

中国不同地区死亡水平差异的模型分析

开 昕

【提要】 本文通过对中国各地区死亡水平差异的分析,提出了死亡水平与社会经济发展程度的非线性模型,用以认识中国各地区人口死亡水平的变化规律,预测未来各地区人口死亡水平的变化趋势。

【作者】 开 昕 首都经济贸易大学人口所,讲师。

1. 问题的提出

从已公布的根据第三、四次人口普查数据计算出的各省(区)生命表看,中国人口平均预期寿命一直保持着东部、沿海地区较高、西部和内陆地区较低的态势。尽管1990年西南、西北各省的平均期望寿命比1982年有较大幅度的提高,但从全国范围来看,仍属较低水平。同时,东部沿海地区的期望寿命也比1982年有不同程度地提高。因此,东西部地区平均期望寿命相差的最大幅度依然没有改变。如何解释东西部死亡水平的这种差距?很自然,我们会想到东西部地区的经济发展程度对死亡水平产生的影响。

导致死亡水平变化的因素很多,如生物、遗传、自然环境及社会经济条件等。但其中社会经济条件对死亡水平有着极为重要的作用。因为,它可以通过改善人口的营养状况、医疗卫生条件和生活居住环境而降低人口死亡水平。

近几十年来,对死亡率社会经济差异的研究越来越受到重视。对中国各地区死亡水平差异与社会经济发展程度的关系,很多学者进行了分析,提出不少有益见解。但从分析方法上看,无论是从历史变化角度还是从时期差异角度看,大多采用了对死亡率(或标准化死亡率)的多元回归分析,即选择多个社会经济变量作为解释变量,对死亡率的变化作路径分析或逐步回归分析,从而揭示各种社会经济变量对死亡率的影响。

多元回归(相关)分析,不失为研究各地区社会经济差异与死亡率变化关系的一种好方法,但它存在以下几点不足:

第一,将被解释变量(死亡率)与各个社会经济变量的关系定义为直线关系,在一定程度上不能准确反映死亡水平与各社会经济因素之间的变化关系;

第二,以人口死亡率作为衡量死亡水平的指标,不能准确反映人口死亡水平,因它受人口年龄结构的影响。即使采用标准化死亡率,也存在一个相对参照系的选择问题,而且不直接;

第三,由于各地区的多个社会经济指标(解释变量)是同时发生变化的,因此,在预测未来死亡率变化时,就必须首先对多个解释变量进行预测,预测的解释变量越多,带入的误差信息就越多,从而无法得到较为准确的死亡水平的变化;

第四,各个解释变量之间共线性的问题,也会影响到多元回归模型的准确性。

因此,能否用一个综合指标来反映各地区的经济发展水平,建立与死亡水平之间的非线性

性模型,较准确地反映两者间的变化关系,以达到预测未来死亡水平变化的目的,是本文所要探讨的内容。

2. 模型建立

从上面的讨论中我们可以得出,死亡水平与社会经济发展程度的关联模型至少应具备两个特性。第一,无论是社会经济发展程度指标还是死亡水平指标,都应具备较强的代表性;第二,两者之间的变化关系是非线性的,并能准确地加以反映。

2.1 指标选择 人口死亡水平一般应以代表人口总体死亡水平的平均期望寿命来衡量。期望寿命反映了一个地区总死亡水平的高低,不受人口年龄构成的影响,因而可以对比不同地区死亡水平总体上的差异。

婴儿死亡概率从另一方面反映了人口死亡水平状况,做为衡量人口生活质量的指标之一,具有特殊意义。因此,同时建立婴儿死亡概率与社会经济发展程度的关系模型,也是十分必要的。

社会经济发展水平,可以用很多指标来衡量。其中国民生产总值(GNP)能较全面地反映出一个地区最终创造的全部财富。考虑到各地区人口数量的差距,本文选择人均GNP指标做为衡量社会经济发展水平的指标。

