1998年第1期
(总第106期)
人口与经济
POPULATION & ECONOMICS
No.1,1998
(Tot.No.106)

由简略生命表生成完全生命表减小误差的最小修正方法

亓 昕　　李 南

在人口分析和人口预测中,我们常常需要将已有的简略生命表化解成完全生命表,如何保留原来死亡模式的主要信息,最大程度地减小误差,这就是本文提出的由简略生命表生成完全生命表的减小误差的最小修正方法。

一、问题的提出

在人口分析和预测中,我们一般是利用生命表技术对未来人口死亡的变化规律做出预测和分析。由普查和抽样调查数据,我们可以编制简略生命表,再利用其他信息,一个地区的死亡模式就成为已知的了,但是,仅仅根据经验的期望寿命的增长方式和同一模式不同死亡水平的简略生命表,我们无法预测未来每一年的人口死亡状况,还必须将简略生命表生成完全生命表,这样未来每一年的死亡水平和模式才是已知的。Heligman 和 Pollard 曾提出了一个由简略生命表生成完全生命表的八参数模型,即:

$$_1q_x = A^{(x+B)^C} + D^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1+GH^x}$$

这一模型已被制成应用软件(Mortpark),被人口学者和有关人员广泛地用于人口预测中了,有了这一软件,将简略生命表生成完全生命表就成为一件非常容易的事情。但是,在应用中我们发现,由这一软件生成的完全生命表并不能够准确地反映原来死亡模式的主要信息,一组不同死亡水平、同一模式的简略生命表,经过该软件运行后所得到的完全生命表并不同属于同一死亡模式,其死亡概率曲线是相交的。如何在生成完全生命表时,保留原简略生命表的主要信息,并使其误差相对较小,这就是本文所要介绍的由简略生命表生成完全生命表的减小误差的最小修正方法。

二、将简略生命表生成完全生命表的方法

1.两种方法简解

将简略生命表生成完全生命表,是人口死亡分析中很关键的一个问题。1980年,Heligman 和 Pollard 提出了一个如下的八参数模型:

$$_1q_x = A^{(x+B)^C} + D^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1+GH^x}$$

其中:A, B, C, D, E, F, G, H 为待定参数;$_1q_x$ 为 x 岁到 x+1 岁的死亡概率。

这种方法已被制成软件,在联合国推广的人口应用软件 MORTPARK.EXE 中被广泛地使用,其中的子软件 UNABR.EXE 就是这种方法的应用软件。

另一种方法是《The Methods and Material of Demography》[1]一书中提出的差分内插方法。它是将死亡概率曲线分为:0~4岁,5~85岁和85岁以上三个部分。在其中的5~85岁年龄段上,可采用四种不同的差分内插方法,在0~4岁和85^+以上是两个已确定的差分内插方法,这在《The Methods and Materials of Demography》一书中有详细介绍,在此略。

这两种方法,虽然都能将简略生命表生成完全生命表,但产生的误差是相当大的。利用《The Methods and Materials of Demography》一书所提供的方法,将简略生命表生成完全生命表后,再将其化成简略生命表,后者与原简略生命表各点死亡概率的相对误差均在50%左右;应用联合国推广的应用软件中的 UNABR.EXE,生成完全生命表,再化成简略表,前后两个简略表的相对误差为19%,期望寿命误差为±0.5岁,后一种方法虽比第一种方法的精确度提高了许多,但是,它对一些模式表的误差仍然很大,而且,最主要的是,对所产生的误差是不能控制的。这是一种较大的缺陷。

本文提出了一种最小修正的方法,通过最小修正后,能够使得生成后的完全生命表期望寿命与原简略表的期望寿命误差在0.02岁以内,完全表化成简略表后,对应原始简略表各点上死亡概率的最大

误差不超过 15%，以下就介绍这种方法。

2.最小修正方法

最小修正方法是对 UNABR 方法的改进。在上面所述中，我们已经知道，由于 UNABR 方法在产生完全表时，对某些模式的生命表所产生的误差太大，并且误差不能够被控制，如果我们能够控制各点${}_5q_x$(完全表所对应的新简略表)上的误差，那么，生成的完全表与原简略表之间的期望寿命误差也会降下来，完全表将会更好地保留原简略表的死亡信息。在生命表中我们知道，死亡概率与留存概率之间存在如下的关系：

