

人口生命表编制的一种新方法:死亡结构法

杨泽祥¹,赵 萍¹,王 欣²

(1.云南财经大学 统计与数学学院;2.云南电网公司物流服务中心,昆明 650221)

摘 要:文章利用我国第五、六两次人口普查数据,采用伐尔、寇尔、指数三种方法,编制中国2000年和2010年男性、女性和全体(平均)人口完全与简略生命表,同时采用了直接计算预期寿命法直接计算预期寿命。探讨了“死亡结构法”编制人口生命表,并编制了“两普查年”中国男性、女性和全体的6张完全生命表以及其他生命表,结果与“三法”基本一致。

关键词:中国;人口普查;生命表;死亡结构法

中图分类号:C9221

文献标识码:A

文章编号:1002-6487(2014)01-0004-03

0 引言

人口生命表自英国统计学家约翰·格兰特(John Graunt,1920~1674)首次编制,几百年来已广泛应用于人口、生物、保险等领域,取得了辉煌成果。

人口生命表编制的核心是:用粗死亡率 m_x 导出死亡概率 q_x ,在假设同时出生的一批人 l_0 下导出尚存人数,进而推算出平均预期寿命。

由于 q_x 的不同计算方法,导致有不同称谓的编表方法。我们用全国2000年和2010年全国人口普查的数据,进行了多方法、多方案模拟,对编表的方法论进行了研究,提出了一个新的编表方法——死亡结构法。

传统的人口生命表编制技术已十分成熟,其表遍布世界,联合国人口司非常重视,以158张男性和158张女性生命表编制了1955年模型生命表;用36张男性和36张女性生命表又专为发展中国家编制1982年模型生命表。但这些表未包括世界人口第一的中国。

1957年,李光荫根据北京市的人口死亡登记资料编制了北京城区1950年及1953年生命表,这是新中国成立后公开发表的第一个生命表。1953年、1964年两次全国人口普查均未编制生命表,只有一些地区编制了该地区生命表。直到1982年,才在著名人口学家蒋正华教授的主持下,利用我国自1957年后各种人口抽样调查、人口普查的死亡等数据,编制了中国区域模型生命表,开创了我国编制和应用人口生命表的先河。此后,中国人民保险公司于1995年末编制出中国人寿保险业第一张经验生命表——《中国人寿保险业经验生命表》,并从1997年4月1日起正式运用于人寿保险业务的经营核算,中国保监会2003年8月启动了第二张《生命表》的编制工作。亓昕等

用1995年全国人口1%抽样调查数据,编制了《中国1995年各省市分性别简略生命表》,李树苗等用全国第五次人口普查数据编制了《2000年全国市镇县以及各省市人口生命表》。台湾十分重视编制和应用人口生命表,每年依据年龄别死亡人数、年中人口数及月龄别出生人数等资料,按男性及女性编算简易生命表。香港编制人口生命表更是制度化、常态化。

1 寇尔、伐尔和指数三种方法编制生命表的实证

1.1 三种方法的死亡概率公式

(1)寇尔法死亡概率公式:

美国人口学家安斯雷·寇尔提出的死亡概率 q_x 公式

$$q_x = \frac{m_x}{1 + (1 - a_x)m_x} \quad (1)$$

式中, a_x 为寿命成数,表示 $x \sim x+1$ 岁间死亡者的平均生存人年数,它的一般经验参数值见表1。

表1 寿命成数 a_x 经验参数值

年龄	a_x	年龄	a_x
0	0.30	75~79	0.50
1~4	0.375	80~84	0.48
5~9	0.45	85~89	0.44
10~74	0.52	90+	0.42

(2)伐尔法死亡概率公式:

$$\begin{cases} q_x = \frac{2m_x}{2 + m_x}, x=0, 1, \dots, \omega-2 \\ q_{\omega-1} \triangleq 1 \end{cases} \quad (2)$$

伐尔法是寇尔法的特殊情况,即将寿命成数 $a_x=1/2$ 。

(3)指数法死亡概率公式:

$$q_x = 1 - e^{-m_x} \quad (3)$$

1.2 生命表的编制步骤

基金项目:国家社会科学基金资助项目(11BTJ013)