2.2 期望寿命与社会经济发展水平的关联模型 在确定了死亡水平指标和社会经济发展水平指标后,首先应考虑的是两者的关系。

纵观人类发展历史,在社会生产力水平由低到高的变化过程中,人口的死亡水平则是经历了由高到低的变化过程。人口死亡水平的下降速度是不均匀的,是一个由快到慢的变化过程。因此,代表社会经济发展程度的人均GNP与反映死亡水平的平均预期寿命的关系是非线性的。基于这一点,本文建立了如下的关联模型:

$$Y = \frac{KK}{1 + \exp(a + b/z)} \quad (1)$$

Y代表平均预期寿命,Z表示经济发展水平,KK、a、b为待定参数。KK用以调整期望寿命的上限。

利用1990年第四次人口普查资料编制出各省市生命表中的期望寿命和各省市人均GNP数据,用最小二乘法,对模型中各参数进行辨识后得到如下结果:

$$Y = \frac{110}{1 + \exp(-0.93 + 1.47/z')} \quad (2)$$

最终参数取值 $KK=110$, $a=-0.93$, $b=1.47$ 。其中 z' 为人均GNP值 x 的变形值, $z' = x/1000 + 2$ 。即将人均GNP值 x 变形后做为自变量。这主要是考虑到人均GNP值 x (元)本身过大。KK是对平均期望寿命上限的界定,是在辨识了一系列 a 、 b 值的基础上,根据最优化原理确定的。

对模型检验结果为:判定系数 $R^2=0.63$,期望寿命最大相对误差小于6%。拟合效果如图1所示。

由判定系数和图1可知,模型较准确地反映了人口平均期望寿命与人均GNP的变化关系。那么,模型具备的性质如何?以下,对模型性质做一粗略的分析。

由式(2) $Y = \frac{KK}{1 + \exp(a + b/z')}$ 可知,

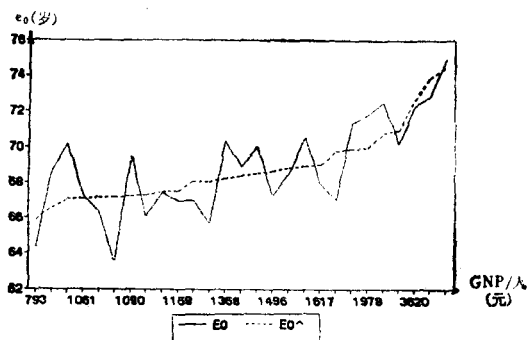


图1 平均期望寿命与人均GNP关系模型的检验

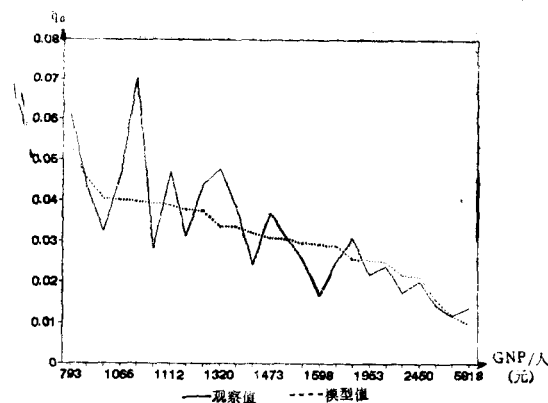


图2 婴儿死亡概率与人均GNP关联模型的检验

(1) 函数是有界的, 即当人均GNP增加时, 平均期望寿命趋于一定值。且Y总为正, 即 $0 < Y \leq M$ (M为正数)。

(2) 平均期望寿命与人均GNP的变化关系是非线性的, 且随着人均GNP的增大平均期望寿命也随之增大, 但增加的速度是递减的。这可以从对式(1)的求导中看出。

求y对 z' 的一阶导数有:

$$\frac{dy}{dz'} = \frac{KK \cdot b}{z'^2} \cdot \frac{\exp(a+b/z')}{[1 + \exp(a+b/z')]^2}$$

