

试论年龄结构检验指数对 普查资料检验的有效性问题

乔晓春

国际上使用的人口年龄结构检验指数通常有如下几种:年龄偏好指数(Indexes of Age Preference),惠普尔指数(Whipple's Index),迈叶斯指数(Myers' Index),年龄准确性指数(Age-accuracy Index)和联合国年龄性别准确性指数(United Nations Age-Sex Accuracy Index),简称联合国综合指数。在以上这些指数中,惠普尔指数、迈叶斯指数和联合国综合指数在我国人口普查的人口年龄资料的检验中经常使用。

在使用中,人们通常只根据它们的计算过程,计算出各指数的结果,并根据给定的判断标准直接进行判断,以此得出结论。从以往的实践中我们发现,如果不了解指数的构造思想和指数的局限性,直接根据标准进行判断,很可能会得出错误的结论。

一、年龄结构检验指数的简单回顾

各指数对年龄结构的检验角度不同,指数的构造方法也不同。相应于每个指数人们都给出了指数的判断标准,并以此标准来判断人口年龄结构质量的好坏程度。

年龄偏好指数是对某一特定年龄的偏好程度进行检验。分析者可以根据以往的经验或某一特定的目的,选取某一年龄,并计算出这一年龄的偏好指数,最后对其结果进行判断。计算公式为:

$$IAP_a = \frac{P_a}{\frac{1}{5}(P_{a-2} + P_{a-1} + P_a + P_{a+1} + P_{a+2})} \times 100$$

这里 a 为年龄, P_a 为 a 岁人口数, IAP_a 为 a 岁的年龄偏好指数。

指数的可能取值范围为0—500。由于人们是为了判断其年龄是否存在偏好,因此,在这个假定下,正常的取值范围应该在100—500之间。如果 IAP_a 的值在100—110之间,则认为在该年龄上不存在堆积;如果 IAP_a 在110—130之间,说明存在堆积现象,但不严重;当 IAP_a 的值大于130,说明堆积严重。

年龄偏好指数只能检验一个特定年龄,即使某一年龄存在偏好,并不能说明其它年龄也存在同样程度的偏好,因而无法由此判断多个年龄或总体年龄结构的好坏。

从国际上看,一般在以“0”和“5”结尾的年龄上堆积最为严重,惠普尔指数正是为了检验一个年龄结构中所有以“0”和“5”结尾的年龄的偏好程度,并以此来推断整体年龄结构的质量,从而弥补了年龄准确性指数的不足。惠普尔指数通常用下面的公式计算:

$$WI = \frac{P_{25} + P_{30} + P_{35} + \dots + P_{55} + P_{60}}{\frac{1}{5}(P_{23} + P_{24} + \dots + P_{60} + P_{61} + P_{62})} \times 100$$

这里 WI 代表惠普尔指数。

该指数的可能取值范围为0—500,正常取值范围为100—500,指数的值越是接近100,说明资料的质量越好;如果其值在100—110之间说明无明显偏好,在110—130之间说明存在年龄偏好,但可以接受;大于130说明偏好严重,不可接受。由于以“0”和“5”结尾的年龄通常是偏好最严重的年龄,以此来反映整体年龄结构质量是可以让人接受的。

虽然如此,人们还是寻求通过对各结尾年龄进行检验来衡量总体年龄结构的质量。对于一个正常人口,如果人口向某一结尾年龄或某几个结尾年龄偏好的话,在这些年龄上必然出现人口的堆积,那么在其它年龄上同时也会出现人口减少的现象,这样就会导致年龄结构分布不均匀。对于这种不均匀程度的判断正是迈叶斯指数的构造思想。迈叶斯指数在计算上的特点是构造了混合和,计算采用了下面的公式:

$$B_i = (i + 1) \sum_{k=1}^8 P_{ki} + (9 - i) \sum_{k=2}^8 P_{ki} \quad (i = 0, 1, \dots, 9)$$
$$MI = \frac{1}{2} \sum \left| \frac{B_i}{\sum B_i} \times 100 - 10 \right|$$

