

Selection of a Stochastic Mortality Model for the Male and Female Population in Iran

Amir Teymor Payandeh Najafabadi^{1*}, Aida Aalabeyki²

Abstract

Forecasting the death rate accurately is of paramount importance. The purpose of this research is to select the best stochastic mortality model in Iran based on a comparison between Lee-Carter models and the age-at-death distribution model (maximum entropy) for the male and female populations separately. It is worth mentioning that one of the strengths of modeling using age-at-death distribution is that it effectively represents the age-specific mortality pattern in Iran and answers the question of whether the mortality pattern in Iran is stable. Considering any model that reports the lowest error based on evaluation criteria, the life table for the year 1400 in Iran has been introduced and compared with the mortality tables TD88-90 in France and ILT1400 in Iran. In this regard, real mortality data in Iran for the years 2011-2019 for ages 0-100 years have been used. Ultimately, it was concluded that the Lee-Carter model performed better for the male population and the Currie model for the female population in Iran. It was also concluded that the existing tables in Iran report a higher mortality rate, and the life table presented in this article may be suitable for pricing insurance products and for more accurate calculation of actuarial indicators, such as the probability of death.

Keywords: Forecasting death rate, Age-at-death distribution, Maximum Entropy, Classical stochastic models, Lee-Carter Model.

Received: 2024-05-05

Accepted: 2024-07-28

¹ . Professor, Department of Actuarial Science, Faculty of Mathematical Science, Shahid Beheshti University, Tehran, Iran (Corresponding Author); amirtpayandeh@sbu.ac.ir

² . PhD Candidate in Actuarial Science, Faculty of Mathematical Science, Shahid Beheshti University, Tehran, Iran; aidaalabeygi@gmail.com

Extended Abstract

Introduction

The death rate holds particular significance in the field of actuarial science. While the global population is growing, stochastic models are increasingly demanded to accurately forecast death rate and life expectancy in the future to support public policies, insurance companies and health institutions to face their risks. Since no dynamic mortality model has been presented in Iran to forecast the death rate by gender, this research aims to select the best stochastic model in Iran based on a comparison between the Lee-Carter family of models and the age-at-death distribution model (maximum entropy) for the male and female populations separately. The three variables of age, period (time), and cohort are used to classify the base model into zero-, one-, two-, or three-factor categories. The zero-factor model is an age-specific rate. One-factor models consider the mortality rate as a function of age. Most forecasting mortality methods use two-factor models of age and period, such as Lee-Carter model. The three-factor models consider mortality as a function of age, period, and cohort. One of the most important criteria in selecting the base model is the interpretability of the parameters. Better interpretation of model parameters will help to more accurately evaluate the parameter trend and its continuation in the future. Mortality modelling boasts a long history in demographic and actuarial science; however, the practice of mortality forecasting is relatively recent. Over the past thirty years, more complex statistical methodologies to forecast mortality have been recommended and used. Most of the models used in modelling and forecasting are based on age-specific death rates, as they are well suited for showing changes in the risk of death over age and time, like the Lee Carter family of models. This family of models utilizes three parameters—age, period, and cohort—to model mortality. Although they provide good information, they cannot answer two questions: (1) How many years does a person live in a population on average? and (2) how to measure the variability of ages at death?

Another effective perspective on the evolution of mortality and changes in the pattern of human lifespan is the use of age-at-death distribution, which captures observed mortality trends in a population and addresses key questions in mortality studies. Age-at-death distribution provides an answer for the above questions, but they have mainly been neglected in the forecasting of future death rate. In addition, it provides unique information and offers valuable insights into mortality patterns and

longevity risk; however, it is seldom used in modelling and forecasting populations. One of the new methods based on the age-at-death distribution is using statistical moments and the density function to forecast age-specific mortality levels. Time series methods are used to extrapolate a limited number of statistical moments. In a more recent development, Marius D. Pascariu et al. (2019) proposed a novel modelling technique using statistical moments. The central idea in this paper involves leveraging statistical knowledge, particularly statistical moments, and using time series methods to extrapolate a limited number of moments.

Methods and Data

One commonly used method for modelling and forecasting the death rate is the family of Lee-Carter models. In this method, the logarithm of the death rate or the logarithm of the probability of death is utilized for modelling. Another effective perspective on the evolution of mortality and changes in the pattern of human lifespan is the use of age-at-death distribution, which captures observed mortality trends in a population and addresses key questions in mortality studies. In this research, ten models based on death rate, along with the maximum entropy model with 4, 5, and 6 moments derived from age-at-death distributions, were employed for modelling and forecasting the death rate. Real data from Iran's mortality, specifically the number of registered deaths and the country's population for the years 2011-2019 for ages 0-100 years, were used. In this research, ten different stochastic mortality models were evaluated with seven accuracy measures.