作者简介:杨泽祥(1940-),男,四川简阳人,教授,研究方向:应用统计与数量经济。

赵 萍(1965-),女,云南嵩明人,教授,研究方向:应用数学和应用统计学。

(1) 普查实际数据年龄别平均人数 $\bar{P}$ 和相应的年龄别死亡人数 D_x , 计算年龄别死亡率:

$$m_x = \frac{D_x}{\bar{P}_x}, x=0, 1, 2, \dots, \omega-1 \quad (4)$$

(2) 用公式(1)或(2)或(3)计算死亡概率 q_x ;

(3) 确定表上死亡人口基数, 一般取 $l_0=10$ 万人或100万人;

(4) 计算表上死亡人数 d_x :

$$d_x = l_x \cdot q_x \quad (5)$$

(5) 计算尚存人数 l_x :

$$l_{x+1} = l_x - d_x, x=0, 1, 2, \dots, \omega-1 \quad (6)$$

(6) 计算平均生存人年数 L_x :

$$L_x = \begin{cases} \frac{1}{4}l_0 + \frac{3}{4}l_1, x=0 \\ \frac{1}{2}(l_x + l_{x+1}) + \frac{1}{24}(d_x - d_{x-1}), x=1, 2, 3, 4 \\ \frac{1}{2}(l_x + l_{x+1}), 5 \leq x \leq \omega-2 \\ \frac{1}{2}l_x, x=\omega-1 \end{cases} \quad (7)$$

(7) 计算平均生存总人年数 T_x (向下累积):

$$T_x = \sum_{i=x}^{\omega-1} L_i, x=0, 1, 2, \dots, \omega-1 \quad (8)$$

(8) 计算平均预期寿命 e_x :

$$e_x = \frac{T_x}{l_x}, x=0, 1, 2, \dots, \omega-1 \quad (9)$$

简略生命表只需将以上完全生命表的年龄组距1换为 n , 如 $n=5$, 即为5岁一组简略生命表。

1.3 中国2000年和2010年生命表

我们根据第五次人口普查数据(2000年), 国务院人口普查办公室和国家统计局人口与就业统计司编《中国2010年人口普查资料》(中国统计出版社2012年4月第1版)的资料, 编制了2000年和2010年生命表, 但由于篇幅所限, 本文并未列出。

2 直接计算预期寿命法的实证

直接计算预期寿命法是在已知一人口系统年龄别死亡率 m_x 和寿命成数 a_x (此计算采用表1)的前提下, 首先导出死亡概率 q_x (此计算采用伐尔概率公式(2)); 其次基于最高年龄组 $x=\omega-1$ 的死亡概率 $q_{\omega-1}=1$, 预期寿命 $e_{\omega-1}=0$, 预期寿命增量 $\Delta e_{\omega-1}=1-a_{\omega-1}$ 的事实, 由预期寿命增量递推公式

$$\Delta e_x = \begin{cases} 1-a_{\omega-1}, x=\omega-1 \\ 1-q_x(\Delta e_{x+1}+a_x), x=\omega-2, \dots, 2, 1, 0 \end{cases} \quad (10)$$

推出预期寿命增量 Δe_x , 最后将 Δe_x 向下累积得到预期寿命

$$e_x = \sum_{i=x}^{\omega-1} \Delta e_i, x=\omega-1, \omega-2, \dots, 2, 1, 0 \quad (11)$$

用中国2000年和2010年第五、六次人口普查资料, 直接法推算得到女性和男性的预期寿命 e_x 分别为:

$$e_0^F(2000)=89.93, e_0^M(2000)=88.48, e_0^{\text{平均}}(2000)=89.26$$

$$e_0^F(2010)=96.46, e_0^M(2010)=95.67, e_0^{\text{平均}}(2010)=95.64$$

不论是2000年还是2010年, 也不论是男性还是女性都远比伐尔等三种方法计算的高, 而且高得离出奇, 不符合实际, 失真不可用。

事实上, 直接法与上述三种方法没有实质性的差别, 它是在粗死亡率的基础上以同样的方法导出死亡概率, 所不同的是它不再尚存人数 l_0 的假设下, 计算尚存人数 l_x 、表上死亡人数 d_x 、平均生存人年数 L_x 、累计生存人年数 T_x , 进而计算预期寿命 e_x 。

3 生命表编制的实证分析

3.1 几种方法的简单比较

将以上寇尔法、伐尔法、指数法和直接法四种方法计算得到的预期寿命列于表2

表2 预期寿命 e_0 对照表 (单位: 岁)