这里 i 为年龄尾数, B_i 为以 i 年龄结尾的人数之和(在10岁—89岁之间), MI 为迈叶斯指数,其取值范围为0—90,其值越接近0,说明资料质量越好,当 MI 在0—5之间时,说明资料质量比较好,在5—10之间说明质量不是很好,但可以接受。如果 MI 大于10说明质量非常不好,不能接受。

前面几个指数都是对按一岁分组的年龄进行检验,人口学中的年龄也经常按五岁进行分组,对按五岁分组的年龄结构进行检验通常利用年龄准确性指数和联合国综合指数。年龄准确性指数(AAI)是在计算各年龄比(${}_5AR_a = {}_5P_a / \frac{1}{2}({}_5P_{a-5} + {}_5P_a + {}_5P_{a+5})$)的基础上,求平均离差。如果 AAI 在0—3之间说明数据质量较好,在3—7之间可以接受,大于7则不可接受。联合国综合指数(UNI)比年龄准确性指数进了一大步,它既可以考虑性别准确性,也可分性别考虑年龄准确性,但在年龄方面,计算的思想与年龄准确性指数一样。其判断标准为:如果 UNI 在0—20之间说明资料是准确的,如果在20—40之间说明资料不够准确,但可以接受。若 UNI 大于40,说明资料很不准确,不能接受。

二、问题的提出

在教学和对指数的使用中,我们发现了以下一些问题:

1. 对辽宁省1987年1%抽样调查的25岁人口进行年龄偏好指数检验,其结果为57.37。从这一结果看,似乎1987年人口调查中对25岁年龄存在着严重的回避现象,很明显,这种认识是不对的。

2. 利用各类指数对我国历次人口普查的年龄结构进行检验(见表1),对各指数检验的结果进行判断,其结论并不完全一致。利用惠普尔指数和迈叶斯指数都可以做出我国第三次和第四次人口普查的年龄结构无偏好现象的判断。但是利用年龄准确性指数和联合国综合指数检验,1982年分别为5.76和26.78,1990年分别为4.62和23.31,根据上述判断标准,其结论应该是:我国人口年龄结构资料不够准确,二者是矛盾的。

3. 对1990年第四次人口普查的年龄偏好进行惠普尔指数检验,其值为100.94,男性为101.05,女性为100.82,说明1990年人口年龄资料的质量是相当好的,而且还好于历次人口普查。用迈叶斯指数检验却发现:1990年普查的年龄资料质量虽然也很好,且好于1982年第三次

人口普查,但不如1953年和1964年的两次人口普查。两个指数检验的结论显然是矛盾的。

表1 我国历次普查(调查)三种指数检验结果

普查年代	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	3.03	3.00	26.78
1987	99.53	99.46	99.63	1.86	1.90	1.91	27.34
1990	100.94	101.05	100.82	2.07	2.06	2.07	23.31

资料来源:乔晓春:“第四次全国人口普查人口性别年龄结构的初步检验”,高校第四次全国人口普查资料分析学术讨论会论文,1991.11.北京。

4. 从国际上的一般情况看:以“0”和“5”结尾的年龄堆积程度最为严重,而其它年龄的堆积则不很明显。因此在正常情况下,惠普尔指数检验的结果应该大于或等于100,而对全国1987年1%抽样调查数据的惠普尔指数检验结果却小于100。自然的解释是:似乎在年龄登记中存在回避“0”和“5”年龄的情况,这实际上是不可能的。这与我们要检验的目的相悖。

三、对年龄偏好的理解

要想找出问题症结所在,必须进一步明确进行各类指数检验的目的是什么。人们都知道指数检验的目的是为了检验年龄结构上是否存在偏好(或堆积),以及偏好的程度。那么首先要明确什么是年龄偏好。年龄偏好指的是被调查者在申报年龄时有意申报自己所喜爱的年龄,而不是真实年龄,从而在年龄结构中出現较一致的、明显的、有规律的年龄“堆积”现象。从定义中可以看出:年龄偏好有两个重要特征,一是申报人有意误报年龄,而不报真实年龄,二是由于误报在年龄结构上表现出一致的、有规律的堆积。

