

第四次全国人口普查人口性别 年龄结构的初步检验

乔晓春

人口性别年龄结构是人口现象的动态和静态的综合反映,它既是出生、死亡运动的结果,又是人口未来发展的决定性因素。在对第四次人口普查10%机器汇总资料进行分析以前,检验和评价普查汇总资料的准确性,特别是检验能够综合反映人口过程和人口状况的性别年龄结构,具有很重要的意义。

对人口性别年龄结构的检验主要有两种方法:一种方法是进行性别和年龄的一致性检验,即利用本次普查与上次普查资料的比较,以及本次普查各资料之间的比较来检验本次普查的性别年龄资料与有关的资料是否存在一致性或合理性;另一种方法是利用检

验指数检查性别年龄结构的偏好程度。下面就依据这两种方法,对中国第四次全国人口普查10%机器汇总资料中人口性别年龄结构数据进行初步的检验。

一、性别结构的一致性检验

如果说1964年和1982年分年龄的性别结构一致性还不是很强的话,那么可能存在下面两种原因:一种原因是两次普查时间间隔过长,这期间两性死亡率差异的变化比较复杂,在不同年龄段上的波动也比较大,变化不规则;另一种原因是第二次普查时间距今比较早,普查的科学性和现代化程度还不高,很可能影响到数据的准确性。在利用1982年

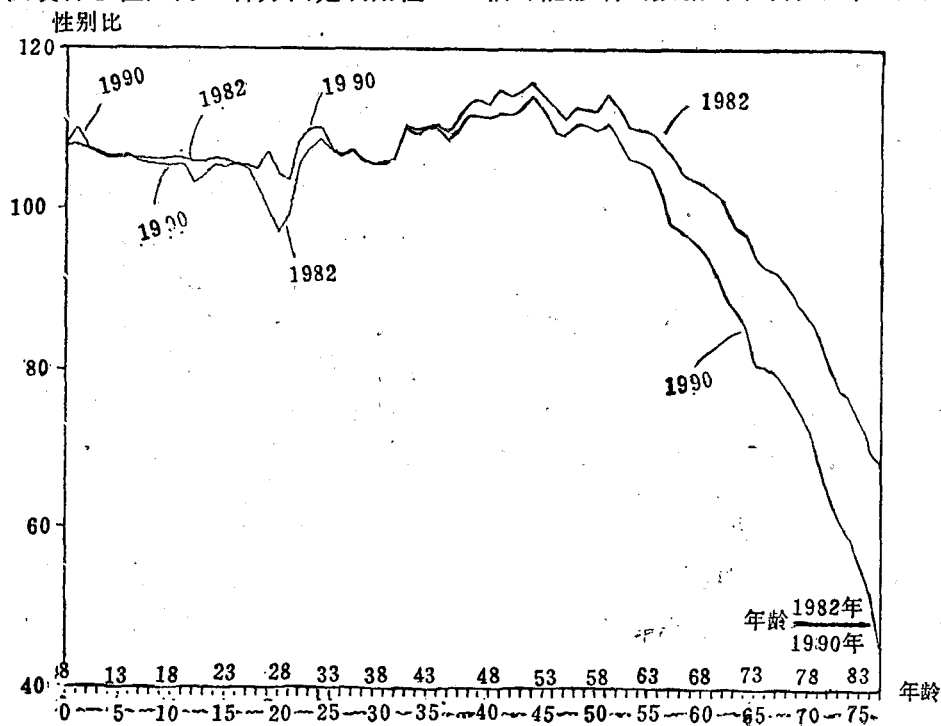


图1 1982/1990年同批人性别比比较

和本次人口普查的性别结构资料进行比较时,就会克服上述的不足。

我们将1982年普查得到的X岁人口性别比与1990年普查时相应的X+8岁人口(为出生同批人)性别比进行比较(见图1),可以明显看出:1982年与1990年同批人性别比具有很好的一致性。二者不相一致的地方主要是由于两资料中都不含现役军人的缘故。而低年龄和高年龄段性别比不仅一致,而且曲线的波动也是一样的。从中反映出人口性别结构资料,特别是25岁以后的人口性别比资料是准确的。

作为同批人性别比的比较,存在着两种可能:一种可能是普查间存活率女性高于男性,那么性别比应该是下降的,即性别比差异为正值;另一种情况是普查间存活率男性高于女性,结果是性别比提高,性别比差异为负值。两性存活率差异越大,性别比差异的绝对值也越大;反之越小。极端的情况是:如果普查间两性存活率相等,性别比也就不会发生变化(假设不存在国际净迁移)。从计算结果(计算结果略)中可以看出:1982年0~6岁人口到1990年8~14岁时,性别比是提高了;从7岁开始性别比出现下降的趋势,而且下降很快。这种下降主要是由于1990年数据未包括现役军人,从而18~23岁人口性别比明显偏低,导致同批人性别比差异为正值;而1982年16岁以后人口的性别比差异又转为负值,而且数值不断增大。这是由于1982年数据未包括现役军人所致;从31岁开始人口的性别比差异变为正值,且二者差异不断加大,这是正常的情况。

