

已知死亡水平和死亡模式时的 生命表的编制

路 磊

一、问题的提出

在编制实际生命表时,我们通常是先根据人口普查或人口调查提供的资料计算出一组分年龄的死亡率 $\{M(x,n)\}$,然后再根据这组实际的分年龄死亡率编制出相应的生命表。在编制模型生命表时,我们通常是根据一组分年龄的死亡概率 $q(x,n)$ 或分年龄的尚存人数 $l(x)$ 来编制生命表。但是在很多情况下,由于各种原因我们不可能获得如此详细的原始资料,因此也就不能用通常的方法来编制生命表了。

比如在两次人口普查之间,我们一般很难得到分年龄的死亡资料,因此也就无法直接编制普查间各年度的生命表。但是我们可以通过使用间接技术估计出普查间的一个平均生命表,以及各年度(或部分年度)的婴儿死亡率或出生时的平均预期寿命。这些估计值向我们提供了有关死亡模式和死亡水平方面的重要信息。

又如在人口预测中,为了进行年龄移算,需要准确地预测各个年龄组死亡率的未来变化趋势。但这是一项十分困难的工作,主要原因有两个:第一,由于缺乏有关的历史资料(尤其是分死因的数据),使得我们无法对死亡率的历史变化过程进行动态分析,因而也就很难对死亡率的未来变化趋势做出十分准确的判断;第二,即使对历史有所了解,但是由于未来往往充满很多不确定性,因此对历史变化过程的外推常常会产生较大的误差。但这并不等于说未来的死亡率变化趋势是无法预测的,因为至少有一点是可以肯定的,即人类的寿命将随着社会经济的发展、医疗卫生水平的提高、营养的改善和教育的普及而不断提高。这个事实早已为世界上许多国家的历史所证明了。一般来讲,预测平均预期寿命的未来变化趋势要比直接预测各个年龄组的死亡率容易得多,也准确得多。我们可以采用如下方法预测死亡率的未来变化趋势:首先根据历史资料提供的变化趋势以及对未来的估计,用模型拟合外推或经验方法对将来的人口出生时的平均预期寿命做出预测(可做多种方案),并且假定死亡模式与某个模型生命表或某个实际生命表相似,然后据此编制出相应的生命表。

在编制模型生命表时也会遇到类似的问题,以上这些问题都涉及到了在已知死亡水平和死亡模式时的生命表的编制问题。在生命表中,死亡水平可以用多种变量来表示。比如,中心死亡率 $m(x,n)$ 、死亡概率 $q(x,n)$ 、尚存人数 $l(x)$ 和平均预期寿命 $e(x)$ 等,而死亡模式通常用 $q(x,n)$ 或 $l(x)$ 表示。本文将讨论一种基于logit变换的生命表编制方法。

二、Logit变换

英国著名人口学家威廉·布拉斯及其同事在研究模型生命表时,首次将概率统计中的logit变换引入了人口学研究。通过引进logit函数,布拉斯等人提出了一个较为简单、灵活的模型生命表生成系统。该模型的提出有力地促进了有关模型生命表、死亡率模拟和死亡率预测等方面的研究。

对一个生命表人口,用 $p(a)$ 表示一个人从出生(确切年龄0岁)活到确切年龄 a 岁的存活概率,即 $p(a) = l(a)/l(0)$ 。定义 $p(a)$ 的 λ 变换为:

$$\lambda[p(a)] = \text{logit}[1 - p(a)] = \frac{1}{2} \ln \left[\frac{1 - p(a)}{p(a)} \right] \quad (1)$$

由(1)式可以看出, $p(a)$ 的 λ 变换即为 $1 - p(a)$ 的logit变换。布拉斯等人的研究表明,两个不同生命

表中 $p(a)$ 的 λ 变换之间存在着明显的线性关系, 即: $\lambda[p(a)] = \alpha + \beta \lambda[p_s(a)]$ (2)

(2) 式是一个双参数模型。其中 $p(a)$ 和 $p_s(a)$ 为两个不同生命表中从出生到确切年龄 a 岁的存活概率, α 和 β 为两个参数。

根据上述模型, 如果选定了一个标准生命表的存活概率 $p_s(a)$, 那么新生命表的存活概率 $p(a)$ 将完全由参数 α 和 β 决定。具体地讲, $p(a)$ 的水平将主要由参数 α 的数值决定, $p(a)$ 的模式将主要由参数 β 的数值决定。特别地, 当 $\beta=1$ 时, $p(a)$ 的模式与 $p_s(a)$ 的模式基本相同。为了简便起见, 我们将忽略参数 β 对 $p(a)$ 的水平的影响, 只考虑其对 $p(a)$ 的模式的作用, 并且假定 $\beta=1$, 即新生成的生命表与原始生命表有相似的模式。于是(2)式可以简化为: $\lambda[p(a)] = \alpha + \lambda[p_s(a)]$ (3)

