

基于卡尔曼滤波的人寿保险原保费收入预测分析

冯海琦, 王琳琳, 樊永红

(鲁东大学 数学与统计科学学院, 山东 烟台 264039)

摘要: 近年来,我国保险业发展迅速,尤其是保费收入增速明显。保费收入不仅会对保险公司未来决策产生影响,也会影响政府对保险业的政策导向,因此使用科学的方法预测原保费收入,可以为提供有效的理论支撑。本文利用卡尔曼滤波算法与指数平滑法,对我国人寿保险原保费收入进行预测分析,发现卡尔曼滤波算法的平均误差比指数平滑法减少了约 18.43%,提高了预测的精度。

关键词: 人寿保险; 原保费; 卡尔曼滤波模型; 预测

中图分类号: O213 **文献标志码:** A **文章编号:** 1673-8020(2021)01-0006-05

现今,居民消费水平的不断提高和人们消费观念的不断转变,使人们对人寿保险的需求越来越多。截至 2019 年末,我国 16~59 岁的劳动年龄人口比上年同期减少了 89 万,60 岁以上老年人口比上年增加 439 万人。随着新生儿数量不断减少,老龄人口的大幅增加,政府的社会福利无法满足老龄化人口的需求。因此,人寿保险公司需要加大对养老保险的投入,拓展新型保障业务,为国家分担人口老龄化带来的经济社会发展压力。

从 1982 年国内恢复寿险业务开始,在国家的支持下,中国人寿保险行业不断发展。在保险业刚进入我国时,我国保费收入仅 4.4 亿元,发展至今保费收入增长迅速,1999 年我国保险收入为 1 175.43 亿元,截至 2019 年 11 月保费收入达到了 39 620 亿元,20 年间增长了约 32 倍。在保费不断增长的过程中,保险从业人员数量也在不断增加。截至 2018 年底,国内保险公司共 229 家,职工人数 1 237 751 人。从上述数据可以看出,保险业发展潜力巨大,竞争力在逐步提高^[1]。

关于人寿保险保费收入的研究分析中,梁来存^[2]通过岭回归法与因子分析方法,结合 1985—2006 年的经济数据,对寿险需求的影响因素进行了相关分析研究,并针对主要影响因素提出关于

提升寿险需求的建议。潘军昌等^[3]采用 1991—2017 年的相关数据,运用统计回归模型进行定量分析,得出对我国寿险需求有抑制和促进作用的不同因素,其中具有促进作用的主要是居民可支配收入,抑制作用较明显的则是通货膨胀。任燕燕等^[4]研究了 48 个国家连续几年的经济数据,运用多元回归模型建模分析影响不同国家寿险需求的原因,发现在不同发展水平的国家,影响寿险的因素并不相同,即使是同一因素对不同国家的影响效果也不同。李国东等^[5]将时间序列分析与数理统计分析相结合,构建新的动态模型,该方法相对于传统方法降低了预测误差。张志浩^[6]利用财险保费收入的时间序列数据,通过分析影响财产保费收入的原因,构建不同的模型,并将单一模型与组合模型进行对比分析,说明组合模型对寿险保费收入预测效果较好。

上述对于寿险原保费收入的预测分析研究,都是建立时间序列模型,但对于提升精度的方法研究较少。此外,大多数学者采用的预测方法为季节 ARIMA 模型^[7]、季节调整 SARIMA 模型^[8]和神经网络^[9]等。本文运用卡尔曼滤波算法模型预测人寿保险原保费收入,在指数平滑法分析的基础上,为预测提供新的思路,并提高预测的准

收稿日期: 2020-09-29; 修回日期: 2020-11-03

基金项目: 国家自然科学基金(11201213); 山东省自然科学基金(ZR2015AM026); 山东省高校科技发展计划(J15LI07)

第一作者简介: 冯海琦(1995—),女,黑龙江七台河人,硕士研究生,研究方向为金融统计和风险管理。E-mail: 1653736339@qq.com

通信作者简介: 王琳琳(1976—),女,山东博兴人,教授,硕士研究生导师,博士,研究方向为金融统计和风险管理。E-mail: wangll_1994@sina.com