二、人口年龄结构的一致性检验

对人口年龄结构的一致性检验,通常是利用观察两次普查间存活率的合理性来进行的。下面利用1982年分性别年龄的人口资料与1990年普查相应性别年龄的人口进行比较,计算出分性别的两次普查间的存活率(见表1和图2、3)。

同批人的人数应该是随着时间的推移而不断减少的。1982年年中X岁的人口活到1990年年中正好是X+8岁。1990年X+8岁的人口除以1982年X岁的人口就得到两次普查间的存活率(见表1)。

根据数据所反映出的规律,我们把两次普查间的存活率按同批人分成下面几段:第一段是1982年0~7岁即1990年8~15岁的人口,这批人在两次普查间的存活率过高,很多年龄是大于1的。这实际上是一种很不正常的情况,对此问题还有待做进一步分析。第二段是1982年8~13岁即1990年16~21岁的人口,这批人的存活率较低,特别是男性,原因是1990年16~21岁人口相对较少。女性虽然在个别年龄上也很低,但总的看影响不大。第三段是1982年14~23岁即1990年22~31岁的人口,这部分人存活率较高,特别是男性,主要原因是1982年的年龄状况数据也未包括现役军人,因此1982年时这批人相对较少,存活率必然偏高。女性也有所偏高,但不严重。第四段就是1982年24岁以上人口即1990年32岁以上人口,总的看这部分人的存活率是比较正常的。

虽然第四段比较正常,但是人们会立即发现,1982年32岁即1990年40岁的人口存活率高于1,从而会想到:1990年时以“0”或“5”结尾的年龄上是否存在着普查间存活率偏高的情况呢?仔细观察果真如此。可以看出1990年32~57岁之间以“0”和“5”结尾的年龄上,无论男女,人口存活率都高于与它相邻年龄人口的存活率。而在59岁(而不是在60岁)也存在着这一现象。以“0”和“5”结尾年龄上人口存活率偏高的原因是1982年时人们在相应年龄上不存在偏好。比如,1990年40岁的人在1982年时为32岁,以“2”结尾的年龄自然偏好不严重。而在1990年时,一些以“0”或“5”结尾年龄附近的人,有意将自己不是以“0”或“5”结尾的年龄错报成以“0”或“5”结尾的年龄。59

表1

1982/1990年分年龄的两次普查间存活率

1982年 年龄	男性	女性	1990年 年龄	1982年 年龄	男性	女性	1990年 年龄	1982年 年龄	男性	女性	1990年 年龄
0	1.059	1.058	8	26	0.984	0.982	34	52	0.905	0.936	60
1	1.040	1.031	9	27	0.986	0.984	35	53	0.883	0.916	61
2	1.054	1.052	10	28	0.986	0.984	36	54	0.880	0.913	62
3	0.988	0.986	11	29	0.985	0.984	37	55	0.864	0.900	63
4	1.017	1.014	12	30	0.974	0.970	38	56	0.848	0.891	64
5	1.000	0.998	13	31	0.992	0.997	39	57	0.839	0.915	65
6	1.004	1.002	14	32	1.005	1.007	40	58	0.812	0.871	66
7	0.995	0.999	15	33	0.989	0.994	41	59	0.797	0.859	67
8	0.977	0.982	16	34	0.988	0.991	42	60	0.761	0.828	68
9	0.978	0.984	17	35	0.964	0.969	43	61	0.764	0.838	69
10	0.969	0.977	18	36	0.973	0.983	44	62	0.738	0.831	70
11	0.949	0.956	19	37	0.984	0.994	45	63	0.712	0.796	71
12	0.972	0.996	20	38	0.968	0.980	46	64	0.692	0.786	72
13	0.955	0.971	21	39	0.967	0.986	47	65	0.653	0.759	73
14	0.997	1.005	22	40	0.957	0.970	48	66	0.630	0.726	74
15	0.100	1.008	23	41	0.946	0.971	49	67	0.605	0.706	75
16	0.982	0.982	24	42	0.976	0.996	50	68	0.565	0.663	76
17	1.006	1.003	25	43	0.947	0.962	51	69	0.545	0.645	77
18	1.017	0.995	26	44	0.957	0.972	52	70	0.516	0.619	78
19	1.021	0.952	27	45	0.941	0.955	53	71	0.469	0.580	79
20	1.055	0.982	28	46	0.931	0.954	54	72	0.450	0.564	80
21	1.099	1.051	29	47	0.939	0.957	55	73	0.392	0.498	81
22	1.021	0.993	30	48	0.924	0.940	56	74	0.374	0.481	82
23	1.017	0.994	31	49	0.910	0.926	57	75	0.342	0.452	83
24	0.985	0.972	32	50	0.907	0.926	58	76	0.332	0.417	84
25	0.973	0.970	33	51	0.928	0.956	59	77	1.050	1.611	85