Findings

This study focuses on selecting the best stochastic mortality model to forecast the death rate in Iran, differentiated by gender and based on actual data. Specifically, it utilizes nine stochastic models for mortality rate forecasting and a maximum entropy model for mortality frequency forecasting. To evaluate the predictive power of each model, seven assessment criteria were employed. Within the specified range (0-100 years), it was determined, among the ten models examined, the Lee-Carter stochastic model performed best for the male population, while the APC stochastic model was the best for the female population. This indicates that the two-parameter and three-parameter models based on the Lee-Carter family provided more accurate results. With slight differences following these models, the Hindman model and the Pleet

model ranked second for the male and female populations in Iran, respectively. Furthermore, life tables for Iran for the year 1400 (2021) were presented, taking into account the predictive performance of the APC and Lee-Carter models for the female and male populations of Iran, respectively. As shown in Tables 4 and 5, the average life expectancy for women is three years higher in childhood, two years higher in middle age, and one year higher in old age compared to men. Therefore, the overall mortality pattern for women differs from that of men, indicating that a single model may not perform equally well for both genders. Additionally, a comparison of the predicted mortality rates for the years 1400 and 1401 (2021 and 2022) with the mortality rates in the existing life tables of Iran was conducted. The comparison results showed that, in most age ranges, the predicted mortality rates (probability of death) from the best-performing model reported lower values than those in the existing life tables in Iran. This finding is particularly significant for the insurance industry and pension funds.

Conclusion and discussion

The lower the probability of death for an insured individual, the longer they are expected to live, allowing us to charge lower premiums for mixed life insurance policies and life insurance with the condition of death. This reduction in premiums can be a factor in increasing demand for such insurance products. Additionally, when issuing life insurance policies with the condition of survival, the decreased probability of death and extended duration of coverage necessitates charging higher premiums to the insured individual. In this study, to compare the predictive accuracy of stochastic mortality models, the crude death rate index was used. It is clear that changes in the mortality index and the evaluation criteria of the models can yield different results.

Acknowledgment

This article is based on the master's thesis of the second author in the field of actuarial science, conducted at the Faculty of Mathematical Sciences at Shahid Beheshti University. The authors extend their gratitude for the financial support provided by the Information and Communications Technology Park of the Ministry of Science, Research, and Technology, under grant code 000553-01-00-02.

Citation:

Payandeh Najafabadi, A.T., Aalabeyki, A. (2024), Selection of a Stochastic Mortality Model for the Male and Female Population in Iran, *Journal of Population Association of Iran*, 19(38), 7-56 <https://doi.org/10.22034/jpai.2024.2027718.1347>

ارجاع:

پاینده نجف‌آبادی، امیرتیمور، اعلائیکی، آیدا (۱۴۰۳). انتخاب مدل مرگ‌ومیر تصادفی برای جمعیت زنان و مردان ایرانی، نامه انجمن جمعیت‌شناسی ایران، ۱۹(۳۸)، ۷-۵۶. <https://doi.org/10.22034/jpai.2024.2027718.1347>

انتخاب مدل مرگ‌ومیر تصادفی برای جمعیت زنان و مردان ایرانی

امیر تیمور پاینده نجف‌آبادی^{۱*}، آیدا اعلاییکی^۲

چکیده

پیش‌بینی هرچه دقیق‌تر نرخ فوت از اهمیت قابل‌توجهی برخوردار است. هدف از این پژوهش انتخاب بهترین مدل مرگ‌ومیر تصادفی کشور ایران با توجه به نتایج حاصل از مقایسه بین خانواده مدل‌های لی‌کارتر و خانواده مدل‌های توزیع سن در زمان فوت (ماکسیمم آنتروپی) برای زنان و مردان است. شایان ذکر است که یکی از نقاط قوت مدل‌سازی به کمک توزیع سن در زمان فوت، این است که به خوبی روند الگوی مرگ‌ومیر خاص سنی در ایران را نشان می‌دهد و پاسخگوی این سوال است که آیا الگوی مرگ در ایران ثابت است؟ با توجه به هرمدلی که براساس معیارهای ارزیابی کمترین خطا را گزارش کند جدول عمر سال ۱۴۰۰ ایران معرفی و بنا عملکرد جداول عمر TD88-90 فرانسه و ILT1400 ایران مقایسه می‌شود. در این راستا از داده‌های واقعی مرگ‌ومیر ایران مربوط به سال‌ها ۱۳۹۰-۱۳۹۸ برای سنین ۰-۱۰۰ سال استفاده شده است. در نهایت نتیجه‌گیری شد مدل لی‌کارتر برای جمعیت مردان و مدل کوری برای جمعیت زنان ایران عملکرد بهتری داشته است. همچنین نتیجه‌گیری شد جداول موجود در ایران، نرخ مرگ‌ومیر بیشتری را گزارش می‌دهند و استفاده از جدول عمر ارائه شده در این مقاله می‌تواند برای قیمت‌گذاری محصولات بیمه‌ای و همچنین محاسبه دقیق‌تر شاخص‌های بیم‌سنجی (احتمال فوت) مناسب باشد.