确性,为人寿保险市场提供科学的监测分析方法。

1 卡尔曼滤波算法

卡尔曼滤波算法是一种最佳的线性测量方法,该算法是应用前一时刻状态变量的估计值来估计当前时刻的观测值。卡尔曼滤波算法的基本原理是:用最小均方误差为估计准则,构造一套递归估计算法来估计状态变量的发展趋势,通过连续的信息反馈、更新、校正,以达到对状态变量的精准估计。以变量的前一时刻的最佳估计量为参照标准,预测当前时刻的变量状态,并同时观察该状态^[10]。通过对观测值与预测量之间的差异进行分析,并结合观测值和预测量对预测值进行校正,从而得到变量当前时刻的最优状态估计^[11]。

卡尔曼滤波可以分为两部分:时间更新方程与状态更新方程。这两类方程的功能各不相同,状态更新方程的功能是最重要的反馈功能,这里反馈是将先前状态和新测量值进行组合。经过比较分析,基于两类状态方程构建改进规则,可以获得最终的最佳估计^[12]。卡尔曼滤波算法的时间更新方程也称为预测方程,测量更新方程称为校正方程,分别表示为^[13]:

$$\begin{cases} x_{klk-1} = Ax_{k-1k-1} + B\mu_{k-1}, \\ P_{klk-1} = AP_{k-1k-1}A^T + Q, \\ K_k = P_{k-1k-1}H^T(HP_{k-1k-1} + R)^{-1}, \\ x_{klk-1} = x_{k-1k-1} + K_k(z_k - Hx_{k-1k-1}), \\ P_{klk} = (I - K_kH)P_{k-1k-1}, \end{cases} \quad (1)$$

$$\begin{cases} x_{klk-1} = Ax_{k-1k-1} + B\mu_{k-1}, \\ P_{klk-1} = AP_{k-1k-1}A^T + Q, \\ K_k = P_{k-1k-1}H^T(HP_{k-1k-1} + R)^{-1}, \\ x_{klk-1} = x_{k-1k-1} + K_k(z_k - Hx_{k-1k-1}), \\ P_{klk} = (I - K_kH)P_{k-1k-1}, \end{cases} \quad (2)$$

式中, x_{klk-1} 、 z_k 分别为 k 时刻系统的状态和观测值, P_{klk} 为 k 时刻的方差, K_k 为 k 时刻的卡尔曼增益; Q 为系统过程噪声方差,即 x_{klk} 到 x_{klk-1} 过程中状态转移和真实过程之间的误差; R 为观测噪声误差,即 x_{klk} 到 z_k 过程中的观测误差^[14]。

卡尔曼滤波模型在分析过程中结合了两种噪声,即系统状态过程中的噪声和测量中的噪声,将时间序列数据观测值直接用作实际值,保留相关参数,对其进行分析。卡尔曼滤波算法表面上看起来很复杂,实际上只有三个参数 P 、 Q 、 R 用于建模分析。这里 P 是初始预测的方差,可以是零以外的任意数字,模型将自动调整为适当的参数; Q 、 R 可以视为常数,根据样本数据的特征选取并进行尝试^[15]。

2 模型实证

保险业在中国起步比较晚,但是发展一直较快,国内外对于寿险的研究一直在进行^[16]。尽管我国大力支持寿险行业发展,但在一线城市的人寿保险市场仍然和发达国家有很大的差距,因此人寿保险行业发展潜力巨大。图1是1999—2018年我国人寿保险保费收入的统计数据(数据来源于中国银保监会网站),可以看到:我国人寿保险保费收入增长在起步阶段发展速度缓慢,但是后续发展速度上升迅猛;从2013年之后发展速度激增,几乎呈指数形式增长。