资料来源：查瑞传《人口普查资料分析技术》，中国人口出版社，第31、34、41页。

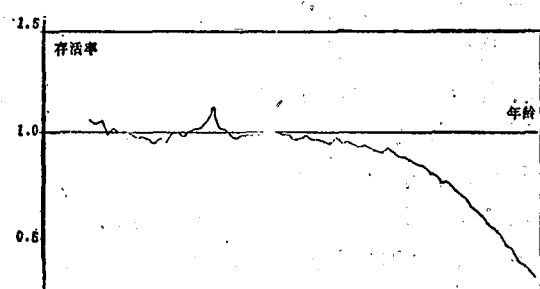


图2 1982/1990年男性分年龄
普查间存活率曲线

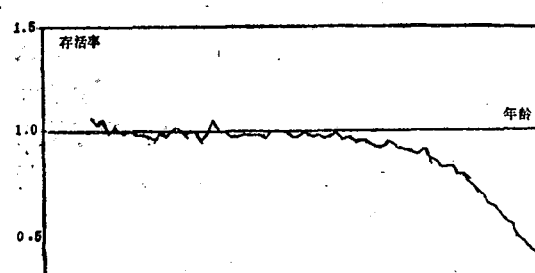


图3 1982/1990年女性分年龄
普查间存活率曲线

岁的存活率偏高的现象，很明显说明人们在有意回避退休年龄。由此可以说明，中国人口年龄登记也在一定程度上存在着年龄偏好现象。这一点必须引起注意。

由此我们会想到另一种相应的情况：1982年普查是否也存在着这种年龄偏好？这一点在普查以后的分析上已做了定论，即不存在年龄偏好。然而从现在的情况看，事实

并不如此。如果1982年普查以“0”和“5”结尾的年龄上也存在一定的年龄偏好的话，而在1990年相应的年龄上（即以“8”和“3”结尾的年龄）不存在年龄偏好，那么就必然会出现同批人普查间存活率较低的现象。对数据进行观察会发现，这也是一种事实，只是这种事实在高年龄段上不明显，这是由于高龄段相邻两个年龄组存活率下降过快，从而掩盖了这种差异。

进一步可考察这种年龄偏好在两性上是否存在差异。如果存在，必然在“0”或“5”结尾的年龄上性别比都偏高或偏低。然而事实上这种情况是不存在的，说明年龄偏好在

性别上不存在明显的差异。

以上用了一种新的方法考察了中国人口年龄的偏好现象，结论是偏好现象存在，但存在的程度并没有给出，这还有待进一步研究。

三、性别年龄结构的指数检验结果

人口性别年龄结构的各种检验指数实际上都有一定的假设，而正是由于假设使这些指数在应用和解释上具有很大的局限性，尤其对中国人口的性别年龄结构来说，客观上的波动十分严重，恰恰与假定相悖。好在指数检验的结果可以得出“准确性比较好”的结论，否则很难对此给出定论。

表2 中国历次普查三种指数检验结果

普查时间 (年)	惠 普 尔 指 数			迈 叶 斯 指 数			联合国 综合指数
	合计	男性	女性	合计	男性	女性	
1953	104.22	104.06	104.40	1.16	1.10	1.26	18.44
1964	103.35	103.10	103.16	0.41	0.42	0.46	18.39
1982	103.47	103.32	103.63	3.00	3.02	3.00	26.78
1987 (1%)	99.53	99.44	99.63	1.86	1.90	1.91	27.34
1990	100.94	101.05	100.82	2.29	2.22	2.36	23.31

资料来源：同表1。

同以往对普查资料的检验一样，对本次普查人口性别年龄结构的惠普尔指数和迈叶斯指数检验都比较理想。男女合计的惠普尔指数值为100.94，好于历次普查的结果（见表2）。迈叶斯指数的值也是比较好的，合计值为2.29。都属于理想的数值范围。在利用联合国综合指数检验时，发现1990年的结果低于1982年和1987年数据检验的结果（分别为26.78和27.34），1990年为23.31，这一数值实际上属于“资料不够准确”的范围，可以明显看出这种判断是不符合实际的，是由于实际与假定不符造成的。

前面我们曾检验出中国人口在一定程度上还存在着年龄偏好现象，为什么用惠普尔

指数和迈叶斯指数却检验不出来呢？其原因是中国人口年龄结构波动过大，而且这种波动在年龄上又没有规律，从而过大的、不规则的波动掩盖了过小的、规则性的偏好。惠普尔指数和迈叶斯指数所能检验出的不只是年龄的不规则，还必须有规律，因为不规则的年龄结构可能在检验时得出好的结果。而联合国综合指数实际上检验的只是年龄结构是否规则，而不必有规律。这就是为什么中国历次普查的惠普尔和迈叶斯指数值都很低，而联合国综合指数却较高的原因。

（本文责任编辑：高春燕）

（作者工作单位：中国人民大学人口研究所）