年龄偏好这两个特征缺一不可。广义的年龄偏好可以指任何有意的年龄误报,比如:为了升学或招工有意改变真实年龄,从而使自己的年龄提高或降低。如果这种误报在总体上并不倾向于某一特定年龄,也不属于本文所讲的可用指数检验的年龄偏好;另一方面,如果在年龄结构上表现出一致的“堆积”现象,但这种现象并不是由于有意的年龄误报,而是由于特定年代出生的“婴儿堆积”造成的,这也不属于年龄偏好的范畴。

我们讲的年龄偏好如上述定义,而现实使用的一系列年龄检验指数,主观目的也是为了检验这样一种年龄偏好。但在我国客观上是否可以较准确地实现这样的目的呢?回答是否定的。

四、对问题原因的一般解释

检验人口年龄结构是否存在偏好应该有一个标准。这个标准当然就是不存在偏好的人口年龄结构。人们用这个标准去衡量由调查得来的人口年龄结构,如果调查的年龄结构与标准年龄结构相差很大,说明年龄偏好现象很严重,如果相差的很小,说明偏好不大或不存在偏好。

年龄性别的各种检验指数实际上都是在利用某一标准年龄结构对特定的年龄结构进行检验。虽然在很多教科书中并没有指出这一点,但在各种指数中实际上都隐含着这样一种假设,即标准的年龄结构是一个分布均匀的结构,从而以一个分布均匀的结构作为标准,判断调查得到的年龄结构是否存在偏好,以及偏好的程度。

下面的关系实际上表示的就是指数检验的实际含义。

标准年龄结构 $\frac{\text{比较差异}}{\text{调查年龄结构}}$

通过将标准年龄结构与调查年龄结构进行比较,找出其绝对差异值的大小,即为调查年龄结构的偏好程度。

还必须注意各检验指数对标准年龄结构的假设并不完全一致。年龄偏好指数和惠普尔指数均假设标准年龄结构是一个分布均匀的年龄结构,迈叶斯指数实际上假定标准年龄结构是一个特定的静止人口的年龄结构。正是由于罗伯特·迈叶斯于1940年利用了这样的一个特定结构作为标准,才构造了通过特殊加权来计算混合和这样一个特定的方法,年龄准确性指数和联合国综合指数则是假定按5岁分组的年龄结构是分布均匀的。

各种年龄检验指数所存在的最大问题是假定标准年龄结构就是真实的年龄结构,或者说假设真实的年龄是一个分布均匀的年龄结构。这种假设可能适用于世界上很多国家,而恰恰与中国的情况相差甚远。因为中国现实的人口年龄结构是一个波动很大的结构,即真实年龄结构与标准年龄结构之间存在较大差异。

历史上中国出生率曾出现过大起大落的现象,出生人数在相邻几年的波动也很大。比如:中国1961年出生人口只有1189万人,1962年为2460万人,而1963年高达2954万人,这种波动必然在年龄结构上反映出来。可见,中国的真实年龄结构是一种非均匀分布的年龄结构,而中国的调查年龄结构与真实年龄结构之间的差异可以说并不是很大。中国90%以上的人都是汉族,都有记忆年龄的办法和习惯,年龄的概念十分明确,虽然存在一定的有意误报,但由于记忆模糊误报的可能性很小。所以,中国的情况是在标准年龄结构与真实年龄结构之间差异较大,而调查年龄结构与真实年龄结构之间差异较小。这时如假定标准年龄结构就等于真实年龄结构,并用标准(分布均匀的)结构直接检验调查的年龄结构就很不合适。

五、对检验指数的选取

从前面分析可以看出,在一般意义上说,在中国使用人口年龄结构检验指数有一定的局限性,但这里不是说前面提到的各种指数都不能用于对中国人口年龄结构的检验。通过分析发现:不同的指数所能达到的目的是不同的,或者说检验的角度不同。对于中国这样一个特殊的年龄结构,相对来说,有些指数还是可用的。