واژگان کلیدی: پیش‌بینی نرخ فوت، توزیع سن در زمان فوت، ماکسیمم آنتروپی، مدل‌های تصادفی کلاسیک، مدل لی‌کارتر

تاریخ پذیرش: ۱۴۰۳/۰۵/۰۷

تاریخ دریافت: ۱۴۰۳/۰۲/۱۶

*۱. استاد آمار، گروه بیم‌سنجی، دانشکده علوم ریاضی، دانشگاه شهید بهشتی، تهران، ایران (نویسنده مسئول)؛

amirtpayandeh@sbu.ac.ir

۲. دانشجوی دکترای بیم‌سنجی، دانشکده علوم ریاضی، دانشگاه شهید بهشتی، تهران، ایران؛

aidaalabeygi@gmail.com

DOI: <https://doi.org/10.22034/jpai.2024.2027718.1347>

مقدمه و بیان مسأله

مدل‌بندی و پیش‌بینی مرگ‌ومیر در دهه‌های اخیر به یکی از موضوعات مهم در حوزه جمعیت‌شناسی و بیم‌سنجی تبدیل شده است. شرکت‌های بیمه و صندوق‌های بازنشستگی از پیش‌بینی نرخ فوت به‌عنوان ورودی در محاسبات قیمت‌گذاری محصولات عمر، پرداخت‌های بازنشستگی و همچنین تعیین حق‌بیمه بیمه‌گذاران خود استفاده می‌کنند. همچنین این پیش‌بینی‌ها برای حقوق بازنشستگی نیز بسیار حائز اهمیت است، زیرا در هنگام بازنشستگی پرداخت‌های ماهانه را براساس امید زندگی افراد محاسبه می‌کنند (Levantesi & Pizzorusso, 2019). عامل اصلی افزایش امید زندگی را می‌توان کاهش میزان مرگ‌ومیر در سنین سالخوردگی نام برد و اگرچه یکی از بهترین دستاوردهای بشر است اما به یکی از مهم‌ترین چالش‌های سازمان بهداشت جهانی، سازمان‌های تامین اجتماعی و شرکت‌های بیمه که محصولات بیمه عمر و مستمری ارائه می‌کنند تبدیل شده است. استفاده از جداول عمر بومی ساخته شده با توجه به میزان مرگ‌ومیر کشور خودمان به تفکیک جنسیت، بر محاسبه دقیق‌تر حق بیمه محصولات نیز اثر چشم‌گیری دارد. لذا در این پژوهش تلاش شده است تا با ارائه جدول عمری مبتنی بر عملکرد بهترین مدل تصادفی انتخاب شده به محاسبات دقیق‌تر نرخ فوت پرداخته به‌گونه‌ای که تقاضا برای محصولات بیمه عمر در ایران بیشتر شود. همچنین از آنجا که نرخ فوت برای هر جامعه و زیرمجموعه جمعیتی هر کشور متفاوت است، نیاز است تا از نرخ‌های فوت متفاوتی برای جمعیت زنان و مردان استفاده شود. زیرا به‌صورت کلی زنان طول عمر طولانی‌تری نسبت به مردان دارند. نرخ فوت در ایران در ابتدا براساس جدول مرگ‌ومیر TD88-90 فرانسه محاسبه می‌شد، که مبتنی بر نرخ فوت مردان فرانسه است. از سال ۱۴۰۰ تاکنون از نرخ فوت جدول ILT1400 ایران در محاسبات استفاده می‌شود اما متأسفانه در این جدول نیز تفکیک جنسیت انجام نشده است. لذا در این مقاله، علاوه بر انتخاب بهترین مدل مرگ‌ومیر تصادفی در پیش‌بینی نرخ فوت کشور ایران به تفکیک جنسیت، به ساخت جدول عمر جداگانه برای زنان و مردان بر مبنای عملکرد بهترین مدل انتخابی پرداخته می‌شود.

جدول عمر یکی از مهم‌ترین ابزارهای جمعیت‌شناسی است که بیانگر وضعیت کنونی جمعیت و ادامه حیات، جامعه مورد بررسی است. برای ساخت جدول عمر می‌توان به کمک میزان مرگ‌ومیر کل عناصر جدول را بدست آورد، اما به کمک آن نمی‌توان دست به پیش‌بینی جمعیتی زد. لذا مدل‌سازی مرگ‌ومیر از نیمه اول قرن هجدهم به یک موضوع ثابت تبدیل شده است و تحت سه رویکرد کلی بررسی می‌شود (Booth & Tickle, 2008). در حالت کلی برای مشخص کردن نرخ فوت در ادبیات بیم‌سنجی از دو روش استفاده می‌شود: (۱) روش مدل‌بندی (۲) روش جدول‌سازی (پاینده نجف‌آبادی، ۱۴۰۱). همچنین به‌طور کلی سه عامل سن، دوره و گروه در مدل‌بندی و پیش‌بینی مرگ‌ومیر به کار می‌روند.

