美国1960、1970和1980年人口普查总人口的净漏报率也分别达3.3%、2.9%和1.4%(Robey, 1989)。在中国这样一个幅员辽阔、经济发展水平还比较落后的11亿多人口的大国,存在诸多不利因素,出生人口仅漏报4%左右,这是一个了不起的成就。但是,由于中国每年出生人口的基数很大,由此造成的出生人数的绝对误差还是相当可观的。不同地区、人群之间漏报率的差异还可能会给生育率的比较研究带来一定的困难。中国第四次全国人口普查出生人口的漏报幅度虽然不大,但绝大部分地区都有不同程度的漏报,具有相当的普遍性。因此,我们必须加强对有关出生登记质量的评估、出生漏报原因和漏报途径的分析以及防止漏报的对策的研究。

除个别地区以外,女婴的漏报比男婴严重。若扣除漏报因素,中国1989年和1990年上半年出生婴儿的性别比将由

表2 全国和各省、自治区、直辖市1989年和1990年上半年  
报告出生人数与预期出生人数之比 (%)

| 地区  | 1989年 |       |      | 1990年上半年 |      |      |
|-----|-------|-------|------|----------|------|------|
|     | 合计    | 男     | 女    | 合计       | 男    | 女    |
| 全国  | 96.7  | 98.2  | 94.9 | 95.1     | 96.8 | 93.1 |
| 北京  | 91.9  | 91.0  | 92.9 | 90.8     | 91.9 | 89.8 |
| 天津  | 96.1  | 97.7  | 94.4 | 96.5     | 97.5 | 95.5 |
| 河北  | 96.7  | 97.7  | 95.5 | 96.1     | 98.1 | 93.9 |
| 山西  | 98.3  | 99.4  | 97.2 | 94.9     | 96.1 | 93.5 |
| 内蒙古 | 97.3  | 98.6  | 96.0 | 95.7     | 98.2 | 92.8 |
| 辽宁  | 97.5  | 98.0  | 96.9 | 96.9     | 98.5 | 95.2 |
| 吉林  | 96.9  | 97.1  | 96.7 | 96.0     | 96.4 | 95.5 |
| 黑龙江 | 97.6  | 97.4  | 97.9 | 97.1     | 97.3 | 96.9 |
| 上海  | 94.3  | 94.1  | 94.5 | 95.3     | 99.7 | 90.9 |
| 江苏  | 96.5  | 98.7  | 93.8 | 94.6     | 96.7 | 92.2 |
| 浙江  | 95.5  | 97.2  | 93.7 | 95.6     | 96.4 | 94.5 |
| 安徽  | 96.6  | 99.0  | 94.0 | 94.2     | 96.3 | 91.9 |
| 福建  | 95.9  | 97.9  | 93.6 | 92.0     | 95.1 | 88.5 |
| 江西  | 94.4  | 98.4  | 90.3 | 92.7     | 95.9 | 89.4 |
| 山东  | 96.5  | 98.5  | 94.2 | 94.5     | 96.6 | 92.1 |
| 河南  | 97.3  | 99.1  | 95.1 | 96.1     | 98.2 | 93.6 |
| 湖北  | 97.9  | 98.9  | 96.9 | 96.3     | 97.0 | 95.6 |
| 湖南  | 97.2  | 98.3  | 95.9 | 95.6     | 97.1 | 94.0 |
| 广东  | 96.9  | 98.3  | 95.4 | 94.0     | 95.2 | 92.6 |
| 广西  | 95.8  | 98.0  | 93.2 | 92.4     | 93.4 | 91.1 |
| 海南  | 96.1  | 96.6  | 95.5 | 95.1     | 97.6 | 92.4 |
| 四川  | 94.8  | 96.9  | 92.5 | 94.7     | 97.3 | 91.9 |
| 贵州  | 96.9  | 98.2  | 95.5 | 94.7     | 96.5 | 93.0 |
| 云南  | 98.6  | 100.2 | 97.0 | 95.7     | 96.6 | 94.6 |
| 西藏  | 97.0  | 94.6  | 99.5 | 95.6     | 95.4 | 95.8 |
| 陕西  | 97.4  | 98.3  | 96.5 | 97.0     | 97.7 | 96.2 |
| 甘肃  | 97.0  | 98.7  | 95.2 | 95.4     | 97.0 | 93.5 |
| 青海  | 96.7  | 96.4  | 97.1 | 94.8     | 97.3 | 92.3 |
| 宁夏  | 99.3  | 100.2 | 98.3 | 95.4     | 96.0 | 94.8 |
| 新疆  | 96.6  | 97.2  | 95.8 | 95.6     | 97.3 | 93.9 |

资料来源：根据第四次全国人口普查1%抽样数据带汇总计算。

原来的114.7和116.9 (根据1%抽样数据带汇总的分性别报告中出生人数计算。它明显高于根据国家统计局的10%汇总资料计算的111.3和111.7) 降至110.8和112.5。由此可见,中国出生婴儿性别比偏高部分是由于出生人口(主要是女婴)的漏报所致,但在考虑了目前能够比较有把握地估算出的漏报出生人数后,出生婴儿性别比仍偏高。这一点应当引起我们足够的重视。

本文所采用的出生、死亡和尚存人口数据之间的一致性检验方法简单,不需要大量的人力、物力和时间,若让经过适当培训的普查员和数据录入员在普查登记和数据录入时以户为单位进行,那么可以在入户登记时和数据录入时发现每户的出生、死亡和尚存人数

之间的矛盾,便于及时查明原因并给予适当的更正。建议今后在进行普查或类似的人口抽样调查时,加强对调查人员进行简单的一致性检验的培训,并对所用数据录入软件增设某些一致性检验的功能;在进行登记质量检查时,除了考虑抽取部分样本复查外,也可考虑进行必要的有关数据的一致性检验。

#### 参考文献

- ① 国务院人口普查办公室:《中华人民共和国国家统计局关于一九九〇年人口普查主要数据的公报》。
- ② 国务院人口普查办公室和国家统计局人口司:《中国1990年人口普查10%抽样资料》,中国统计出版社,1991。
- ③ Robey, B. (1989): "The 1990 U.S. census: How good is good enough?" 《Asian and Pacific Population Forum》1989年第3期,第1~8页。

(本文责任编辑:徐莉)(作者工作单位:北京大学人口研究所)