年龄偏好指数是为了检验一个特定年龄的偏好问题,通过检验,若被检验年龄人口明显多于相邻年龄人口,则认为在这个年龄上存在年龄偏好。但对中国这样波动性很大的年龄结构,如果某年龄人口明显多于相邻年龄的人口,并不能确定这是由于这批人在出生时数量就明显多于相邻年份,还是由于调查时有很多人在这个年龄上有明显的偏好而误报该年龄。如果年龄偏好指数能证明被检验人口与相邻年龄人口在出生时没有明显的波动,而调查人口的年龄出现波动,则检验才是有意义的,否则就不可用。

惠普尔指数是为了检验最易出现堆积的以“0”和“5”结尾的年龄。能够给出这种判断的前提条件是以“0”和“5”年龄结尾的人口如果多,并不是由于相应的人口在出生时就多。在中国这一前提条件似乎是可以成立的。因为以“0”和“5”岁结尾的这些人中可能会有某一批或某几批人在以往出生时人数比较多,但不存在每隔五年出现一次出生人数多的现象。事实上可能是某些年出生人数多,五年后的某些年出生人数可能又会少,从总体上看多和少在一定程度上可相互抵消。可见,虽然年龄结构的波动可能对惠普尔指数的检验结果产生一定的影响,但影响并不会很大。这样利用惠普尔指数检验我国人口年龄结构的准确性还是可行的,其结果在一定程度上也是可信的。

迈叶斯指数是通过检验按同一尾数分组的人口分布均匀程度,来判断年龄偏好的。为了消

除各尾数年龄人口受死亡率影响而出现的梯形结构,罗伯特·迈叶斯1940年根据特殊的稳定人口年龄结构构造了“混合和”,使按尾数分布的年龄结构由梯形转化为矩形,并以此作为判断年龄结构是否正常的标准。迈叶斯指数将按同一结尾年龄分组的人数转化为百分比分布,在总体为100且分布均匀的情况下,每一年龄结尾的人口应占比例为10%,若某一年龄结尾的人数多,必然出现其它一些年龄结尾的人数少的情况,各年龄人口与10%的绝对离差越大,说明迈叶斯指数的值也会越大。从迈叶斯指数的构造上看,指数值大的必要条件是某一个或几个结尾的年龄人数较多。而每十年恰好都出现出生人数多的情况并不多见,那么出现某一结尾年龄人数多,就可以认为是有意误报造成的。以上说明用迈叶斯指数判断年龄偏好也是可行的。

年龄准确性指数的目的是对按五岁分组的年龄结构的准确性问题进行检验,它实际上是计算按五岁分组的人口的年龄比。在指数中实际上假定正常的人口年龄结构是分布均匀的,即年龄比近似于100,离100越远(或大或小)都会导致指数的值越大。从年龄准确性指数的构造方法上看,该指数似乎并不是在检验由人们对某一年龄存在偏好,因为人们只能对某一特定年龄存在偏好,而有意向某五岁年龄组偏好的情况并不存在。既然按五岁组间的偏好并不存在,那么年龄准确性指数的高低也就不是在检验年龄偏好,而是判断按五岁分组年龄结构的均匀或波动程度,这似乎背离了我们检验年龄结构的本意,当然也就不能用于解释调查质量的好坏。

联合国综合指数与年龄准确性指数存在同样问题,也不能用于判断年龄结构的登记质量。

六、对前面提出问题的具体解释

根据对年龄结构检验指数的认识,我们可以对本文第二部分所提出的问题作具体解释,由此可以进一步加深对各类指数的理解。

1. 为什么对辽宁省1987年1%抽样调查的25岁人口进行年龄偏好指数的检验其结果只有57.37呢?原因是年龄波动本身造成的,而并不能认为人们在登记时对25岁年龄存在回避现象。1961年下半年到1962年上半年是辽宁省最严重的生育低谷时期,这段时间出生的婴儿到1987年中正好是25周岁,这样25岁的人数必然要少。按照年龄偏好指数进行计算其值必然很小,这是一种客观存在,而不是主观上的年龄误报,可是年龄偏好指数并未起到这样的作用。