یکی دیگر از دیدگاه‌های کارآمد و متفاوت برای مدل‌سازی و پیش‌بینی مرگ‌ومیر استفاده از توزیع سن در زمان فوت است. این توزیع بیانگر مرگ‌ومیر تجربی یک جمعیت است و اطلاعات مفیدی در مورد معیار کلیدی طول عمر مانند متوسط مدت زمان زندگی یک جمعیت، نابرابری تجربه شده توسط جمعیت در زمان فوت و یا طول عمر فراتر از حد انتظار افراد بدست می‌آورد. علی‌رغم مزایای ذکر شده برای این روش، عموماً در مدل‌سازی و پیش‌بینی مرگ‌ومیر نادیده گرفته می‌شود. با توجه به اهمیت مدل‌بندی و پیش‌بینی مرگ‌ومیر، در این مقاله از داده‌های مرگ‌ومیر کشور ایران برای مدل‌سازی و پیش‌بینی نرخ فوت استفاده شده است. داده‌های واقعی شامل تعداد فوتی‌ها و جمعیت کشور ایران، برای ۹ سال متوالی (۱۳۹۰-۱۳۹۸) و برای بازه سنی ۰ تا ۱۰۰ سال و به تفکیک هر دو جنسیت است که به ترتیب از سازمان ثبت احوال کشور و مرکز آمار ایران اخذ شده است. برای تعیین میزای درستی داده‌ها در زمینه مرگ‌ومیر، در ابتدا داده‌ها ارزیابی شده و سپس به مدل‌سازی پرداخته می‌شود. برای اینکار داده‌های ثبتی سالانه فوت از نظر کم‌ثبتی ارزیابی و اصلاح شده‌اند. بدین صورت که ماتریس توزیع سنی و سالانه فوت‌های هر جنسیت به‌صورت جداگانه بررسی شده است. یکی از روش‌های اصلاح آن هموارسازی میانگین محرک است. هموارسازی در واقع بعد از محاسبه نرخ فوت و در هنگام ساخت جدول عمر اعمال شده است. در این روش ابتدا میزان فوت مرکزی

(تقسیم تعداد فوت در هر سن بر جمعیت همان سن) در یک سال مشخص با میزان فوت مرکزی فرضی که به صورت کاهش سالیانه ۵ در صدهزارم از میزان فوت مرکزی سال پایه، سال ۱۳۹۰، محاسبه شده و مقایسه می شود. اگر میزان فوت محاسبه شده از میزان فوت فرضی کمتر باشد مقدار فرضی جایگزین می شود و تعداد فوت جدید به دست می آید. با کمک این روش می توان کم‌ثبتی‌های موجود در یک سال مشخص را شناسایی و اصلاح کرد. اما به دلیل اینکه سال ۱۳۹۰ به عنوان پایه انتخاب شد اصلاحات انجام شده به روش هموارسازی داده‌های این سال را تغییر نمی دهد. همچنین می توان گفت کم‌ثبتی اغلب برای گروه‌های سنی میانسال رخ می دهد و همچنین تاثیر آن در اصلاح کم‌ثبتی ناچیز است. لذا اصلاح کم‌ثبتی در این مقاله به روش هیل انجام شده است. در واقع بعد از اعمال اصلاحات در الگوی فوت با کمک روش هموارسازی، داده‌های حاصل در ارزیابی داده‌های فوت به روش هیل در فاصله بین ۱۳۹۰-۱۳۹۸ مورد استفاده قرار می گیرد. در این روش علاوه بر داده‌های فوت داده‌های دو سرشماری ۱۳۹۰ و ۱۳۹۵ نیز وجود دارد. داده‌های سرشماری به منظور استفاده در این روش به میانه سال سرشماری جابه‌جا شده است. بنابراین داده‌های فوت بین دو سرشماری باید از اول مهر ۱۳۹۰ تا ۳۱ شهریور ۱۳۹۵ را دربرگیرد. بدین ترتیب داده‌های سالانه ثبت فوتی برای سال‌های ۱۳۹۱ تا ۱۳۹۴ به طور کامل و با فرض توزیع یکنواخت فوت‌ها برای سال‌های ۱۳۹۰ و ۱۳۹۵ است. با کاربرد این داده‌ها روش هیل ضریب پوشش داده‌های فوت را برای مردان و زنان به ترتیب ۸۱ و ۶۲ درصد برآورد کرده است. کاربرد روش هیل برای اصلاح داده‌های فوتی ثبتی در فاصله دو سرشماری ۱۳۹۰ و ۱۳۹۵ بدین معناست که تنها ۸۱٪ از فوت‌های دوره ۶ ساله برای مردان و ۶۲٪ برای زنان ثبت شده است. به منظور اصلاح کم‌ثبتی تعداد فوتی‌های ثبتی دوره ۶ ساله باید بر میزان‌های مذکور تقسیم شود. همچنین برای محاسبه میزان مرگ و میر باید جمعیت سرشماری سال ۱۳۹۰ بر عدد ۰.۸۱ تقسیم شود تا کم‌شماری آن نسبت به سال ۱۳۹۵ اصلاح شود. شایان ذکر است که ضرایب پوشش فوت در روش هیل برای کل دوره محاسبه می شوند و ضریب اعمال شده برای تمام سال‌ها و سنین دوره مطالعه یکسان است اما استفاده از ضریب یکسان