显然(3)式是一个单参数(α)模型。由(1)式和(3)式可以解得:

$$p(a) = \frac{1}{1 + \exp(2\alpha + 2\lambda[p_s(a)])} = \frac{p_s(a)}{p_s(a) + [1 - p_s(a)] \exp(2\alpha)} \quad (4)$$

(4) 式给出了 $p(a)$ 与参数 α 之间的解析关系, 是我们用来编制新生命表的基本方程。从(4)式可以看出, 当 $p_s(a)$ 固定时, $p(a)$ 是一个关于参数 α 的严格单调递减的连续函数。当 $\alpha < 0$ 时, 有 $p(a) > p_s(a)$; 当 $\alpha = 0$ 时, 有 $p(a) = p_s(a)$; 当 $\alpha > 0$ 时, 有 $p(a) < p_s(a)$ 。显然 $0 < p(a) < 1$ 。并且当 α 趋于 $+\infty$ 时, $p(a)$ 趋于0; 当 α 趋于 $-\infty$ 时, $p(a)$ 趋于1。当我们根据某个死亡水平估计出相应的 α 值后, 将该值代入(4)式即可得到一组存活概率 $\{p(a)\}$, 进而可以编制出完整的生命表。下面我们就来建立参数 α 与生命表中各种(水平)变量之间的关系。

三、已知 $m(x, n)$ 时 α 值的确定

现在我们讨论参数 α 与中心死亡率 $m(x, n)$ 的关系, 即给定一个 $m(x, n)$ 值后, 如何确定相应的 α

$$\begin{aligned} \text{值。由于 } m(x, n) &= \frac{d(x, n)}{L(x, n)} = \frac{d(x, n)}{n l(x+n) + a(x, n) d(x, n)} \\ &= \frac{l(x) - l(x+n)}{n l(x+n) + a(x, n) [l(x) - l(x+n)]} \\ &= \frac{p(x) - p(x+n)}{n p(x+n) + a(x, n) [p(x) - p(x+n)]} \end{aligned} \quad (5)$$

由(4)式和(5)式可以解得:

$$\alpha = -\frac{1}{2} \ln \left[\left(\frac{1}{m(x, n)} - a(x, n) \right) \left(\frac{p_s(x) - p_s(x+n)}{n p_s(x) p_s(x+n)} \right) - \frac{1 - p_s(x)}{p_s(x)} \right] \quad (6)$$

(6) 式即是参数 α 与 $m(x, n)$ 之间的解析关系, 从中可以看出参数 α 与 $m(x, n)$ 成正比。只要给定一个 $m(x, n)$ 的值, 就可以根据(6)式计算出相应的 α 值。下面我们以联合国发展中国家模型生命表General 1模式中 $e(0) = 60$ 岁的男性生命表为标准生命表来说明 α 的计算。由该表可知 $p_s(20) = 0.88195$, $p_s(25) = 0.87051$ 。假定 $m(20, 5) = 0.00100$, 并取 $a(20, 5) = 2.6$ 。将上述各值代入(6)式, 可得 $\alpha = -0.52165$ 。

四、已知 $q(x, n)$ 时 α 值的确定

在本节中我们讨论参数 α 与死亡概率 $q(x, n)$ 的关系, 即给定一个 $q(x, n)$ 值后, 如何确定相应的 α

$$\text{值。由于 } q(x, n) = \frac{d(x, n)}{l(x)} = \frac{l(x) - l(x+n)}{l(x)} = \frac{p(x) - p(x+n)}{p(x)} \quad (7)$$

由(4)式和(7)式可以解得:

$$\alpha = -\frac{1}{2} \ln \left[\frac{p_s(x) - p_s(x+n)}{q(x, n) p_s(x) p_s(x+n)} - \frac{1 - p_s(x+n)}{p_s(x+n)} \right] \quad (8)$$

(8)式即是参数 α 与 $q(x, n)$ 之间的解析关系,从中可以看出参数 α 与 $q(x, n)$ 成正比。只要给定一个 $q(x, n)$ 的值,就可以根据(8)式计算出相应的 α 值。下面我们以联合国发展中国家模型生命表General模式中 $e(0) = 60$ 岁的男性生命表为标准生命表,来说明 α 的计算。由该表可知 $p_s(20) = 0.88195$, $p_s(25) = 0.87051$ 。假定 $q(20, 5) = 0.00500$ 。将上述各值代入(8)式,可得 $\alpha = -0.52039$ 。