2. 为什么利用惠普尔指数和迈叶斯指数进行检验的我国历次普查的年龄结构质量都很好,而利用联合国综合指数检验则认为不好呢?这是因为中国人口年龄结构的波动性很大,而且很不规则,而只要不规则,联合国综合指数就能反映出来。不规则的程度越大,指数值也越大,而惠普尔和迈叶斯指数所检验的不是实际年龄结构的不规则,而是每隔五年或每隔十年存在的那种周期性波动,而我国人口年龄结构的不规则,并不具有明显的周期性规律。因此惠普尔和迈叶斯指数的值就比较小,联合国综合指数值则比较大。

3. 为什么1990年人口普查的惠普尔指数和迈叶斯指数在与以往各次普查检验结果进行比较时存在矛盾?这是因为两个指数计算方法不同所致。对于1990年来说,1953年和1964年普查的惠普尔指数比较高,是由于在当时中国确实还存在对“0”和“5”结尾的年龄偏好,虽然程度并不高。而迈叶斯指数对“0”和“5”结尾年龄的检验只是整个指数检验的一部分,1953和1964年在年龄结构上还未出现明显的波动,这样指数值相对于1990年就比较低。这并不说明1990年资料的质量不如1953年和1964年好。

4. 为什么对1987年小普查资料的惠普尔指数检验,出现小于100的现象?如果从惠普尔指数构造和通常的经验看,这种现象似乎不可能出现。因为从一般情况看,逢“0”和“5”结尾的年龄堆积是最严重的,即使不向这一年龄偏好,也不会出现回避的问题。很明显惠普尔指数小于

100,不是有意误报造成的,只能是由于当时特定的年龄结构产生的,即在1987年调查时点上以“0”和“5”结尾年龄上的人口相对来说少于其它年龄结尾的人口。

七、对我国人口普查年龄数据质量的判断

综上所述,对于中国不规则的人口年龄结构,惠普尔指数和迈叶斯指数还是可用的。分析这两个指数对我国历次人口普查的检验结果会发现,从总体上看,人口年龄偏好现象在我国并不存在,即使在年龄结构没有明显波动的1964年和1953年,惠普尔和迈叶斯指数都很低,而且随着社会经济的发展和人们文化素质的提高,其偏好程度也会越来越低,惠普尔指数已经表现出了这种趋势。迈叶斯指数的提高主要是由于1982年以后的年龄结构波动造成的。

1990年人口普查我国的惠普尔指数只有100.94,可以说是相当低了,根据这一结果是否可以判断,我国1990年人口年龄登记不存在偏好,质量相当好呢?其实不然。

我们可以通过计算两次普查间的存活率,并比较二者一致程度来判断普查年龄资料的准确性。由于1982年23岁以下人口受现役军人人口年龄结构影响^①,因而只计算23岁到60岁各队列人口存活率。由计算结果可以发现:1990年32岁至57岁之间从“0”和“5”结尾年龄上的人口存活率,无论男女都高于与它相邻年龄人口存活率。在59岁(而不是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”和“5”结尾年龄附近的人,有意将自己的年龄错报成以“0”或“5”结尾的年龄。59岁人口存活率偏高的现象,很明显说明人们在有意回避退休年龄。由此可以说明:我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年人口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。从现在情况看,事实上并非如此。如果1982年人口普查以“0”和“5”结尾的年龄上也存在一定年龄偏好的话,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。其原因是由于高龄相邻两组存活率下降过快而掩盖了这种差异。