برای سال‌های مختلف چندان درست نمی‌باشد. بنابراین برای بهبود ثبت داده‌های فوت در طی دوره مورد بررسی از شاخص نسبت داده‌های مرگ‌ومیر به کل تعداد مرگ‌ومیر در هر سال مورد بررسی به‌عنوان شاخص وزنی مناسب استفاده شده است. در واقع بعد از اطمینان از صحت داده‌ها، از نرخ فوت خام ایران در سال‌های ۱۳۹۰ تا ۱۳۹۵ برای برازش مدل‌ها استفاده می‌شود و نرخ فوت برای سال‌های ۱۳۹۶-۱۳۹۸ پیش‌بینی می‌شود. سپس با مقایسه نرخ مرگ‌ومیر مشاهده شده با نرخ مرگ‌ومیر پیش‌بینی شده براساس هریک از مدل‌های معرفی شده در ادامه به رتبه‌بندی مدل‌ها براساس دقت پیش‌بینی هریک پرداخته می‌شود. براساس مدلی که رتبه نخست را اخذ کرده است پیش‌بینی نرخ فوت تا افق ۱۴۰۲ انجام می‌شود. و در نهایت نرخ مرگ‌ومیر پیش‌بینی شده براساس عملکرد بهترین مدل تصادفی با نرخ مرگ‌ومیر جداول TD88-90 فرانسه و ILT1400 ایران مقایسه می‌شود. با داشتن نرخ فوت برای کل دوره (۱۴۰۲-۱۳۹۰) می‌توان جداول عمر متناظر با جنسیت را برای هر سال ساخت. ساخت جدول عمر با استفاده از نرم‌افزار R و بسته "demography" ساخته شده است.

مرور پیشینه پژوهش

می‌توان گفت منشأ مدل‌سازی مرگ‌ومیر به‌عنوان یک زمینه تحقیقاتی و جمعیت‌شناسی به‌عنوان یک رشته علمی، به‌طور کلی با کار جان گرانت تاجر مشهور لندنی آغاز شد. (Graunt, 1662) با توجه به مشاهدات طبیعی خود که براساس ثبت دویست هزار مرگ‌ومیر در بولتن‌های کلیسای لندن بود فهرستی شبیه جدول مرگ‌ومیر را که براساس علل مرگ‌ومیر طبقه‌بندی شده بود تدوین کرد که سرآغاز تدوین جدول‌های تحلیلی مرگ‌ومیر شد. بعد از آن (Halley, 1693) ریاضی‌دان و ستاره‌شناس مشهور، جدول کامل‌تری را به تفکیک سن بنا نمود. وی حتی موفق شد روشی را برای انجام محاسبات سالیانه عمر براساس این جدول ارائه دهد. سپس محققان دیگری به این مسئله پرداختند و جدول‌هایی در مقیاس ملی تدوین کردند و تفاوت مرگ‌ومیر در نقاط شهری و روستایی را مطرح ساختند؛ تا آنکه نخستین جدول عمر بر حسب

گروه‌های سنی ساخته شد و از آن به بعد مطالعه مرگ‌ومیر به یکی از موضوعات عمده جمعیت‌شناختی تبدیل شد. شایان ذکر است که جدول عمر مهم‌ترین ابزار تحلیلی جمعیت‌شناسان است اما بدون دسترسی به داده‌های مرگ با جزئیات سنی و جنسی کار دشواری است. بنابراین، در مناطق و کشورهایی که داده‌های ثبت مرگ با جزئیات مورد نیاز در دسترس نیست (Moultrie and et.al, 2013) و یا به دلیل ضعف نظام ثبت، دچار کم‌ثبتی یا خطای گزارش باشند ساخت جدول عمر به مراتب سخت‌تر است. در بسیاری از کشورهای کمتر توسعه‌یافته و در حال توسعه از جمله ایران، با وجود همه پیشرفت‌های صورت گرفته در مقایسه با گذشته، پوشش ثبت واقعه مرگ هنوز صددرصد نیست (عزیزی‌شاکر، ۱۳۹۸). در کشور ایران نیز، به‌منظور در نظر گرفتن تغییرات جمعیتی، طبق آیین نامه شماره ۶۸ شورای عالی بیمه، نهاد ناظر صنعت بیمه، بیمه مرکزی جمهوری اسلامی ایران موظف است حداقل هر پنج سال یک بار جدول عمر اعلام شده را به‌روزرسانی نماید (مصوبات شورای عالی بیمه، ۱۳۹۰).