五、已知 $l(x)$ 时 α 值的确定

在本节中我们讨论参数 α 与尚存人数 $l(x)$ 的关系,即给定一个 $l(x)$ 值后,如何确定相应的 α 值。由于 $l(x) = l(0)p(x)$ (9)

$$\text{由(4)式和(9)式可以解得: } \alpha = \frac{1}{2} \ln \left[\left(\frac{l(0)}{l(x)} - 1 \right) \left(\frac{p_s(x)}{1 - p_s(x)} \right) \right] \quad (10)$$

(10)式即是参数 α 与 $l(x)$ 之间的解析关系,从中可以看出参数 α 与 $l(x)$ 成反比。只要给定一个 $l(x)$ 的值,就可以根据(10)式计算出相应的 α 值。下面我们以联合国发展中国家模型生命表General模式中 $e(0) = 60$ 岁的男性生命表为标准生命表,来说明 α 的计算。由该表可知 $p_s(20) = 0.88195$, $p_s(25) = 0.87051$ 。假定 $l(20) = 94986$ 。将上述各值代入(10)式可得 $\alpha = -0.46523$ 。

六、已知 $e(x)$ 时 α 值的确定

在本节中我们讨论参数 α 与平均预期寿命 $e(x)$ 的关系,即给定一个 $e(x)$ 值后,如何确定相应的 α 值。

$$\begin{aligned} \text{由于 } e(x) &= \frac{T(x)}{l(x)} = \frac{1}{l(x)} \int_x^{\infty} l(a) da = \frac{1}{p(x)} \int_x^{\infty} p(a) da \\ &= \int_x^{\infty} \frac{p_s(a)[p_s(x) + (1 - p_s(x))\exp(2\alpha)]}{p_s(x)[p_s(a) + (1 - p_s(a))\exp(2\alpha)]} da \end{aligned} \quad (11)$$

由于很难求出参数 α 与 $e(x)$ 之间的解析关系,我们将采用一种二分迭代法根据给定的 $e(x)$ 值估计相应的 α 值。从(11)式可以看出 $e(x)$ 是参数 α 的函数,记作 $e(x, \alpha)$ 。在(11)式两边对参数 α 求导可得,

$$\frac{de(x, \alpha)}{d\alpha} = \int_x^{\infty} \frac{2p_s(x)p_s(a)\exp(2\alpha)[p_s(a) - p_s(x)]}{[p_s(x)(p_s(a) + (1 - p_s(a))\exp(2\alpha))]^2} da \quad (12)$$

由于 $p_s(a)$ 是一个关于年龄 a 的严格单调递减的函数,所以 $p_s(a) < p_s(x)$,于是由(12)式可知:

$$de(x, \alpha)/d\alpha \leq 0 \quad (13)$$

(13)式表明 $e(x, \alpha)$ 是一个关于参数 α 的严格单调递减的函数。下面我们来讨论如何根据给定的 $e(x)$ 值确定相应的 α 值。首先根据研究的需要选择一个合适的标准生命表 $\{p_s(a)\}$,然后再人为地选择两个 α 值,分别记作 α_{min} 和 α_{max} (满足 $\alpha_{min} < \alpha_{max}$),将 $p_s(a)$ 和 α 值代入(4)式可以计算出两组存活概率,进而可以得到两个完整的生命表。记这两个生命表中 x 岁的平均预期寿命分别为 $e_{max}(x) = e(x, \alpha_{min})$ 和 $e_{min}(x) = e(x, \alpha_{max})$ 。注意,在选择 α_{min} 和 α_{max} 时应使计算出来的 $e_{max}(x)$ 足够大(比如为80岁), $e_{min}(x)$ 足够小(比如为2岁),以保证任何可以预测到的 $e(x)$ 值介于 $e_{min}(x)$ 和 $e_{max}(x)$ 之间。由于 $e(x, \alpha)$ 的严格单调递减性,对任意给定的 $e^*(x)$,只要它满足 $e_{min}(x) < e^*(x) < e_{max}(x)$,则一定存在一个 α^* (满足 $\alpha_{min} < \alpha^* < \alpha_{max}$)使得 $e(x, \alpha^*) = e^*(x)$ 。二分迭代法的步骤如下:

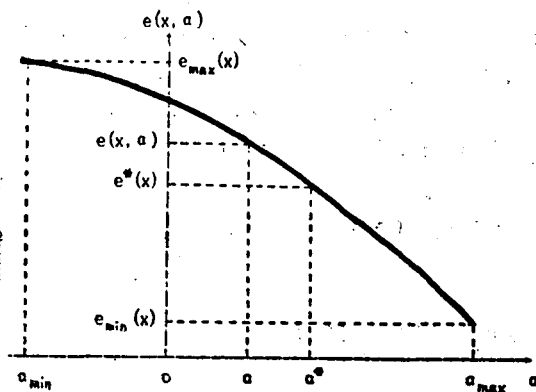
第一步:确定收敛精度 ϵ , $\epsilon > 0$;