$${}_np_x = 1 - {}_nq_x$$

其中：${}_np_x$ 为 x 到 x+n 岁的留存概率，${}_nq_x$ 为 x 到 x+n 岁的死亡概率。

并且：${}_np_x$ 与 x 到 x+n 岁各单岁年龄上的留存概率 $P_i(i=x, x+1, \cdots x+n-1)$，存在如下关系：

$${}_np_x = {}_1p_x \cdot {}_1p_{x+1} \cdots {}_1p_{x+n-1}$$

因此，若能控制生成完全表的留存概率与原简略表之间的误差的话，那么，死亡概率的误差也就解决了。

最小修正方法，就是基于这样一种设想。那么，如何来构造目标函数呢？我们知道，原始表有一留存概率${}_np_i^*$，完全表也有${}_np_i$，在无误差的条件下应有：${}_np_x^* = p_ip_{i+1}\cdots p_{x+n-1}$成立。因此，目标函数必须使得所有的$\sum(\prod_{i=1}^{x+n-1} p_i - {}_np_i^*)^2$ 最小，同时，还应考虑到，修正后的表与未修正的表的差异不能太大。基于上述两点，我们设目标函数 F_2 如下：

$$设:F_2 = \alpha(\ln P_i^* - \sum_{j=1}^{N}\ln P_{ij})^2 + (1-\alpha)\sum_{j=1}^{N}(\ln P_{ij} - \ln P_{ij}^*)^2$$

其中：P_i^* 为原简略生命表第 i 年龄组的留存概率。

P_{ij}为完全生命表第 i 年龄上的留存概率，它是修正以后的值。

P_{ij}^* 表示未修正的完全生命表第 i 年龄上留存概率。

N 为简略生命表的年龄组间隔。

求解 $\quad \mathrm{Min}F_2$

可得 $\quad P_{ij}$

令 $\quad T_i = \ln P_i^*$

$$\bar{R}_{ji} = \ln P_{ij}$$

$$R_{ji} = \ln P_{ij}^*$$

则 $F_2 = \alpha(T_i - \sum_{j=1}^{N}\bar{R}_{ji})^2 + (1-\alpha)\sum_{j=1}^{N}(\bar{R}_{ji} - R_{ji})^2$

欲使 $\min F_2(\bar{R}_{ji})$，对 F_2 求偏导：

$$\frac{\partial F_2}{\partial \bar{R}_{Ki}} = 0 \qquad (K 取法同 N)$$

$$-\alpha(T_i - \sum_{j=1}^{N}\bar{R}_{ji}) + (1-\alpha)(\bar{R}_{Ki} - R_{Ki}) = 0$$

整理得：

$$V\bar{R} = \alpha T_i I + (1-\alpha)R$$

$$V = \begin{pmatrix} 1 & \alpha & \cdots & & \alpha \\ \alpha & 1 & \alpha & \cdots & \alpha \\ \vdots & & & & \\ \alpha & & \cdots & & 1 \end{pmatrix}_{N\times N} \qquad \bar{R} = \begin{pmatrix} \bar{R}_{1i} \\ \vdots \\ \bar{R}_{Ni} \end{pmatrix}$$

$$R = \begin{pmatrix} R_{1i} \\ \vdots \\ R_{Ni} \end{pmatrix} \qquad I = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$可以证明:V^{-1} = \begin{pmatrix} X & Y & \cdots & Y \\ Y & & & \\ \vdots & & & \\ Y & \cdots & Y & X \end{pmatrix} \tag{1}$$

$$X = \frac{-(N-2)\alpha - 1}{(N-1)\alpha^2 - (N-2)\alpha - 1} \tag{2}$$

$$Y = \frac{\alpha}{(N-1)\alpha^2 - (N-2)\alpha - 1} \tag{3}$$

$$即:\bar{R}_{ki} = [X + (N-1)Y]\alpha T_i + (1-\alpha) \cdot (X-Y)R_{Ki} + (1-\alpha)Y + \sum_{j=1}^{N}R_{ji} \tag{4}$$

$$令: \quad T_i - \sum_{j=1}^{N}R_{ji} = R_{1i} \tag{5}$$

$$\Delta = T_i - \sum_{j=1}^{N}\bar{R}_{ji} \tag{6}$$

由(2)、(3)、(4)有：

$$\bar{R}_{Ki} = R_{Ki} - Y(1-\alpha)R_{1i} \tag{7}$$

$$\sum_{K=1}^{N}\bar{R}_{Ki} = \sum_{K=1}^{N}R_{Ki} - YN(1-\alpha)R_{1i} \tag{8}$$

将(5)、(6)代入(8)

$$T_i - \Delta = T_i - R_{1i} - YN(1-\alpha)R_{1i}$$

$$\Delta = R_{1i}[1 + YN(1-\alpha)] \tag{9}$$

将(3)代入(9)

$$\Delta = -\frac{(\alpha-1)^2}{(N-1)\alpha^2 - (N-2)\alpha - 1}R_{1i} \tag{10}$$