八、结论

中国人口年龄结构波动性在一定程度上限制了各种年龄偏好指数在中国的使用,由于惠普尔和迈叶斯指数只有在一定程度上受客观的年龄结构波动影响,从一般意义上说还都可以使用。但特定的年龄结构可能会导致惠普尔指数有时偏低,有时偏高。对1990年普查资料检验发现,惠普尔指数的值偏低,相对来看,迈叶斯指数的值目前还是比较合理的,而年龄偏好指数、年龄准确性指数和联合国综合指数在我国使用还是存在很大问题的。当然对于年龄堆积比较严重或人口年龄分布比较均匀的特定地区和特定人口来说,这些指数还是可用的。

^① 得不到按岁分组的现役军人年龄结构。

表2

1982—1990年普查间存活率

1982年		存活率		1990年		1982年		存活率		1990年	
年龄	男性	女性	年龄	年龄	男性	女性	年龄	年龄	男性	女性	年龄
23	1.0170	0.9941	31	42	0.9762	0.9958	50				
24	0.9845	0.9724	32	43	0.9473	0.9622	51				
25	0.9730	0.9703	33	44	0.9572	0.9724	52				
26	0.9839	0.9818	34	45	0.9413	0.9549	53				
27	0.9861	0.9837	35	46	0.9311	0.9539	54				
28	0.9858	0.9835	36	47	0.9393	0.9567	55				
29	0.9850	0.9837	37	48	0.9238	0.9398	56				
30	0.9737	0.9704	38	49	0.9104	0.9257	57				
31	0.9918	0.9965	39	50	0.9065	0.9264	58				
32	1.0050	1.0070	40	51	0.9282	0.9557	59				
33	0.9888	0.9940	41	52	0.9049	0.9358	60				
34	0.9880	0.9910	42	53	0.8834	0.9156	61				
35	0.9640	0.9690	43	54	0.8804	0.9134	62				
36	0.9727	0.9830	44	55	0.8640	0.9000	63				
37	0.9842	0.9936	45	56	0.8478	0.8905	64				
38	0.9682	0.9801	46	57	0.8391	0.9145	65				
39	0.9670	0.9861	47	58	0.8115	0.8709	66				
40	0.9569	0.9697	48	59	0.7965	0.8586	67				
41	0.9462	0.9711	49	60	0.7613	0.8281	68				

(作者工作单位:中国人民大学人口研究所)

中外动态

北京举行当代中国妇女地位 国际学术研讨会

[本刊讯]根据中国政府与联合国人口基金达成的协议,由中国社会科学院人口研究所主持的、以妇女地位的现状与变动趋势为中心议题的当代中国妇女地位国际学术研讨会1993年10月31日—11月4日在北京举行。中国计划生育协会会长宋平,以及中国社科院、国家计生委、全国妇联、国家统计局、国务院理论局、北京大学人口所的领导、联合国计划署的官员出席了会议的开幕式。参加研讨会的正式代表共计52人,其中中方代表35人,外方代表17人。会议共收到论文51篇。与会者根据历史的和实证的调查资料,围绕当代中国妇女地位调查研究、妇女家庭地位、妇女社会地位、妇女的就业、职业与教育状况、妇女地位与生育等专题,作了广泛的理论上的探讨和

研究。许多代表在发言中谈到,中国政府一贯重视妇女解放,重视妇女地位的提高,重视妇女在建设中的作用。当今的改革开放和市场经济条件,使广大妇女获得了广泛的机遇,进一步调动了她们的积极性,从而有利于并能从根本上促进广大妇女社会经济和家庭地位的提高。代表们指出,从另一方面看,重男轻女的现象在我国许多地区,特别是农村地区还很严重,这是导致部分农村地区生育率居高不下的重要原因。进一步提高中国农村妇女地位,纠正重男轻女的陋习,应成为全社会关注的议题。

会议由中国社科院人口所副所长沙吉才研究员主持,所长田雪原研究员在结束时就当前中国人口科学深入开展研究须紧密结合社会主义市场经济体制进行、应加强市场经济体制与流动人口问题研究,以及广泛开展学术交流等问题发表了意见。

(本刊记者)