第二步: $\alpha = \frac{1}{2}(\alpha_{min} + \alpha_{max})$, 代入(4)式计算相应的 $e(x, \alpha)$;

第三步:如果 $|e(x, \alpha) - e^*(x)| < \epsilon$, 则取 $\alpha^* = \alpha$, 迭代终止, 否则转第四步;

第四步:如果 $e(x, \alpha) < e^*(x)$, 说明 $\alpha > \alpha^*$, 令 $\alpha_{max} = \alpha$, 转第二步, 否则说明 $\alpha < \alpha^*$, 令 $\alpha_{min} = \alpha$, 转第二步。

下图是二分迭代法的示意图。下面我们来证明上述二分迭代法是收敛的。在上述迭代过程中,我们实际上是构造了许多个区间,而每个区间都包含精确解 α^* 。记第 n 次迭代对应的区间为 $[a_n, b_n]$, 相应的区间长



度记为

$d_n = |b_n - a_n|$, 则有 $d_n = \frac{1}{2} d_{n-1}$ 。由此可

$$\text{得: } d_n = \frac{1}{2^n} d_0 = \frac{1}{2^n} |a_{\max} - a_{\min}| \quad (14)$$

显然当 $n \rightarrow +\infty$ 时, $d_n \rightarrow 0$ 。又由于 $a_n \leq \alpha^* \leq b_n$,

$$\text{所以 } 0 \leq |a_n - \alpha^*| \leq |b_n - a_n| = d_n \quad (15)$$

$$0 \leq |b_n - \alpha^*| \leq |b_n - a_n| = d_n \quad (16)$$

由此可知 $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \alpha^*$, 所以该算法收敛。

七、结束语

美国著名人口学家安斯雷·寇尔和保罗·德曼,在区域模型生命表中曾经提出一种根据死亡水平和死亡模式计算生命表的方法。该方法是目前使用最多的方法之一。不过它明显存在一些弱点:(1)由于该模型中的方程是建立在经验数据基础之上的,而这些经验数据又存在着明显的时空局限性,所以该模型的普遍适用性受到限制。(2)该模型只能生成四种模式的生命表,即模型的灵活性较低。如果所研究的死亡模式与区域模型生命表中的所有模式都不相符,那么该模型不适用。(3)该模型只能生成简略生命表,要想得到完全生命表,还必须进行插值。

联合国人口司在发展中国家模型生命表中也提出了一种根据死亡水平和死亡模式计算生命表的方法。该模型在实际中应用很广泛,而且其灵活性要高于寇尔和德曼的模型,但是它也存在一些问题:(1)由于主因素向量是建立在经验数据基础之上的,所以有可能不具有普遍适用性。(2)该模型也只能生成简略生命表,要想得到完全生命表,必须进行插值。

本文中作者讨论了一种根据死亡水平和死亡模式估计生命表的方法。与上述两种方法相比,本文的模型有如下几个优点:(1)模型的适用性强。模型中的关系不受经验数据的影响,适用于各种情况。(2)模型的灵活性好。用户可以根据实际需要提供各种各样的死亡模式,通过本模型都可以作成与之有相似模式的生命表。(3)既可生成简略生命表,又可生成完全生命表。本模型可根据用户提供的原始信息的粗细情况生成相应的简略的或完全的生命表。

参考文献:

1. Arriaga, Eduardo, Patricia Anderson and Larry Heligman. 1976. Computer Programs for Demographic Analysis. U.S. Government Printing Office.
2. Coale, Ansley J., and Paul Demeny with Barbara Vaughan. 1983. Regional Model Life Tables and Stable Populations. Academic Press.
3. Keyfitz, Nathan. 1977. Applied Mathematical Demography. John Wiley & Sons.
4. Rogers, Andrei. 1975. Introduction to Multiregional Mathematical Demography. John Wiley & Sons.
5. Schoen, Robert. 1988. Modeling Multigroup Populations. Plenum Press.
6. United Nations. 1982. Model Life Tables for Developing Countries. New York: United Nations.
7. United Nations. 1983. Manual X: Indirect Techniques for Demographic Estimation. New York: United Nations.
8. United Nations. 1988. MortPak-Lite: The United Nations Software Package for Mortality Measurement-Interactive Software for the IBM-PC and Compatibles. New York: United Nations.
9. United Nations. 1990. Step-by-Step Guide to the Estimation of Child Mortality. New York: United Nations.

(本文责任编辑:徐莉)

(作者工作单位:中国人民大学人口研究所)