将(10)、(7)合并消去 R_{1i}

得：$\bar{R}_{ki} = R_{Ki} - \frac{\alpha}{\alpha-1}\Delta$

将(10)式化简

$$R_{1i} = (1 - \frac{\alpha}{\alpha-1}N)\Delta$$

$$\frac{R_{1i}}{\Delta} = [1 - (\alpha/\alpha-1)]N$$

令： $f(\alpha) = 1 - \frac{\alpha}{\alpha - 1}N$

则： $f(\alpha) \in C(0,1)$

$$f'(\alpha) = \frac{N}{(\alpha-1)^2} > 0$$

$$f(0) = 1 \qquad f(1) = +\infty$$

故　当 $1 < \frac{R_{1i}}{\Delta} < +\infty$

有唯一解 $\alpha_i \in (0,1)$，

使： $\frac{R_{1i}}{\Delta} = 1 - \frac{\alpha_0}{\alpha_0 - 1}N$

记 α_0 的存在条件式：

$$1 < \frac{R_{1i}}{\Delta} < +\infty \qquad (11)$$

当(1)成立：

$$\frac{\alpha_0}{\alpha_0 - 1} = \frac{1}{N}\left(1 - \frac{R_{1i}}{\Delta}\right) \qquad (12)$$

将(12)代入

$$\bar{R}_{Ki} = R_{Ki} - \frac{\alpha}{\alpha - 1}\Delta \quad \text{式中}$$

$$\text{得 } \bar{R}_{Ki} = R_{Ki} - \frac{1}{N}(\Delta - R_{1i})$$

记：

$$R_{Zi} = P_i^* - \prod_{j=1}^{N} P_{ji}^* \qquad (13)$$

$$E = P_i^* - \prod_{j=1}^{N} P_{ji}$$

则
$$\ln\left(1 - \frac{R_{Zi}}{P_i^*}\right) = \ln\left(\frac{P_i^* - R_{Zi}}{P_i^*}\right) = \ln\prod_{j=1}^{N} P_{ji}^* - \ln P_i^* = \sum R_{ji} - T_i = R_{1i}$$

$$\ln\left(1 \frac{E}{P_i^*}\right) = \ln(P_i^* - E) - T_i = \sum \bar{R}_{ji} - T_i = \Delta$$

$$\bar{R}_{Ki} = R_{Ki} - \frac{1}{N}\left[\ln\left(1 - \frac{E}{P_i^*}\right) - \ln\left(1 - \frac{R_{Zi}}{P_i^*}\right)\right]$$

$$\ln P_{Ki} = \ln P_{Ki}^* - \frac{1}{N}\ln\frac{1 - \frac{E}{P_i^*}}{1 - \frac{R_{Zi}}{P_i^*}}$$

$$P_{Ki} = P_{Ki}^* \sqrt[N]{\frac{1 - E/P_i^*}{1 - R_{Zi}/P_i^*}} \qquad (14)$$

令：$E_1 = \left(\prod_{j=1}^{N} P_{ji} - P_i^*\right)/P_i^*$

则：

$$E = E_1 P_i^* \qquad (15)$$

由(13)、(14)、(15)式，就得到修正解 R_{Ki}。

在上述方法的基础上，可以编制程序。首先，利用 Mortpark 运行软件，产生一张完全生命表，将这张完全表做为最小修正法程序的输入文件，同时输入的还有原始简缩生命表。从终端输入所要求的精度 $E_1(0 < E_1 < 1)$，根据完全生命表预期寿命与原简略表之间的误差，调整 E_1，选择 E_1，使其所修正后的完全表的预期寿命与原简略表的预期寿命之间的误差，达到满意的程度为止。

3.对上述方法的测试

我们将这一最小修正方法用于各种模式的简略生命表生成完全生命表中，将所得到的完全生命表再化成简略生命表，与原简略生命表比较，其平均期望寿命的误差均在 0.02 岁以内，对应各年龄点上死亡概率的相对误差不超过 15%。而且生成的同一模式的完全生命表是不相交的，较好地保留了原死亡模式的信息。

最小修正方法是一种很实用的方法，它使得原简略生命表的死亡信息被最大程度地保存下来。在实际人口预测中，具有较高的应用价值。

参考文献：

1　U.S. BUREAU OF CENSUS. THE METHODS AND MATERIALS OF DEMOGRAPHY. U.S. GOVERNMENT PRINTING OFFICE, 1971

(作者工作单位：

亓昕　首都经贸大学人口所　北京市　100026

李南　西安交通大学人口所　西安市　710049)