	2000年			2010年		
	女性	男性	全体	女性	男性	全体
寇尔法	74.32	70.64	72.42	80.38	75.61	77.29
伐尔法	64.49	61.33	63.00	80.36	75.59	77.27
指数法	74.29	70.63	72.40	80.36	75.60	77.27
直接法	89.93	88.48	89.26	96.46	95.67	95.64
结构法	72.92	71.36	72.03	79.17	75.86	76.39

表2看出, 直接计算法结果比其他几种方法的结果都高, 不可采用。伐尔法的死亡概率计算简单易行, 指数法是在离散化后的实用计算公式, 也简单易行, 但它们都是假设死亡事件在其所在区间服从均匀分布; 寇尔法是美国人口学家安斯雷·寇尔创立, 其死亡概率公式, 誉为精确的计算死亡概率公式。寇尔的创新之处, 是将粗死亡率的计算公式(4)

$$m_x = \frac{D_x}{\bar{P}_x}, x=0, 1, 2, \dots, \omega-1$$

的分母平均人口 $\bar{P}_x$ 作一修补, 添加了 $(1-a_x)D_x$, 即是

$$q_x = \frac{D_x}{\bar{P}_x + (1-a_x)D_x}$$

我们姑且不论为何要将平均人口进行这样的修补? 这样为何就成为了死亡概率? 寇尔只将上式变形, 即得其死亡概率公式

$$q_x = \frac{m_x}{1 + (1-a_x)m_x}$$

从实证看, 寇尔法与其特殊情况(寿命成数 $a_x=1/2$)的伐尔法, 与另行思路的指数法, 这“三法”计算的结果基本一样, 当然伐尔法最为简单, 指数法计算也很容易, 但寇尔法需要确定该人口系统的寿命成数 a_x , 我们认为这比确定粗死亡率的难度还大, 而且更增加了预期寿命的变数。

现行的编制人口生命表, 都是把实际统计数据计算得到的粗死亡率转化成所谓的死亡概率。我们认为年龄别粗死亡率, 是直接用实际行动统计, 年龄别死亡人数与该年龄组平均人数计算得到的, 从概率论来讲它仅是死亡频

率的概念,直接将死亡频率改头换面成为所谓死亡概率,有悖于频率与概率的关系,也有改变死亡结构之嫌。导致推算出的预期寿命的真实性、可靠性、可用性遭到质疑。

4 结构法编制生命表的探讨

基于上述各种方法编制人口生命表的实证分析与质疑,其根本问题有三:一是现行的由年龄别粗死亡率推导出的年龄别死亡概率,在一定程度上改变了该人口系统的死亡结构,有别于真实的年龄别死亡结构;二是在同时出生的一群人的假定下,计算出这群人的平均预期寿命;三是以这假设一群人的预期寿命就直接作为现实人口系统的预期寿命。我们探讨了直接用真实的年龄别死亡人数的结构编制真值生命表,称为“死亡结构法”。死亡结构法编制生命表的基本思想是,所研究的人口系统0岁的人群,按其所在现实人口系统中的年龄别死亡结构,编制其生命表,计算出这群0岁人口的平均预期寿命。

4.1 “死亡结构法”编表的步骤

(1)由人普等人口调查取得所编人口系统的年龄别死亡人数的实际数据;

(2)计算死亡结构:年龄别死亡人数占总死亡人数的比例;

(3)计算年龄别尚存人数 l_x :

以该人口0岁的人数作为初始尚存人口 l_0 :如,此次取0岁平均人口数(人)为

$$l_0^{\text{女性}}(2000)=5748208, l_0^{\text{男性}}(2000)=6896316, l_0^{\text{全体}}(2000)=12644523$$

$$l_0^{\text{女性}}(2010)=7194841, l_0^{\text{男性}}(2010)=8586163, l_0^{\text{全体}}(2010)=15781004$$

并按其年龄别死亡结构比例,计算年龄为1,2,..., $\omega-1$ 的尚存人数 l_x ;

(4)平均生存年数 L_x 、平均生存总人年数 T_x 和平均预期寿命 e_x ,与一般编制生命表类似,分别按式(7)、(8)和(9)计算。中国2000年和2010年普查数据,用“死亡结构法”编制的人口生命表(表略)。