显然: $\frac{dy}{dz'} > 0$, 它表明随 z' 增加, y也增大。进一步对 z' 求二阶导数有:

$$\frac{d^2y}{dz'^2} < 0 \text{ (表达式略)}$$

它表明, 随着人均GNP的增大, 平均期望寿命的增加速度是递减的。这与实际情况十分相符。

3. 婴儿死亡概率与人均GNP关系模型的建立

由于婴儿死亡概率的特殊性, 本文认为有必要建立婴儿死亡概率与社会经济发展水平的关系模型。

根据1990年生命表中的婴儿死亡概率, 本文建立的模型如下:

$$Y = \frac{1}{1 + \exp[a + b(z - 1/z)]} \quad (3)$$

式中: Y为婴儿死亡概率, z为人均GNP值X的变换: $z = \frac{X}{2000}$, 同样由最小二乘法, 经辨识后, 参数a取值3.68, b取值0.37。分子1表示最大婴儿死亡概率y将不超过1, 是对婴儿死亡概率y的界定。

经检验, 模型判定系数 $R^2 = 0.73$ 。拟合效果如图2所示。

同样, 模型建立之后, 应对其性质进行分析。

由式(3) $Y = \frac{1}{1 + \exp[a + b(z - 1/z)]}$ 可知, 婴儿死亡概率y与人均GNP之间的变化关

系是非线性的。具有以下几个性质：

(1) $0 < y < 1$ ，这与婴儿死亡概率本身的性质相符；

(2) $\frac{dy}{dz} < 0$ （表达式略），这表示当人均GNP增加时，婴儿死亡概率是下降的；

(3) 当人均GNP值增大到某一值时，婴儿死亡概率将趋近于零。即： $\lim_{z \rightarrow +\infty} y \rightarrow 0$

4. 对模型的进一步讨论及结论

上述两个模型式(2)、式(3)都采用了罗吉斯蒂曲线的结构。其特点是保证了平均期望寿命(或婴儿死亡概率)与社会经济水平(人均GNP)之间一一对应关系。由于连续高阶可导，避免了奇异点的出现并保证了曲线的光滑。而且，重要的是两个模型都较准确地反映出平均期望寿命与人均GNP、婴儿死亡概率与人均GNP之间非均匀的关系。模型结构简单、合理、便于应用。

从模型(2)和(3)的表达式或拟合图(图1、2)还可以得出，中国各省市人均期望寿命和婴儿死亡概率与人均GNP之间具有如下的关系：当人均GNP在1000元以下时，人均GNP每增加100元，人口平均期望寿命将提高0.4岁，婴儿死亡概率将下降4%左右；当人均GNP达到1500元后，人均GNP每增加100元，人口平均期望寿命仅提高0.3岁，婴儿死亡概率仅下降1.4%左右；当人均GNP增加到2000元以后，人均GNP每上升100元，期望寿命仅提高0.2岁，相应婴儿死亡概率下降幅度也减小为0.7%左右。

从以上对模型性质讨论和拟合情况看，我们只能说模型较准确地反映了人口死亡水平随社会经济发展而变化的这种发展趋势，它不是精确的描述。从图1、2可以看出，实际值总是围绕模型值而上下波动。这种波动正好说明了两个问题：一是影响死亡水平变化的因素错综复杂，对中国各地区死亡水平差异的研究仍有待于进一步完善；二是正是社会经济水平的差异导致了中国各地区死亡水平的巨大差异。

本文建立的中国各省市社会经济发展水平与死亡水平之间的关联模型，可以作如下应用：当我们掌握了一个地区的经济发展水平时，就能大致估计出该地区人口平均预期寿命和婴儿死亡概率的范围，再结合模型生命表，就能了解到该地区人口死亡状况的比较完整信息；反之，利用死亡信息，也能大致估计该地区的经济发展水平。这是模型的外推，对于这一步工作，尚有待进一步研究和探讨。

参考文献

- 1 门俐俐.影响中国人口死亡率水平的社会经济因素分析.中国人口科学, 1992. 5
- 2 孙文生等.影响中国死亡率水平的社会经济因素的实证分析.人口与经济, 1995. 4
- 3 梁鸿.经济与人口死亡率的模型分析.人口与经济, 1994. 4
- 4 李树茁.80年代中国人口死亡水平和模式变动分析.人口研究, 1994. 2
- 5 高凌.中国不同职业人口的死亡率差异分析.中国人口科学, 1995. 4
- 6 蒋正华.人口分析与规划.陕西科技出版社

(本文责任编辑：徐莉)