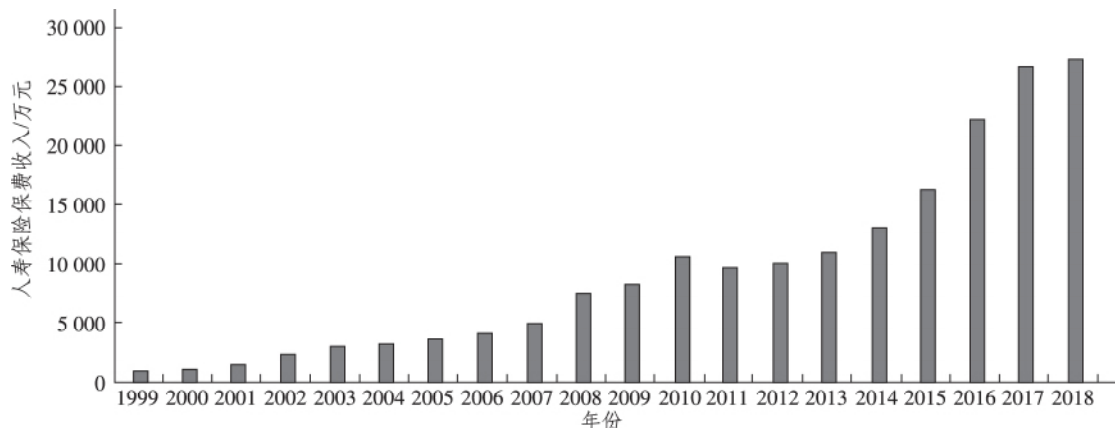


图1 1999—2018年人寿保险原保费收入

Fig.1 Original premium income of life insurance premiums from 1999 to 2018

2.1 模型建立

本文使用 2010 年 1 月—2019 年 12 月人寿险业的月原保费收入数据作为观测值,对该数据进行拟合分析,建立相应的卡尔曼滤波模型。通过以上模型预测 2020 年 1—7 月的原保费收入,与银保监会公布的真实数据进行对比分析,观察模型的准确度。

设定模型中参数 $P_0 = 1$ 经过不断尝试,其它

参数为 $Q = 0.06$, $R = 0.0001$, 拟合效果较好。运用 R 软件计算,得到 2010 年 1 月—2019 年 12 月我国人寿保险原保费收入的拟合值,从结果中可以发现,模型的平均误差较低,仅为 6.4%。为检验上述拟合效果,利用 SPSS 软件对数据构建指数平滑法模型进行对比,得到相关参数(表 1)。表 1 中 $R^2 = 0.988$, 拟合优度非常好,说明模型精度也很好,但是对比卡尔曼滤波模型的平均误差,指数平滑法平均误差高出了 18.43%。

表 1 指数平滑法模型统计量
Tab.1 Exponential smoothing model statistics

| 预测变量数 | 模型拟合统计量 | | Ljung-Box Q(18) | | |
|-------|-----------|-------|-----------------|----|-------|
| | 平稳的 R 方 | R 方 | 统计量 | DF | Sig. |
| 0 | 0.136 | 0.988 | 130.319 | 15 | 0.000 |

2.2 模型结果分析

对于 2020 年各月原保费收入数据的预测,由于卡尔曼滤波无法单独进行向后预测,因此,需要借助指数平滑法模型拟合的预测数据进行卡尔曼滤波分析,在指数平滑模型预测的基础上提供其预测精度。将由上述方法得到的 2020 年数据作为原始值,通过卡尔曼滤波模型来拟合 2020 年 1—7 月的人寿保险原保费收入的数值;为检验模型的预测效果,也将指数平滑模型预测的数据输入灰色预测模型中进行预测分析,由此说明卡尔

曼滤波可以提升指数平滑法预测数据的精度(表 2)。

根据表 2 数据,卡尔曼滤波模型在指数平滑模型的基础上提高了预测精度。在 2020 年 1—2 月,由于新冠病毒的影响,人们对于保险的认知度提高,增加了人寿保险的保费投入,因此,预测误差较大。由此可以看出:随着人们生活逐步恢复正常,长期误差减小;在 2020 年 7 月卡尔曼滤波预测误差仅为 1.8%,后续误差小,准确率较高。因此,与灰色预测模型结果对比,卡尔曼滤波整体预测效果良好,对于提升准确性具有较好的作用。