4.2 “死亡结构法”编制的生命表的统计检验

生命表元素存在着随机性、比例性以及平均性,对生命表进行统计检验是必要的。由于我们编制了2000年和2010年不同时期的生命表,因而,检验有两个层次:一是不同时期(2000年和2010年)的全体对全体、男性对男性、女性对女性的差异性检验;二是同一时间的全体对男性、全体对女性、男性对女性的差异性检验。具体检验内容这里只分别对死亡结构和预期寿命,在其正态总体假设下,设定显著水平为 $\alpha=0.05$,用Z检验法,对其原假设差值为0,进行无差异检验,在Excel实现,检验结果拒绝原假设,即男性与女性及集体是有差异的。

4.3 “死亡结构法”与“三法”编制的生命表的比较

对于“死亡结构法”除用两普查年数据计算外,对其他有年龄别死亡人数的生命表,也进行了实证。计算结果表明,他与一般常用的伐尔法、寇尔法和指数法所得的预期寿命,结果相近(详见表2);特征也类同,他们都是2010年

生命表没有出现预期寿命矛盾现象,而2000年都出现了预期寿命矛盾现象:

女性是

$$e_0^F(2000)=72.92, e_1^F(2000)=76.36, e_2^F(2000)=75.73,$$

$$e_3^F(2000)=74.97,$$

$$e_4^F(2000)=74.15, e_5=73.29$$

1岁与0岁的预期寿命高出近4岁,一致到5岁的预期寿命都高于0岁。

男性是

$$e_0^M(2000)=71.36, e_1^M(2000)=73.15, e_2^M(2000)=72.45,$$

$$e_3^M(2000)=71.67$$

两性对比可看出,女性远远高于男性,而且持续的时间也比男性长。

全体是

$$e_0^{\text{全体}}(2000)=72.03, e_1^{\text{全体}}(2000)=74.53, e_2^{\text{全体}}(2000)=73.85,$$

$$e_3^{\text{全体}}(2000)=73.08, e_4^{\text{全体}}(2000)=72.26$$

究其原因主要是,0岁婴儿死亡比例远高于1岁,全体、男性、女性分别为:4.65%>0.44%,3.79%>0.41%,5.75%>0.47%,特别是女婴的死亡比例非常高,导致预期寿命矛盾现象的发生。这种矛盾现象是生命表特有现象,因为生命表既研究一个Cohort人群的生命过程,同时又研究他的死亡过程,生命函数是单调减少的,而死亡函数却是单调增加函数,因而生命表中就可能出现平均预期寿命随着人口年龄增长而上升的矛盾现象。

“三法”表的质量完全取决于死亡概率 q_x , q_x 取决于死亡率 m_x , m_x 是生命表的生命。而 m_x 决定于实际年龄别死亡人数 D_x 和相应的年龄别平均人数 $\bar{P}_x$ 两大要素。这样就带来两大问题:一是原始统计数据 D_x 和 $\bar{P}_x$,二是由 m_x 导出 q_x ,它们都直接影响着生命表的质量。长期来,众多人口学家,如寇尔等在 $\bar{P}_x$ 上潜心研究,最终也只是对 $\bar{P}_x$ 作了一些修补的创新工作。

“结构法”表的质量只取决于死亡人数 D_x 的真实、可靠、准确,无须考虑年龄别平均人数 $\bar{P}_x$ 。显然其质量优于“三法”表。

另外,由于粗死亡率与死亡结构比例概念不同,从下图1看出,在80岁以前它们基本重合,拟合得很好。在80岁以后由于粗死亡率逐渐增大,而死亡结构比例,由于平均人口逐渐减少,死亡人数当然也随之逐渐减少,故其占比也就逐渐减少而趋于零。因而形成一个剪刀差。

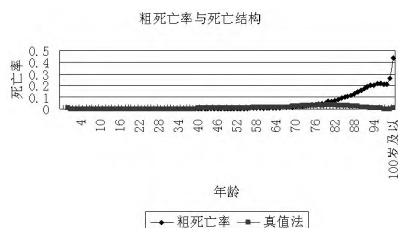


图1

5 结论

国外CPI偏差的测定方法及启示

雷怀英, 乔睿蕾, 贾文军
(天津工业大学 管理学院 300384)