در مورد پیشینه مدل‌سازی نظری مرگ‌ومیر می‌توان گفت که در ابتدا با (De Moivre, 1725) آغاز شد. وی فرض کرد که مدل فوتی‌ها از توزیع یکنواخت پیروی می‌کنند. بعدها گومپرتز در سال ۱۸۲۵ با اتخاذ یک رویکرد بیولوژیکی به مدل‌سازی ریاضی پرداخت. او فرض کرد که شدت مرگ‌ومیر x در بزرگ سالی افزایش تقریباً تصاعدی را نشان می‌دهد، زیرا بر این باور بود که قوای جسمانی افراد با افزایش سن کاهش می‌یابد. در واقع، این مدل یک تابع نمایی است که در آن دو پارامتر مدل مثبت هستند. (Makeham, 1867) مدل گومپرتز را که برای سنین بالا مناسب نبود اصلاح کرد، همچنین در این مدل یک پارامتر ثابت برای در نظر گرفتن مرگ‌ومیر به دلایل غیرمرتبط با سن اضافه گردیده است. تا اوایل قرن بیستم، مدل‌های گومپرتز و مکهام از جمله مدل‌های رایج مرگ‌ومیر باقی ماندند تا اینکه اقتصاددان ایتالیایی، پارتو نظریات خود را که بر مبنای مشکلات جامعه بود در مدل وارد کرد. برخلاف مدل‌سازی مرگ‌ومیر، پیش‌بینی مرگ‌ومیر یک تلاش نسبتاً جدیدتر است. به‌طور کلی مدل‌سازی مرگ‌ومیر به سه رویکرد تقسیم می‌شود: ۱- رویکرد مبتنی بر تجربه: در رویکرد ذهنی پیش‌بینی مرگ‌ومیر مبتنی

بر تجربیات افراد خبره و متخصصان است. مزیت این رویکرد وارد کردن دانش تخصصی در یک روش کیفی است اما این روش به دلیل ذهنی و محافظه‌کارانه بودن نظرات خبرگان، اغلب منجر به کم برآورد کردن مرگ‌ومیر می‌شود (Waldron, 2005). ۲- رویکرد مبتنی بر برون‌یابی: اکثر روش‌های پیش‌بینی مرگ‌ومیر مبتنی بر رویکرد برون‌یابی هستند. در این روش از نظم موجود در الگوهای سنی در طول زمان برای پیش‌بینی مرگ‌ومیر آتی استفاده می‌شود. این رویکرد از برون‌یابی خطی ساده تا روش‌های مبتنی بر مدل‌های دو عاملی (سن و دوره) را شامل می‌شود. در مدل‌های دو عاملی نرخ فوت، تنها تابعی از سن بیمه‌گذار نیست و دوره زمانی نیز نقش اساسی در تغییرات نرخ مرگ‌ومیر ایفا می‌کند. ۳- رویکرد مبتنی بر علت: در این روش از مدل‌های مبتنی بر علت فوت استفاده می‌شود، که می‌توان دانش پزشکی و اطلاعات در مورد تغییرات رفتاری و محیطی را وارد مدل کرد. نمونه‌ای از چنین مدلی رابطه توضیحی بین سیگار کشیدن و سرطان ریه است. اگرچه استفاده از این روش به علت عدم‌درک کامل از روابط بین عوامل ریسک‌ومرگ‌ومیر با نااطمینانی همراه است، اما کاربرد این روش در شبیه‌سازی اثر شیوع بیماری بر مرگ‌ومیر است. در واقع می‌توان گفت که در طول سه دهه گذشته است که مجموعه جدیدی از مدل‌های مرگ‌ومیر، با تکیه بر روش‌های آماری و تصادفی مشخص می‌شوند و با هدف صریح پیش‌بینی دقیق طول عمر مورد استفاده قرار گرفته‌اند. لازمه پیش‌بینی مرگ‌ومیر، تصریح یک مدل پایه از داده‌ها و تصریح مدلی برای پیش‌بینی است.

از طرفی یکی دیگر از دیدگاه‌های متفاوت در مورد تحولات مرگ‌ومیر و تغییرات الگوی طول عمر انسان، استفاده از توزیع سن در زمان فوت است که نمایانگر مرگ‌ومیر تجربی در یک جمعیت است و به این ترتیب پاسخ‌گوی دو سوال کلیدی در مطالعات مرگ‌ومیر خواهد بود. (۱) انسان‌ها به‌طور متوسط چند سال عمر می‌کنند؟ (۲) سن در زمان فوت چقدر متغیر است؟ پاسخ به سوال دوم می‌تواند زمینه را برای پرسش مطرح شده در مقاله کمیجانی و همکاران (۱۳۹۳) فراهم کند. اینکه آیا نرخ مرگ‌ومیر ویژه سنی در طول زمان در ایران تغییر نکرده است؟

علی‌رغم مزایای ذکر شده برای به تصویر کشیدن الگوهای مرگ‌ومیر و مطالعه نابرابری طول عمر، توزیع سن در زمان فوت عموماً در مدل‌سازی و پیش‌بینی نادیده گرفته می‌شود. تنها چند تلاش برای استفاده صریح از آن‌ها برای پیش‌بینی مرگ‌ومیر آینده انجام شده است. اولین بار (Pearson, 1987)، اقدام به توصیف الگوی مرگ‌ومیر انسانی به کمک توزیع سن در زمان فوت کرد. (Dellaportas, Smith and Stavropoulos, 2001) از تعداد فوتی‌ها برای تطبیق مدل هلیگمن پولارد با روش‌های بیزی استفاده کرده است. همچنین (Mazucco, Scarpa and Zanutto, 2018) مرگ‌ومیر را با برازش یک توزیع نیمه نرمال^۱ و توزیع چوله نرمال چند بعدی^۲ به تابع توزیع تجربی سن در زمان فوت مشاهده شده مدل می‌کنند. ایده جدید دیگری توسط (Basellini and Camarda, 2019) معرفی شد که در آن از یک تابع سه پارامتری برای مدل‌سازی و مقایسه پویایی توزیع مرگ‌ومیر بزرگسالان حول سن معین استفاده می‌کنند که این پارامترها می‌توانند با استفاده از مدل‌های سری زمانی استاندارد برون‌یابی شوند. یک روش نوآورانه دیگر برای مدل‌سازی و پیش‌بینی مرگ‌ومیر استفاده از گشتاورهای آماری است که توسط (Pascariu, Lenart and Canudas-Romo, 2019) ارائه شد. روش‌های سری زمانی برای برون‌یابی تعداد محدودی از گشتاورها استفاده می‌شود و سپس بازسازی توزیع سن در زمان فوت برای آینده انجام می‌شود. استفاده از دانش آماری در مورد شکل توزیع و اینکه چگونه سطح مرگ‌ومیر در یک سن خاص کاملاً به سطوح مرگ‌ومیر در تمام سنین دیگر وابسته است، مزیت بزرگی برای روش‌های برون‌یابی مبتنی بر فراوانی مرگ است. این مقاله نشان می‌دهد که با استفاده از گشتاورهای آماری و اطلاعات توزیع سن در زمان فوت می‌توان پیش‌بینی‌های دقیقی از میزان مرگ‌ومیر خاص سنی به دست آورد. در این مقاله مدل ماکسیمم آنتروپی را با سایر مدل‌های کلاسیک مرگ‌ومیر تصادفی به خوبی برازش داده شده مقایسه

1. Half-Normal

2. Skew-bimodal-Normal distribution

می‌شوند و رتبه هریک از مدل‌ها باتوجه به قدرت پیش‌گویی نرخ فوت آینده تعیین می‌شود. قابل‌توجه است که وجه اصلی تمایز این مقاله با اکثر پژوهش‌های انجام‌دهنده‌های مورد استفاده است. بیشتر مدل‌سازی‌های صورت گرفته در ایران با استفاده از بانک اطلاعات مرگ‌ومیر (HMD) انجام شده است که این بانک اطلاعاتی شامل داده‌های مرگ‌ومیر ۵۱ کشور است اما متأسفانه داده‌های مربوط به کشور ایران در آن بارگذاری نشده است. از جمله این پژوهش‌ها می‌توان به مقالات کشمرزی و شعاعی (۱۴۰۱) و ذکایی و آل‌حسینی (۱۳۹۳) اشاره کرد. اگرچه الگوی مرگ‌ومیر کشور فرانسه به الگوی مرگ‌ومیر ایران تاحدودی شبیه است اما استفاده از داده‌های ملی در برآورد دقیق‌تر و درست‌تر نرخ فوت کمک زیادی به جمعیت‌شناسان و بیم‌سنج‌ها می‌کند. همچنین باتوجه به مطالب گفته شده در بیشتر مدل‌سازی‌ها از خانواده مدل‌های لی‌کارتر استفاده می‌شود، اما در این پژوهش عملکرد پیش‌بینی مدل‌های کلاسیک با رویکرد نسبتاً جدید مدل‌سازی بر مبنای توزیع سن در زمان فوت مقایسه می‌شوند.