表 2 卡尔曼滤波预测数据对比

Tab.2 Comparison of Kalman filtering prediction data

| 日期 | 实际值/亿元 | 卡尔曼滤波预测值/亿元 | 误差/% | 灰色预测模型预测值/亿元 | 误差/% |
|---------|-----------|-------------|---------|--------------|---------|
| 2020.01 | 6 686.60 | 5 454.930 | 0.225 8 | 5 454.930 | 0.225 8 |
| 2020.02 | 8 203.79 | 7 389.565 | 0.110 2 | 8 480.125 | 0.032 6 |
| 2020.03 | 10 797.93 | 9 926.333 | 0.087 8 | 9 691.346 | 0.114 2 |
| 2020.04 | 12 182.79 | 11 314.330 | 0.076 8 | 11 075.566 | 0.100 0 |
| 2020.05 | 13 634.61 | 12 879.662 | 0.058 6 | 12 657.495 | 0.077 2 |
| 2020.06 | 15 618.92 | 15 018.655 | 0.040 0 | 14 465.372 | 0.079 7 |
| 2020.07 | 16 881.53 | 16 576.151 | 0.018 4 | 16 531.469 | 0.021 2 |

3 人寿保险的发展建议

通过对人寿保险行业的月度原保费收入数据进行相关分析,观察人寿险业的原保费收入,并对卡尔曼滤波模型的拟合结果进行研究分析发现:疫情期间人寿保险实际原保费收入比预测值高,在一定程度上说明人们在当前事件发生后,对

于人寿保险的认知提高,人寿保险需求增加,人寿保险公司销售量也增加;加大保险在人们生活中的保障,也提高了保险在风险防范中的作用。

“十三五”保险业发展规划^[17]提出我国保险行业的发展计划:增强创新能力,增强风险监管能力,将保险行业做大做强。因此,寿险行业也需要不断提升行业要求。通过发挥监管和市场的调节作用,充分驱动人寿保险市场竞争,尽快建立一个

基本社会功能发挥完善、产品结构协调有序的人寿保险行业。

居民消费水平的提升、健康和保障需求的增加、重大突发疾病、慢性病的频发以及人口老龄化问题的加剧,这些都给人寿保险行业提供了发展新思路。保险的主要功能之一是分散风险,因此,保险公司可以结合相关实际问题,建立新型保障机制。从今年突发的公共卫生事件中发现,人寿保险业也是经济发展的一道防线,在一定程度上为人们的经济提供风险保障。针对以上研究和发现,提出下面几点建议:

首先,寿险公司有必要对寿险市场进行探索与研究,寻找更大的发展空间。各大寿险公司可以结合市场行情和公司的实际发展状况,综合原保费收入数据,建立相应的预测模型;应用相关数据模型监管其他业务,将日常管理与数据分析相结合,探索多种发展模式。

其次,提升风险应对机制。自 2020 年 1 月以来,新冠病毒席卷全球,对全球的经济均产生影响,对人寿保险的原保费收入也产生了一定程度的影响。利用本文的分析结果,结合公司的实际原保费收入进行有效对比研究,并对这次疫情产生的效果和影响进行相关政策分析,为公司经营提供相关的数据支持。同时,为能够在突发事件发生时作出快速应对,并减少风险损失,提供了参考数据。

然后,加强和完善监管体系。通过对比各保险公司发布的年报数据,及时发现异常数据,进而加强监管措施,使人寿保险市场更加公平合规,更加有效地进行风险管控。

最后,提高大众对保险的认知。面对我国人口老龄化问题,人寿保险公司应该在配合养老保险发展的同时,研究创新性保险产品,为社会养老保险分担压力。在人口老龄化不断发展的今天,通过引导大众认识保险的重要性,提高其对保险的认知,扩大保险普及度,开发并制定适合个人的寿险保单,在增加保险覆盖率的同时提高保障水平。

4 结语

本文通过指数平滑法预测拟合,将得到的数据再利用卡尔曼滤波算法进行分析。通过与灰色

预测模型预测的数据进行对比,得出结论:利用卡尔曼滤波算法可以提高指数平滑法拟合数据的精确度。这不仅为保险公司未来预测保费时提供参考模型,也有助于政府进行有效监管。

参考文献:

- [1] 中华人民共和国统计局.中国统计年鉴[M].北京:中国统计出版社,2018.
- [2] 梁来存.我国寿险需求的实证分析[J].数量经济技术经济研究,2007,24(8):80-89.
- [3] 潘军昌,杨军.我国寿险需求的影响因素分析[J].经济管理,2010,32(2):131-137.
- [4] 任燕燕,刘金娥.寿险需求影响因素的实证分析[J].统计与决策,2008(22):97-100.
- [5] 李国东,臧鸿雁,陈文义,等.人寿保险收入的数学模型与预测[J].哈尔滨理工大学学报,2003(3):122-124.
- [6] 张志浩.保险行业保费收入预测及发展趋势实证分析[D].哈尔滨:黑龙江大学,2017.
- [7] 陈黎明,赵元元.季节 ARIMA 模型在保费总收入预测中的应用[J].福建金融管理干部学院学报,2018(4):3-10.
- [8] 刁莉,王宁.基于 X12-LSTM 模型的保费收入预测研究[J].计算机科学,2020,47(增刊1):512-516.
- [9] 周桦,卢志源,郑敏.基于 TEI@I 方法的中国保险业保费收入预测[J].管理评论,2020,32(7):166-179.
- [10] 彭丁聪.卡尔曼滤波的基本原理及应用[J].软件导刊,2009,8(11):32-34.
- [11] 王艳,王珍.基于卡尔曼滤波的浙江省住户存款预测[J].金融理论与教学,2020(4):7-11.
- [12] 舒服华,马勇军.卡尔曼滤波算法在我国钢产量预测中的运用[J].徐州工程学院学报(自然科学版),2018,33(2):8-12.
- [13] 王艳.基于卡尔曼滤波算法的新疆棉花产量预测[J].国际纺织导报,2018,46(6):57-60.
- [14] 朱中,杨兆升.基于卡尔曼滤波理论的实时行程时间预测模型[J].系统工程理论与实践,1999(9):3-5.
- [15] 舒服华.基于卡尔曼滤波的四川省社会消费品零售额预测[J].阿坝师范学院学报,2019,36(4):43-48.
- [16] BROWNE M J, KIM K. An international analysis of life insurance demand[J]. The Journal of Risk and Insurance, 1993, 60(4): 616-634.
- [17] 王和.《中国保险业发展“十三五”规划纲要》解读[J].保险理论与实践,2016(9):44-50.

Prediction Analysis of Original Premium Income of Life Insurance Based on Kalman Filter

FENG Haiqi , WANG Linlin , FAN Yonghong

(School of Mathematics and Statistics Science ,Ludong University ,Yantai 264039 ,China)

Abstract: In recent years ,insurance industry has developed rapidly ,especially the growth rate of insurance premiums.The collection of premiums will not only affect the future decision-making of insurance companies , but also affect the government 's policy orientation towards the insurance industry.Therefore a scientific approach of predicting the original premium input can provide an effective theoretical support for the insurance industry. This paper predicted and analyzed the original premium income of life insurance by using the Kalman filtering algorithm and the forecast accuracy was improved based on the exponential smoothing method.Compared with the exponential smoothing method ,the average error of the model is reduced by 18.43% ,which provides more possibilities for future prediction.

Keywords: life insurance; original premium; Kalman filtering model; prediction

(责任编辑 顾建忠)

版权声明

根据《中华人民共和国著作权法》《信息网络传播权保护条例》等法律法规的规定,本刊作如下声明:

1. 作者向本刊投稿,即意味着将作品的发表权、删改权、信息网络传播权、数字化汇编权、数字化复制权、数字化制品形式(包括光盘、互联网出版物)出版发行权等权利授予本刊,并视同许可本刊官方新媒体免费转载以及与有关数据库的合作(本刊不再另行支付费用)。如不同意以上授权,请在投稿时说明。

2. 本刊载刊的全部编辑内容归《鲁东大学学报(自然科学版)》编辑部所有,非经书面同意,任何单位和个人不得转载、摘编、刊印或以其他方式使用在本刊发表的文章等。如有违反,本刊保留一切法律追究的权利。

3. 本刊版面、栏目等受著作权保护,对复制、仿制、假冒者将追究法律责任。

4. 已在本刊发表的作品,本刊有免费结集出版精华本、合订本,以及相关电子产品的权利,有特别声明者除外。

《鲁东大学学报(自然科学版)》编辑部