摘要: CPI偏差是衡量价格指数能否准确度量实际价格变化程度的重要指标, 偏差大小是决定是否需要
进行价格指数编制方法改革的重要依据。文章对国外CPI偏差的测定进行了梳理和总结, 为我国CPI偏差的测
定提供了可借鉴的依据。

关键词: CPI偏差; 测定; 价格指数

中图分类号: F222.3 **文献标识码:** A **文章编号:** 1002-6487(2014)01-0007-04

0 引言

CPI是测量实际通货膨胀的重要指标, 是进行投资等经济决策的重要参考依据, 也是调整存贷款利率的重要依据, CPI还与工资、政府福利的指数化, 以及国家预算等都有很大的关系, CPI的高估(低估)将导致社会福利的实际增加(减少)。由于我国现阶段是以相对固定的消费品篮子编制的CPI, 反映的是不同时期固定篮子内商品和服务价格的变化, 该指数与居民的实际生活费用价格的变化存在一定的偏差, 国内外学者普遍将CPI与COLI(Cost of Live Index生活费用指数)之间的偏差定义为CPI偏差。

CPI偏差是衡量价格指数能否准确度量实际价格变化程度的重要指标, 偏差大小是决定是否需要价格指数编制方法改革的重要依据, 伴随着对价格指数的各种争议和评价, 学术界与相关统计部门开始了对价格指数的偏差进行定量的评估。20世纪90年代期间, 美国引发了

一场有关价格指数的争论, 参与者非常广泛, 争辩也较为激烈。多方评估的结果虽说不完全相同, 但争论的问题和结论对价格指数编制理论的完善和实践应用均具有深刻的意义, 既推动了价格指数偏差度量方法的完善, 也引起了政府部门以及学者对价格指数偏差调整问题的重视。

近年来我国社会各界对CPI准确性的质疑声也不断。本文通过对国外文献资料的认真梳理和总结, 力求为我国CPI偏差的准确度量及其调整提供可借鉴的依据。

1 固定篮子CPI指数与生活成本指数(COLI)及其偏差来源

国外研究将CPI与COLI(生活成本指数)之间的差异定义为CPI偏差, 为了进一步明确偏差的来源, 我们先看二者在概念界定上的差异。

设报告期和基期有 n 种商品和服务可供消费者选择, $t=0$ 代表基期, $t=1$ 代表报告期, 报告期和基期市场上产品

基金项目: 国家社科基金资助项目(11BTJ018)

作者简介: 雷怀英(1969-), 女, 山西苛岚人, 博士, 教授, 研究方向: 数理统计、价格指数。

用人口普查数据编制完全生命表是十分必要的, 是真实可靠、可信可用的, 今后的生命表编制应在此基础上, 再用我国年末逢“5”的全国人口1%的抽样调查和死亡、生育等专项调查数据编制逢“5”生命表, 逢“0”和逢“5”表结合编制我国逐年以及推测的生命表, 使之制度化、规范化、常态化, 成为无论保险业还是其他各业研究、使用的基准表。

编表类型上, 需编制男性和女性以及全体(平均)生命表, 需编制全国和省(市、区)等地区生命表, 需编制分民族表以及专业性如劳动就业、教育等生命表, 构成以蒋再华模型生命表为核心的中国人口生命表体系。

编表方法上, 可采用“死亡结构法”, 他理论有据, 简单易行。

- [1] 胡英. 中国分城镇乡村人口平均预期寿命探析[J]. 人口与发展, 2010, (2).
- [2] 李永胜. 人口统计学[M]. 成都: 西南财经大学出版社, 2002.
- [3] 黄荣清. 经验生命表与普通生命表[J]. 市场与人口分析, 2001, (5).
- [4] 张广裕. 甘南藏族自治州男性人口生命表和死亡分析[J]. 西北人口, 2009, (3).
- [5] Ansley J. Coale. The Growth and Structure of Human Population—A Mathematical Investigation[M]. Princeton: Princeton University Press, 1972.
- [6] J. H. Pollard. Mathematical Models for the Growth of Human Populations[M]. London: Cambridge University Press, 1973.
- [7] Nathan Keyfitz. Applied Mathematical Demography(2nd Edition) [M]. New York: Springer-Verlag, 1985.

(责任编辑/亦 民)

参考文献: