

# 中国人口预期寿命预测的 Double-Gap 方法及其改进<sup>\*</sup>

赵 明 方玉昕 王晓军

**【摘 要】**文章通过引入灰色预测 GM(1,1)模型改进 Double-Gap 方法对寿命差距进行建模,并将预测结果与传统 Double-Gap 方法、Lee-Carter模型和贝叶斯分层模型进行比较,以探寻适合中国人口预期寿命预测的科学方法。研究表明:(1)从建模原理看,Double-Gap 方法不仅能在建模中将中国人口预期寿命与人类最优表现预期寿命进行关联,还能避免 Lee-Carter 模型等间接预测法中死亡模式固定的缺陷,同时方法客观、可操作性较强。(2)从模型的拟合与短期预测效果看,改进 Double-Gap 方法具有较小的平均绝对百分误差,并能有效弥补 Lee-Carter 模型在人口预期寿命预测中出现的过拟合问题。(3)从模型长期预测效果看, Lee-Carter 模型和贝叶斯分层模型在预测中国人口预期寿命时较为保守,而 Double-Gap 方法预测的中国人口预期寿命能保持与最优表现预期寿命之间的合理变动趋势,并且改进 D-G 方法预测的出生人口平均预期寿命增长趋势的斜率值小于传统 D-G 方法,能够更好地反映出人口死亡率降低到一定程度后寿命增长减速的趋势。

**【关键词】**预期寿命 Double-Gap 方法 Lee-Carter 模型 贝叶斯分层模型

**【作 者】**赵 明 首都经济贸易大学金融学院,副教授;方玉昕 中国人民大学统计学院,硕士研究生;王晓军 中国人民大学统计学院,教授。

## 一、引 言

随着经济的增长和医疗水平的提升,世界上大部分国家经历了不同程度的人口死亡率下降和预期寿命提高,全球范围内的人口寿命延长已成为不可逆转的趋势(Wilmoth 等,2000)。人口寿命延长伴随着人口老龄化和长寿化,将会冲击经济主体的财务稳定性(Booth 等,2002)。此外,人口预期寿命是衡量全年龄人口死亡率变动的综合指标,也是反映人口质量的重要维度,科学判断人口预期寿命的变动趋势能为制定人口长期均衡

<sup>\*</sup> 本文为首都经济贸易大学北京市属高校基本科研业务费专项资金资助项目“2022 年校级青年学术创新团队建设——大数据背景下健康老龄化综合评价指标体系与精算测度研究”的阶段性成果。

发展战略提供依据。因此,本研究从梳理人口预期寿命的预测方法出发,在总结评价当前研究贡献与不足的基础上,对已有研究方法进行改进,并采用中国数据进行实证检验,以期为中国人口预期寿命预测的方法选择提供借鉴。

国内外关于人口预期寿命动态预测的研究主要分为两类:一是通过年龄别人口死亡率的预测间接获得人口预期寿命的预测值;二是直接对人口预期寿命建立预测模型。

首先,从年龄别人口死亡率的预测出发,通过生命表中各要素间的构造关系,可以间接计算预期寿命。Lee 等(1992)创建了随机死亡率模型,假设对数死亡率由相互独立的年龄效应和时间效应构成,模型具有待估参数少、拟合过程简单和预测结果稳健等优势,得到国内外学者的普遍应用。随着研究的深入, Lee-Carter 模型逐渐暴露出一些问题,对模型的完善与改进主要包括参数估计方法的完善(Lee 等,2001)、模型假设的放宽(Li 等,2013)、有限数据情况下的模型拓展(Li 等,2004)和死亡率模型年龄效应因子的动态改进(Li 等,2013)等方面。近年来,越来越多的研究发现,人口死亡率预测是一个系统过程,仅考虑单一人口群体的死亡率建模,在长期预测中会导致预测结果出现违背人口生物特征的交叉或背离(Li 等,2005)。因此,多人口视角下的死亡率建模逐渐成为新的研究热点,通过将两个或多个人口群体联合建模,考虑不同群体间死亡率变动趋势的一致性与差异,能够有效提升死亡率长期预测的合理性(Enchev 等,2017)。

其次,与人口死亡率模型的进展相比,直接对人口预期寿命建模的研究相对滞后,研究方法主要包括三类:一是时间序列 ARIMA 模型。从时间维度上看,人口预期寿命是一组时间序列数据,适用于 ARIMA 方法建模(Torri 等,2012)。由于 ARIMA 模型对数据的平稳性要求较高,各国研究中对女性人口预期寿命建立 ARIMA 模型较为普遍(Gao 等,2009),其原因是女性人口死亡率下降趋势领先于男性,其预期寿命增长趋势更具代表性。二是联合国的贝叶斯分层模型(BHM)。贝叶斯分层模型以预期寿命为研究对象,是一个带有非恒定漂移项的随机游走模型,漂移项是当前预期寿命的非线性函数,能够反映不同国家在不同寿命水平下预期寿命的增长速度,Raftery 等(2013)使用 Double-Logistic 函数表示这个增量。贝叶斯分层模型预测人口预期寿命,考虑了不同国家、不同性别之间人口预期寿命间的内在关联,但该方法较为复杂,应用不便。三是基于人类最优表现预期寿命的预测方法。Oeppen 等(2002)提出最优表现预期寿命的概念,即在医疗技术、公共卫生等领域做出最大努力后,全世界范围内的最高预期寿命水平。最优表现预期寿命的增长具有规律的线性趋势,因此可以通过对最优表现预期寿命进行建模,并在此基础上对目标国家预期寿命和最优表现预期寿命的差距进行建模分析。其中,极值理论是最优表现预期寿命建模的有效方法(Medford,2017);Double-Gap 方法(以下简称“D-G 方法”)给出了人口预期寿命预测的系统化框架(Pascariu 等,2018),被认为是近年来人口预期寿命预测的经典方法(Nigri 等,2021)。

目前针对人口预期寿命预测的研究仍存在以下不足:(1)当前以 Lee-Carter 模型(以下简称“L-C 模型”)为代表的死亡率预测方法,假设死亡率年龄模式不发生变化,但高寿命国家的经验显示,人口预期寿命的增长背后伴随着死亡率年龄模式的变化,因此通过人口死亡率的预测间接预测预期寿命的方法存在缺陷。(2)ARIMA 模型仅考虑单一国家或性别人口预期寿命的时间序列变动趋势,若不加以限制,在长期中人口预期寿命将会无限增长,违背人类生物规律。(3)贝叶斯分层模型考虑了不同国家和性别人口之间的内部关联,但模型方法复杂,且模型的先验信息设置具有较强的主观性,在经济和社会领域中的应用不够便利。可见,基于最优表现预期寿命的 D-G 方法,是研究人口预期寿命动态预测的较好选择,不仅能够在建模中系统地将不同人口群体预期寿命进行关联,还能够避免间接预测法中死亡模式固定的缺陷,同时方法可操作性较强。然而,传统 D-G 方法存在一定缺陷,其对寿命差距所建立的时间序列模型需要满足平稳性要求和白噪声假设,且需要较大的样本量。此外,不同国家人口的基因特征、医疗水平、生活环境等因素差异较大,未来寿命差距的变动趋势具有较大的不确定性,而传统 D-G 方法对寿命差距存在较强的主观性假设。为了弥补以上缺陷,本文通过引入灰色预测的 GM(1,1)模型对 D-G 方法进行改进。其中,GM(1,1)模型无需假设特定概率分布,且能够减少时间序列的随机性,在小样本下预测效果较为理想,适合中国人口预期寿命的建模与预测。

## 二、Double-Gap 方法及其改进

### (一) Double-Gap 方法的基本原理

基于最优表现预期寿命的预期寿命预测方法,最早由 Oeppen 等(2002)提出,由 Torri 等(2012)的研究不断发展,并由 Pascariu 等(2018)将建模思想系统化,命名为 D-G 方法。运用 D-G 方法预测目标国家人口预期寿命的思路包括:(1)确定最优表现预期寿命,并对此建立线性模型。确定最优表现预期寿命的目的是掌握人类预期寿命的变动规律,Pascariu 等(2018)给出了最优表现预期寿命的选择标准,即人口预期寿命越高,意味着死亡率下降程度越大,此时的预期寿命会更加接近人类极限水平,则其变动趋势越具有代表性。在本文所使用的样本数据周期内(1994~2019 年),日本女性人口预期寿命一直保持世界最高水平,因此用其代表最优表现预期寿命。(2)计算目标国家女性预期寿命与最优表现预期寿命间的差距(Gap1),并对差距数据建立经典时间序列模型。该步骤利用统计学中差分的思想,通过对数据进行差分处理,能够提高时间序列的平稳性,更加适于建立经典的 ARIMA 模型。通过第一个差距的预测值和第一步中最优表现预期寿命的预测值,可以得到目标国家女性人口预期寿命的预测值。(3)利用经典时间序列模型预测目标国家男女预期寿命之间的差距(Gap2),并结合第二步中目标国家女性人口预期寿命的预测值,可以获得目标国家男性人口预期寿命的预测值。(4)预测区间由

多元正态分布构成,其中均值为零,协方差矩阵由前几步拟合 3 个模型的残差构成。综上所述,D-G 方法可以用以下两个核心模型进行概述。

第一,国家  $k$  的  $x$  岁女性在  $t$  年的预期寿命( $e_{x,t,k}^f$ )预测值,可以通过最优表现预期寿命( $e_{x,t}^b$ )预测值减去第一差距( $D_{x,t,k}$ )预测值得到,即:

$$e_{x,t,k}^f = e_{x,t}^b - D_{x,t,k} \quad (1)$$

第二,国家  $k$  的  $x$  岁男性在  $t$  年的预期寿命( $e_{x,t,k}^m$ )预测值,可以通过该国家女性预期寿命( $e_{x,t,k}^f$ )预测值减去第二差距( $G_{x,t,k}$ )预测值得到,即:

$$e_{x,t,k}^m = e_{x,t,k}^f - G_{x,t,k} \quad (2)$$

## (二) 方法改进与中国应用

D-G 方法构建了人口预期寿命预测的框架体系,但其部分假设仍有改进的空间。一方面,对中国女性预期寿命与最优表现预期寿命间的差距(Gap1)建模,由于不同国家人口的基因特征、医疗水平、生活环境等因素差异较大,未来的变动趋势具有较大的不确定性,无论是采用线性模型(Pascariu 等,2018),还是采用均值回复过程或几何布朗运动建模(Torri 等,2012),在描述未来寿命差距之间的变动趋势时均存在主观性。因此,本文引入灰色预测系统的 GM(1,1)模型,构建一个以时间为变量的连续微分方程,该预测方法能够弱化人口预期寿命变动中的未知因素,强化已知因素的影响。另一方面,对中国男女人口预期寿命差距(Gap2)建模,本文采用分段的截断数据自回归模型。由于中国男女人口面临的经济水平、医疗技术等背景因素相同,二者之间的寿命差异主要体现在性别上的生物特征,未来的变动趋势具有较强的一致性,因此分段的截断数据自回归模型能够很好地刻画这一特征,使预期寿命的性别差距预测值收敛为固定数值。基于以上方法改进,本文对中国人口预期寿命建立动态预测模型的步骤如下。

第一,选取日本女性人口为最优表现预期寿命,并对其建立线性模型,表达式为:

$$e_x^B(t) = \alpha_{x0} + \alpha_{x1}t + \varepsilon_x(t) \quad (3)$$

其中, $\alpha_{x0}$  是线性模型的截距项, $\alpha_{x1}$  是线性模型的斜率项。利用式(3)即可得到最优表现预期寿命预测值  $e_x^{B*}(t)$ 。

第二,对中国女性人口预期寿命与最优表现预期寿命的差距建立 GM(1,1)模型。令预期寿命差距为:

$$D_x(t) = e_x^B(t) - e_x^f(t) \quad (4)$$

对  $D_x(t)$  生成一阶累加序列  $D_x^{(1)}(t)$ ,该序列满足下列一阶微分方程,即:

$$dD_x^{(1)}(t)/dt + aD_x^{(1)}(t) = b \quad (5)$$

采用上述方法对累加序列进行预测后,还需对结果累减还原,得到预期寿命差距的预测值  $D_x^*(t)$ 。进一步可以得到中国女性人口预期寿命的预测值,即:

$$e_x^{f*}(t) = e_x^{B*}(t) - D_x^*(t) \quad (6)$$



第三,对中国男女人口预期寿命差距建模,令预期寿命差距为:

$$G_x(t)=e_x^f(t)-e_x^m(t) \tag{7}$$

在此基础上,对中国男女人口预期寿命差距建立分段的截断数据自回归模型,即:

$$G_x(t)=\begin{cases} \beta_0+\beta_1G_x(t-1)+\beta_2G_x(t-2)+\beta_3[e_x^f(t)-\tau]_++\varepsilon_x(t), & \text{if } e_x^f(t)\leq A \\ G_x(t-i)+\varepsilon'_x(t), & \text{if } e_x^f(t)>A \end{cases} \tag{8}$$

该截断数据自回归模型分为两段,第一段是线性模型,第二段则是不含漂移项的随机游走模型。若女性预期寿命小于寿命上限  $A$ ,则性别差为前两年的性别差  $G_x(t-1)$ 、 $G_x(t-2)$  和一个与女性预期寿命有关的附加项  $[e_x^f(t)-\tau]_+$  的线性组合,  $\tau$  是预期性别差停止扩大时的预期寿命水平;当女性预期寿命超过上限  $A$  时,随机游走模型用于进一步拟合和预测性别差距。其中,  $\beta_0$ 、 $\beta_1$ 、 $\beta_2$ 、 $\beta_3$ 、 $A$  和  $\tau$  均为待估计参数。将参数估计结果带入模型,可以得到预期寿命差距的预测值  $G_x^*(t)$ 。

基于上述改进方法得到的预期寿命差距的预测值  $G_x^*(t)$ ,可以得到中国男性人口预期寿命的预测值,即:

$$e_x^{m*}(t)=e_x^{f*}(t)-G_x^*(t) \tag{9}$$

第四,与 D-G 方法相同,预测区间由多元正态分布构成,其中均值为零,协方差矩阵由前三步拟合 3 个模型的残差构成。

### 三、数据来源、假设与模型拟合

#### (一) 数据来源与假设

本文选取中国国家统计局公布的 1994~2019 年全国分年龄、分性别人口死亡率数据和日本女性人口平均预期寿命数据进行研究。其中,中国人口死亡率数据来自《中国人口统计年鉴》和《中国人口和就业年鉴》;日本女性人口预期寿命数据来自人类死亡率数据库(HMD)。由于 HMD 数据质量较高,日本女性人口预期寿命数据可以直接应用。然而,在所选取的中国人口死亡率数据期间,只有 2000 和 2010 年为普查数据,且存在一定程度漏报,其他年份的抽样调查数据也存在统计误差。因此,本文根据研究需要和数据特征进行以下处理与假设:(1)插补年龄别死亡率的缺失数据。中国国家统计局公布的各年度年龄别死亡率数据中,人口变动抽样调查的抽样比为 1‰,而死亡是一个小概率事件,会出现抽样死亡人数为 0 的情况(王广州,2021)。由于女性人口死亡率低于男性,这种情况在女性人口中更为常见。因此,本文采用线性插值法对人口死亡率为 0 的数据在年龄维度上进行填补。(2)修正中国新生婴儿死亡率和高龄人口死亡率数据。中国人口死亡率数据有效积累不足、抽样调查年份误差较大,甚至普查年份新生婴儿和高龄人口死亡人数也存在一定程度的漏报(郭志刚,2011)。大量研究表明,中国第三次人口普查的死亡数据质量较高(孙福滨等,1993),在此基础上可以对第五次人口普查和第六次

人口普查的死亡率数据进行检验与修正。本文采用 Brass-Logit 生命表法,以 1981 年死亡率为基准,修正各年份的 0~4 岁低龄人口和 60~99 岁高龄人口死亡率。其中,对 0~4 岁人口的修正使用 5~59 岁数据建模,对 60~99 岁人口死亡率的修正使用 30~59 岁数据建模(王金营,2013)。(3)假设死亡人口在 1 年内服从均匀分布假设,并将各年份的死亡率( $m_{x,t}$ )转化为死亡概率( $q_{x,t}$ ),其公式为: $q_{x,t}=m_{x,t}/(1+0.5m_{x,t})$ 。(4)构造中国人口动态生命表。以 10 万人口为基准,根据调整后的年龄别人口死亡率数据,对各年龄的死亡人口数进行调整,构造中国人口动态生命表。基于动态生命表,可以得到平均预期寿命的计算公式: $e_{x,t}=T_{x,t}/l_{x,t}$ 。其中, $l_{x,t}$  为  $t$  年  $x$  岁的生存人数, $T_{x,t}$  为  $t$  年  $x$  岁的累积生存人年数。

(二) 中国人口预期寿命调整与计算

根据上述的数据处理与假设,可以构造 1994~2019 年中国人口动态生命表,并可以通过生命表函数计算各年男性和女性人口预期寿命。

1. 人口平均预期寿命计算值比较

1994 年以来,中国国家统计局公布了 2000、2005、2010 和 2015 年的男女 0 岁人口预期寿命,即出生人口平均预期寿命。为了检验人口死亡率数据调整的合理性,本文将计算结果与中国国家统计局公布的历年人口预期寿命和王金营(2013)修正数据进行比较,结果如表 1 所示。由表 1 可见,通过本文的基本假设和数据处理方法计算得到的人口预期寿命值与王金营(2013)修正结果较为接近,并且男性预期寿命低于国家统计局公布数据,女性预期寿命高于国家统计局公布数据。

2. 中国人口预期寿命的调整结果

为了比较不同年龄人口预期寿命的变动趋势,本文选取 0 岁和 60 岁两个重要年龄为代表,分析预期寿命的变动趋势。由表 2 可见,无论是出生人口平均预期寿命,还是 60 岁时平均预期寿命,均随时间的推移呈线性增长趋势。总体来说,1994~2019 年男性出生人口平均预期寿命年均增长 0.20 岁,女性出生人口平均预期寿命年均增长 0.33 岁;男性 60 岁时平均预期寿命年均增长 0.09 岁,女性 60 岁时平均预期寿命年均增长 0.17 岁。可见,男性出生人口平均预期寿命整体增长幅度小于女性,男性 60 岁时平均预期寿命增长幅度也小于女性。这意味着 1994~2019 年女性人口预期寿命与男性人口预期寿命的差距不断增大,其中出生人口平均预期寿命的差距从 1994 年的 4.11 岁增加到 2019 年的 7.24 岁、60 岁时平均预期寿命的差距从 1994 年

表 1 中国人口平均预期寿命修正值的比较 岁

	2000 年	2005 年	2010 年	2015 年
本文修正				
男性	69.26	70.62	71.89	73.08
女性	75.02	76.91	78.61	80.14
王金营(2013)				
男性	68.76	-	71.58	-
女性	74.47	-	78.26	-
统计局公布				
男性	69.63	70.83	72.38	73.64
女性	73.33	75.25	77.37	79.43

表 2 1994~2019 年中国人口平均预期寿命修正值 岁

年份	男性 $e_0$	男性 $e_{60}$	女性 $e_0$	女性 $e_{60}$	年份	男性 $e_0$	男性 $e_{60}$	女性 $e_0$	女性 $e_{60}$
1994	68.63	16.56	72.74	19.82	2007	71.17	17.49	77.81	22.05
1995	67.68	16.29	73.07	19.98	2008	70.30	17.24	78.34	22.24
1996	67.99	16.29	73.26	19.99	2009	71.98	17.74	78.76	22.62
1997	68.44	16.56	74.26	20.68	2010	71.90	17.77	78.71	22.42
1998	68.80	16.88	74.67	20.81	2011	71.71	17.61	79.00	22.67
1999	69.42	16.75	74.42	20.65	2012	72.08	17.74	78.80	22.54
2000	69.15	16.73	74.96	20.65	2013	72.25	17.99	79.28	22.86
2001	69.42	16.75	74.42	20.65	2014	72.66	18.17	79.87	23.42
2002	69.44	16.82	76.24	21.33	2015	74.20	18.85	81.00	23.69
2003	69.99	17.13	75.41	21.05	2016	74.24	18.92	80.96	23.78
2004	70.28	17.20	75.88	21.13	2017	74.02	18.79	80.97	23.90
2005	69.53	16.89	76.74	21.45	2018	73.82	18.57	80.54	23.67
2006	71.05	17.55	77.08	21.56	2019	73.70	18.80	80.94	24.17

的 3.26 岁增加到 2019 年的 5.37 岁。

此外,历史数据表明男女人口预期寿命的差距不断增加,若将该变动趋势作为假设带入预测模型,将会导致未来中国人口预期寿命的性别差距进一步增大,显然不符合人口生物特征,因此在对性别差距进行建模时不仅要考虑历史数据特征,更要考虑长期变动趋势的合理性。

(三) 模型参数拟合结果

本文提出 D-G 模型改进方法,针对中国女性人口与最优表现预期寿命之间的第一差距(Gap1)建立 GM(1,1)模型,针对中国男性和中国女性预期寿命之间的第二差距(Gap2)建立分段的截断数据自回归模型。

首先,对于 GM(1,1)模型预测精度的检验常用后验差方法,该方法包含两个指标:后验差比值(C)和小概率误差(P),其中 C 值越小表明残差的方差相对越小,P 值越大表明残差与残差均值之差的绝对值相对越小。当前被学术界普遍认可和广泛采用的预测精度数值标准为:当 P 大于 0.80、C 小于 0.50 是较好的拟合结果;当 P 大于 0.95、C 小于 0.35 是很好的拟合结果。本文对第一差距建立的 GM(1,1)模型预测精度如表 3 所示,其中女性出生人口平均预期寿命的差距预测精度等级为很好( $P=1.00, C=0.25$ ),60 岁时平均预期寿命的差距预测精度等级为较好( $P=0.92, C=0.25$ )。上述结果显示,GM(1,1)模型能够较好地拟合中国女性人口与最优表现预期寿命之间的第一差距。此外,针对中国男性和女性预期寿命之间的第二差距建立分段的截断数据自回归模型,其参数拟合结果如表 4 所示。

参数  $\beta_0$  表示截距水平,可以解释为男性和女性之间预期寿命的生物差距,参数值

体现出生人口平均预期寿命的性别差大于 60 岁时平均预期寿命的性别差。参数  $\beta_1$  和  $\beta_2$  表示过去两年的性别差对当年性别差的影响,系数含义与传统的 AR(2)模型相同。参数  $\beta_3$  的估计值为负,表明性

别差呈收敛趋势, $\beta_3$  的绝对值给出了女性和男性预期寿命之间的收敛速度,参数估计结果表明出生人口平均预期寿命比 60 岁时平均预期寿命性别差收敛得更快。参数  $A$  是女性预期寿命的上限;由于目前难以对高龄性别差的未来模式进行假设(Raftery 等, 2013),所以当女性预期寿命超过上限  $A$  时,使用随机游走模型拟合性别差。参数  $\tau$  是预期性别差停止扩大时的预期寿命水平, $A$  和  $\tau$  的估计值均使用极大似然方法获得。

四、人口预期寿命预测的方法比较

在比较人口预期寿命预测方法时,首先对改进前后的 D-G 方法进行比较,并与间接预测的 L-C 模型进行比较,对于传统 D-G 方法和 L-C 模型分别采用 Pascariu 等(2018)、李志生和刘恒甲(2010)的方法进行拟合,其中 L-C 模型的参数估计采用加权最小二乘法。然后,结合联合国世界人口司的贝叶斯分层模型进一步验证预测结果的合理性。

(一) 样本内拟合优度比较

根据 3 种模型的参数估计结果,可以计算出 1995~2019 年人口预期寿命的样本内拟合值。通过 MAPE(平均绝对百分误差)检验 3 种方法的拟合效果,结果如表 5 所示。对于女性出生人口平均预期寿命,改进 D-G 方法拟合结果具有最小的 MAPE 值,拟合效果最优;传统 D-G 方法和 L-C 模型的拟合结果较为接近。对于男性出生人口平均预期寿命,传统 D-G 方法拟合结果具有最小的 MAPE 值,拟合效果最优;改进 D-G 方法和 L-C 模型的拟合结果较为接近。可见,对于出生人口平均预期寿命,直接预测法(D-G 方法)比间接预测法(L-C 模型)在拟合优度上具有一定优势,但整体而言二者的差异较小。此外,对于女性 60 岁时平均预期寿命和男性 60 岁时平均预期寿命,L-C 模型拟合结果具有相对较小

表 3 GM(1,1)模型预测精度

第一差距	P	C	预测精度等级
$D_0(t)$	1.00	0.25	很好
$D_{60}(t)$	0.92	0.50	较好

表 4 分段的截断数据自回归模型参数估计结果

性别差	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$A$	$\tau$
$G_0(t)$	2.210	0.244	0.432	-0.034	88.458	82.222
$G_{60}(t)$	0.553	0.473	0.430	-0.018	28.026	25.916

表 5 不同方法下样本内拟合结果比较

性别	直接预测方法		间接预测方法
	改进 D-G 方法	传统 D-G 方法	L-C 模型
女性 $e_0$	0.465	0.509	0.506
男性 $e_0$	0.681	0.643	0.685
女性 $e_{60}$	0.958	0.942	0.900
男性 $e_{60}$	1.428	1.120	1.050



的 MAPE 值;改进 D-G 方法和传统 D-G 方法的拟合结果较为接近,略大于 L-C 模型拟合结果。由于本文修正 60 岁以上人口死亡率时采用的方法与 L-C 模型类似,均是一种线性方法,因此会提升 L-C 模型的拟合优度,而拟合优度的提升是由模型结构带来的,还需进一步通过样本外预测评价模型的效果。

(二) 历史期内回溯预测效果比较

以 1994~2009 年人口预期寿命作为样本值,预测 2010~2019 年人口预期寿命,并使用现有的 2010~2019 年人口预期寿命实际值与不同方法的预测值进行比较。通过 MAPE 指标比较 3 种方法的预测值与实际值之间的误差(见表 6)。无论是男性还是女性,也无论是出生人口平均预期寿命还是 60 岁时平均预期寿命,改进 D-G 方法的 MAPE 值均显著小于传统 D-G 方法和 L-C 模型方法。由此可见,改进 D-G 方法的短期预测效果显著优于传统 D-G 方法和 L-C 模型方法。此外,对比传统 D-G 方法和 L-C 模型方法的 MAPE 值可见,间接预测的 L-C 模型方法在短期预测中的表现优于直接预测的传统 D-G 方法,但通过本文的改进能够提高直接预测法的短期预测效果。结合

表 6 不同方法下历史回溯期预测结果比较

性别	直接预测方法		间接预测方法
	改进 D-G 方法	传统 D-G 方法	L-C 模型
女性 $e_0$	0.516	1.196	0.585
男性 $e_0$	0.894	1.566	1.149
女性 $e_{60}$	0.905	1.897	0.946
男性 $e_{60}$	1.591	1.762	2.544

表 5 和表 6 看,L-C 模型在 60 岁人口预期寿命中的拟合优度表现出色,但短期预测效果不如改进 D-G 方法,意味着 L-C 模型存在过拟合问题,改进 D-G 方法能够较好地解决过拟合问题带来的预测陷阱。

(三) 长期预期效果比较

以 1994~2019 年的人口预期寿命作为样本值,预测 2020~2035 年中国人口预期寿命,结果如图 1 和图 2 所示。首先,将本文预测结果与第七次人口普查结果进行比较。采用改进 D-G 方法得到的 2020 年男性出生人口平均预期寿命均值预测结果为 74.88 岁,95%置信区间下界和上界分别为 73.10 岁和 76.66 岁;女性出生人口平均预期寿命均值预测结果为 81.75 岁,95%置信区间下界和上界分别为 80.99 岁和 82.51 岁。李睿等(2021)的研究显示,第七次全国人口普查漏报率仅为 0.05%,数据质量较高,其中男性和女性出生人口平均预期寿命分别为 75.37 岁和 80.88 岁。本文预测结果与第七次全国人口普查结果较为接近。图 1 对不同方法的预测结果进行比较,从中可以看出,对直接预测方法中的改进 D-G 方法和传统 D-G 方法的预测结果进行比较,无论是男性还是女性,改进 D-G 方法预测的出生人口预期寿命变动趋势的斜率均小于传统 D-G 方法。其原因是人口全年龄段死亡率随时间推移降低到一定程度后,人口预期寿命增长的速度将有所减缓,改进 D-G 方法能够有效反映这一趋势,因此出生人口预期寿命预测结果变

动趋势的斜率小于传统 D-G 方法。对直接预测方法和间接预测方法(L-C 模型)的结果进行比较,无论是男性还是女性,间接方法 L-C 模型预测的人口预期寿命的增长趋势较为平缓,即预期寿命预测值显著低于直接预测方法,意味着在长期预测中 L-C 模型预测的人口死亡率相对较高,预期寿命值则相对较低。本文认为直接预测方法结果较为理想,因为改进 D-G 方法预测值与最优表现预期寿命预测值呈一致性的变动趋势,即中国女性人口预期寿命的变动趋势在参照人类最高寿命的趋势变动,而 L-C 模型预测的中国女性出生人口平均预期寿命增长速度则显著低于人类最高寿命增长速度、缺乏与最优表现预期寿命之间的关联。随时间推移,人口死亡年龄模式在不断转变,但 L-C 模型将人口死亡率年龄效应参数设定为常数,实际应用中普遍存在高估人口死亡率、低估人口平均预期寿命的问题(Plat,2011),而改进 D-G 方法预测的较高人口平均预期寿命值能够弥补 L-C 模型预测结果偏保守的问题。

图 2 显示,无论男性还是女性,改进 D-G 方法预测的 60 岁人口预期寿命变动趋势的斜率大于传统 D-G 方法,这一特征与出生人口平均预期寿命的预测有差异。其原因在于,随着中低年龄段人口死亡率降低到一定程度后,未来老年人口死亡率下降将成为人口预期寿命提升的主要动因(Wilmoth 等,2000),改进 D-G 方法能够有效反映这一趋势。

(四) 与联合国人口司预测结果比较

将本文预测结果与联合国《世界人口展望(2019)》的预测结果进行比较。由于联合国发布的人口预期寿命数据是 5 年的平均数,为了增强可比性,将本文的预测结果处理后列于表 7 中。联合国预测中国人口预期寿命时,对历史数据并未进行过多的调整,历史期的人口预期寿命整体与中国国家统计局公布的结果接近。由于历史数据存在差异,必然会导致不同方法下的预测结果不能直接比较,因此本文通过比较预期寿命的增长幅度,以反映不同方法在长期预测中的合理性。

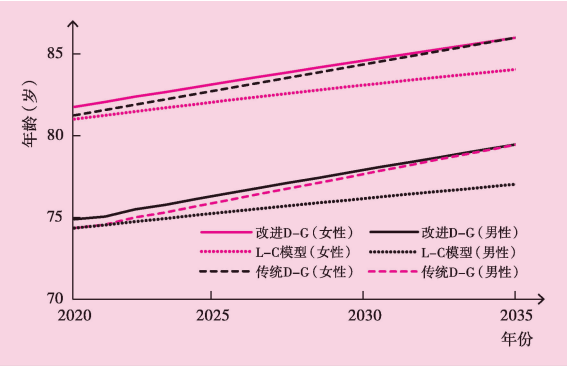


图 1 2020~2035 年中国出生人口平均预期寿命预测结果比较

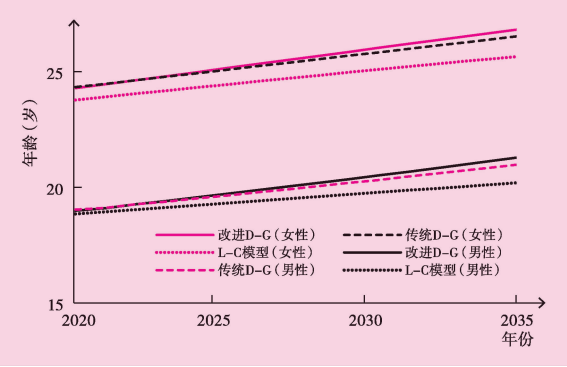


图 2 2020~2035 年中国 60 岁时平均预期寿命预测结果比较

表 7 不同方法下人口预期寿命预测结果比较 岁

年 份	最佳寿命	女性人口平均预期寿命( $e_0$ )				男性人口平均预期寿命( $e_0$ )			
		改进 D-G	传统 D-G	L-C 模型	BHM	改进 D-G	传统 D-G	L-C 模型	BHM
2020~2025	88.40	82.37	81.90	81.47	79.73	75.46	74.98	74.74	75.36
2025~2030	89.29	83.86	83.54	82.57	80.47	77.09	76.75	75.69	76.26
2030~2035	90.19	85.29	85.17	83.57	81.17	78.67	78.54	76.59	77.17

注: BHM 为联合国采用分层贝叶斯方法的预测结果。

从表 7 中 BHM 方法的预测结果可以看出,2020~2035 年男性出生人口平均预期寿命年均增长 0.18 岁、女性出生人口平均预期寿命年均增长 0.14 岁,该结果显著小于 1995~2019 年中国男性和女性出生人口平均预期寿命的增长幅度(男性出生人口平均预期寿命年均增长 0.24 岁、女性出生人口平均预期寿命年均增长 0.32 岁)。从表 7 中 L-C 模型的预测结果可见,2020~2035 年男性出生人口平均预期寿命年均增长 0.18 岁、女性出生人口平均预期寿命年均增长 0.21 岁,该结果与 BHM 方法的预测结果较为接近,其中女性人口的寿命增长幅度略有提高。BHM 方法和 L-C 模型预测的未来人口预期寿命增长幅度均显著小于中国历史数据中的寿命增长幅度。由于 L-C 模型在应用中普遍存在的高估人口死亡率、低估人口预期寿命问题(Plat, 2011),因此 BHM 方法在预测中国人口预期寿命时可能也存在预测结果过于保守的问题。

由表 7 中的最优表现预期寿命可见,2020~2035 年最优表现预期寿命年均增长 0.18 岁。这意味着,随着时间的推移,最优表现预期寿命呈线性增长趋势。由于最优表现预期寿命是人类最高的预期寿命,历史数据表现为显著的线性增长趋势,当前也并无证据表明该趋势的终止,因此预测结果具有稳健性。在最优表现预期寿命的基础上,可以得到改进 D-G 方法和传统 D-G 方法预测的中国人口预期寿命。2020~2035 年,改进 D-G 方法预测的男性和女性出生人口平均预期寿命年均分别增长 0.32 和 0.29 岁;传统 D-G 方法预测的男性和女性出生人口平均预期寿命年均分别增长 0.36 和 0.33 岁。可见,两种方法预测的未来人口预期寿命增长幅度均能与最优表现预期寿命的变动趋势保持一致,其中改进 D-G 方法预测的人口预期寿命增长幅度与中国历史寿命变动幅度较为接近,而传统 D-G 方法预测的未来预期寿命增长幅度相对历史增速偏高。此外,改进 D-G 方法预测的中国女性人口预期寿命增速大于 Oeppen 等(2002)研究的 1840~2000 年最优表现预期寿命增速(年均增长 0.25 岁)。由于本文预测的是 2020~2035 年,未来中国在经济、医疗技术和公共卫生等方面均会优于 1840~2000 年同期水平,且中国人口平均预期寿命与最优表现预期寿命之间尚存在一定差距,因此本文得出的相对较高的寿命增速具有合理性。

图 3 显示,随着时间的推移,改进 D-G 方法和传统 D-G 方法预测结果表明中国女性出生人口平均预期寿命与最优表现预期寿命之间的差距呈不断减小的趋势,但改进

D-G 方法的下降趋势更加平缓,反映出不同国家间寿命差距的缩小是一个长期且缓慢的过程;L-C 模型预测结果表明中国女性人口寿命与最优表现预期寿命的差距呈水平趋势,BHM 模型预测结果表明中国女性人口寿命与最优表现预期寿命的差距呈不断增加的趋势。无论是水平趋势还是增加趋势,这样的预测结果显然过于保守。随着国际经济一体化的发展,世界各国的经济水平与医疗技术之间的差距不断减小,不同国家之间的人口寿命差距也将不断缩小。

图 4 显示,随着时间的推移,改进D-G方法和传统 D-G 方法预测结果表明中国女性人口寿命与男性的差距呈先减少并不断收敛的趋势,两种方法的预测结果较为接近,均由 2020~2025 年的 6.90 岁逐渐下降到2030~2035 年的 6.61 岁,并呈逐渐收敛趋势;L-C 模型预测结果表明中国女性人口寿命与男性的差距不断增长,该结果与历史数据中的性别差距的变动趋势一致,但该趋势在长期中难以保持不变,反映出 L-C 模型存在过拟合问题;BHM 模型预测结果表明中国女性人口寿命与男性的差距呈线性递减的趋势,若将时间向外推移,按照该趋势,中国男女人口预期寿命的差距将会缩小为零,该结果忽略了男女人口基因差异导致的寿命差距。因此,基于 D-G 方法预测的中国男女人口预期寿命差距的变动趋势,能够吻合最优表现预期寿命国家的人口预期寿命性别差距先减小并逐渐收敛的经验结果。

综合样本期内的拟合优度、历史期内回溯预测效果和长期预测效果,可见改进 D-G 方法不仅能有效避免出现 L-C 模型的过拟合问题,还能提升短期内的预测效果,甚至弥补 L-C 模型和 BHM 方法预测结果较为保守的问题,可作为一种适合中国人口预期寿命预测的科学方法。

### 五、结论与讨论

本文通过对人口预期寿命预测的 D-G 方法进行改进,并将预测结果与传统 D-G 方

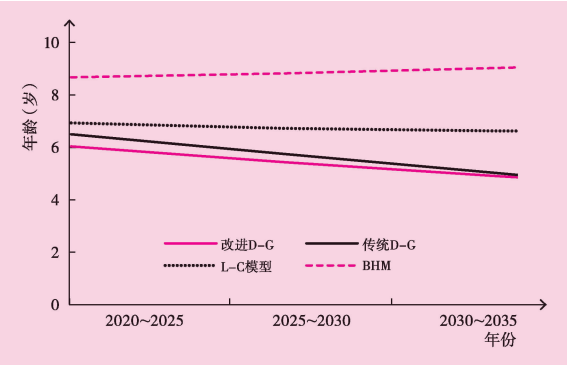


图 3 2020~2035 年中国女性人口与最优表现预期寿命差距

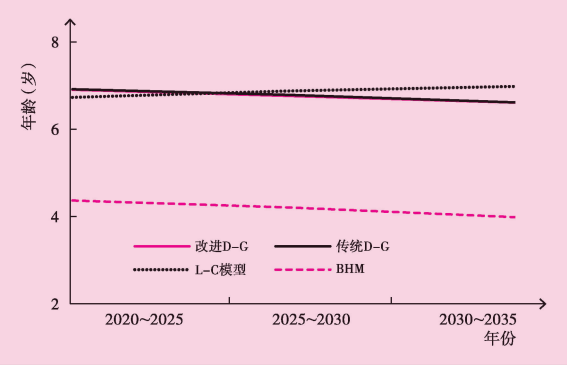


图 4 2020~2035 年中国男性与女性人口预期寿命差距



法、L-C 模型和联合国 BHM 方法进行比较,探究适合中国人口预期寿命预测的方法,以便合理判断未来人口预期寿命变动趋势和增长态势,并为促进人口长期均衡发展、实施积极应对人口老龄化战略提供理论依据。研究结果表明:(1)从建模原理看,基于最优表现预期寿命的 D-G 方法,不仅能在建模中将中国人口预期寿命与人类寿命经验进行关联,还能够避免 L-C 模型等间接预测法中死亡模式固定的缺陷,同时方法客观、可操作性较强。通过引入灰色预测 GM(1,1)模型改进 D-G 方法对寿命差距进行建模,能够更好地捕获人口预期寿命的变动趋势。(2)从模型的拟合与短期预测效果看,无论是间接预测的 L-C 模型还是直接预测的 D-G 方法,均能够获得较低的 MAPE 值,对中国人口预期寿命的拟合不存在明显差异。然而,改进 D-G 方法能够显著提升短期预测效果,并能有效弥补 L-C 模型在 60 岁时平均预期寿命预测中出现的过拟合问题。(3)从模型长期预测效果看,L-C 模型和 BHM 方法在预测中国人口预期寿命时较为保守,而改进 D-G 方法预测的中国人口预期寿命维持了与最优表现预期寿命差距合理的变动趋势,更加符合人的生物规律。

与从年龄别人口死亡率出发间接预测人口预期寿命的方法和联合国世界人口司的贝叶斯分层模型的复杂方法相比,D-G 方法是直接从人口预期寿命出发建立预测模型的有效方法,该方法简洁易懂、可操作性较强。针对人口预期寿命预测的 D-G 方法的研究,未来仍具有较大的拓展空间。若未来最优表现预期寿命的线性趋势发生改变,在 D-G 方法中需要对最优表现预期寿命的建模方法进行改进,以此为基础的中国女性和男性人口预期寿命也会随着最优表现预期寿命的变动呈现新的趋势。因此,本文建议在使用改进 D-G 方法预测中国人口预期寿命时,应根据每年更新的样本数据进行动态预测,这样能更好地发挥数据驱动模型的优势。然而,本文还存在一定的不足。首先,本文以第三次人口普查的年龄别死亡率为基础对样本数据进行修正,尽管学界普遍认为第三次人口普查数据质量较高,但对其数据准确度的评估仍有待进一步研究。其次,本文修正各年份低龄和高龄人口死亡率所采用的 Brass-Logit 方法,假设人口死亡模式不发生变化,未来基于第七次人口普查结果对历史数据的再评估,将有助于挖掘中国人口死亡模式转变规律,为数据修正提供更全面的参数信息。最后,尽管李睿等(2021)研究显示第七次人口普查漏报率已整体达到国际上公认的低水平,但通常情况新生儿和高龄人口死亡漏报较为集中,将会对人口预期寿命产生影响,因此对第七次人口普查中年龄别死亡率数据的评估有待进一步研究。

#### 参考文献:

1. 郭志刚(2011):《六普结果表明以往人口估计和预测严重失误》,《中国人口科学》,第6期。
2. 李睿等(2021):《七普实践与经验:普查数据质量是如何保障和提高的?》,《人口研究》,第5期。

3. 李志生、刘恒甲(2010):《Lee-Carter 死亡率模型的估计与应用——基于中国人口数据的分析》,《中国人口科学》,第3期。
4. 孙福滨等(1993):《中国第四次人口普查全国及部分省区死亡漏报研究》,《中国人口科学》,第2期。
5. 王金营(2013):《1990年以来中国人口寿命水平和死亡模式的再估计》,《人口研究》,第4期。
6. 王广州(2021):《中国人口平均预期寿命预测及其面临的问题研究》,《人口与经济》,第6期。
7. Booth H., Maindonald J., Smith L. (2002), Applying Lee-Carter under Conditions of Variable Mortality Decline. *Population Studies*. 56(3):325-336.
8. Enchev V., Kleinow T., Cairns A. J. G. (2017), Multi-Population Mortality Models: Fitting, Forecasting and Comparisons. *Scandinavian Actuarial Journal*. 2017(4):319-342.
9. Gao Q., Hu C. (2009), Dynamic Mortality Factor Model with Conditional Heteroskedasticity. *Insurance: Mathematics and Economics*. 45(3):410-423.
10. Lee R. D., Carter L. R. (1992), Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*. 87(419):659-671.
11. Lee R., Miller T. (2001), Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality. *Demography*. 38(4):537-549.
12. Li N., Lee R. (2005), Coherent Mortality Forecasts for a Group of Populations: An Extension of the Lee-Carter Method. *Demography*. 42(3):575-594.
13. Li N., Lee R., Gerland P. (2013), Extending the Lee-Carter Method to Model the Rotation of Age Patterns of Mortality Decline for Long-term Projections. *Demography*. 50(6):2037-2051.
14. Li N., Lee R., Tuljapurkar S. (2004), Using the Lee-Carter Method to Forecast Mortality for Populations with Limited Data. *International Statistical Review*. 72(1):19-36.
15. Nigri A., Levantesi S., Marino, M. (2021), Life Expectancy and Lifespan Disparity Forecasting: A Long Short-term Memory Approach. *Scandinavian Actuarial Journal*. (2):110-133.
16. Medford A. (2017), Best-Practice Life Expectancy: An Extreme Value Approach. *Demographic Research*. 36(34):989-1014.
17. Oeppen J., Vaupel J. W. (2002), Broken Limits to Life Expectancy. *Science*. 296(5570):1029-1031.
18. Pascariu M. D., Canudas-Romo V., Vaupel J. W. (2018), The Double-Gap Life Expectancy Forecasting Model. *Insurance: Mathematics and Economics*. 78:339-350.
19. Plat R. (2011), One-year Value-at-risk for Longevity and Mortality. *Insurance: Mathematics and Economics*. 49(3):462-470.
20. Raftery A. E., Chunn J. L., Gerland P., Ševčíková H. (2013), Bayesian Probabilistic Projections of Life Expectancy for All Countries. *Demography*. 50(3):777-801.
21. Torri T., Vaupel J. W. (2012), Forecasting Life Expectancy in an International Context. *International Journal of Forecasting*. 28(2):519-531.
22. Wilmoth J. R., Deegan L. J., Lundstrom H., Horiuchi S. (2000), Increase of Maximum Life-span in Sweden, 1861-1999. *Science*. 289(5488):2366-2368.

(责任编辑:李玉柱)