چارچوب نظری پژوهش

(۱) نمادها

ساده‌ترین معیار مرگ‌ومیر تعداد مرگ‌ومیر است. با این حال، برای اهداف عملی کاربرد چندانی ندارد، زیرا به شدت تحت تأثیر تعداد افرادی است که در معرض خطر مرگ هستند. به همین دلیل، جمعیت‌شناسان معمولاً به جای تعداد مرگ‌ومیر از نرخ فوت استفاده می‌کنند. ساده‌ترین نرخ فوت قابل‌تصور از تقسیم تعداد کل فوت‌های یک جمعیت در یک دوره زمانی بر کل جمعیت مورد بررسی به‌دست می‌آید. به این معیار نرخ فوت خام می‌گویند. دوره زمانی مورد بررسی معمولاً یک سال تقویمی است.

❖ m_x (نرخ فوت خام): نرخ فوت خام از تقسیم تعداد فوتی افراد در سن x در دوره زمانی سال (t) بر جمعیت میانه آن سال به‌دست می‌آید. اگر جمعیت پایا باشد (جمعیتی که تعداد

افراد در سن x که در سال t فوت می‌کنند برابر است با تعداد افرادی که در این سال به سن x می‌رسند، می‌توان به جای جمعیت در میانه سال از جمعیت در پایان سال استفاده کرد و می‌توان گفت اصولاً نرخ فوت برابر است با امید ریاضی فوت‌ها تقسیم بر تعداد افرادی که در معرض خطر قرار دارند. $m_{x,t} = d_{x,t}/e_{x,t}$

❖ $q_{x,t}$ (احتمال مرگ): نشان‌دهنده احتمال فوت فرد x ساله در سال تقویمی t است. همچنین رابطه بین دو شاخص نرخ فوت و احتمال مرگ به صورت زیر بیان می‌شود:

$$m_{x,t} = -\log(1 - q_{x,t})$$

اکثر مدل‌های تصادفی مرگ و میر که از نرخ فوت ($m_{x,t}$) یا احتمال مرگ ($q_{x,t}$) برای مدل‌سازی استفاده می‌کنند به فرم زیر هستند:

$$\log m_{x,t} = \sum_t \beta_x^{(i)} k_t^{(i)} \gamma_{t-x}^{(i)}$$

$$\text{logit } q_{x,t} = \sum_t \beta_x^{(i)} k_t^{(i)} \gamma_{t-x}^{(i)}$$

که در آن پارامتر وابسته به سن، k_t پارامتر وابسته به زمان و $\gamma_{t,x}$ پارامتر مربوط به گروه است.

استفاده از نرخ فوت و یا احتمالات مرگ و میر بیشترین کاربرد و علاقه را در بین بیم‌سنج‌ها در امر مدل‌سازی و پیش‌بینی مرگ و میر دارد. رویکرد اکثر روش‌های پیشنهادی و تثبیت‌شده برای مدل‌سازی و پیش‌بینی مرگ و میر انسان استفاده از نرخ فوت و برآورد پارامترهای مدل‌های تصادفی مرگ و میر براساس تابع ماکسیمم درست‌نمایی است. یکی از دلایل اصلی استفاده از این شاخص جمعیت‌شناختی در مطالعه و پیش‌بینی مرگ و میر این است که آن‌ها به راحتی تحولات مرگ و میر را در طول سن و زمان به تصویر می‌کشند و در واقع نمایش مستقیمی از تغییرات در الگوی سنی مرگ و میر ارائه می‌کنند. همچنین استفاده از لگاریتم مرگ و میر به ما این اطمینان را می‌دهد که ضرورتاً نرخ‌ها مثبت خواهند بود.

در این بخش، ابتدا یک دسته‌بندی کلی در خصوص مدل‌های استفاده شده در این مقاله انجام می‌شود و در مقاله ای دیگر به صورت مفصل به معرفی هریک از آنها پرداخته شده است.

جدول ۱: معرفی مدل‌ها

نام مدل	فرمول مدل	شناسایی‌پذیری مدل
مدل‌های کلاسیک تصادفی		
لی کارتر	$\ln m(x, t) = \alpha_x + \beta_x k_t + \epsilon_{x,t}$	مشکل شناسایی‌پذیری دارد.
رینشوهابرم	$\log(m_{x,t}) = \alpha_x + \beta_x^{(1)} k_t + \beta_x^{(2)} \gamma_{t-x} + \epsilon_{x,t}$	مشکل شناسایی‌پذیری دارد.
کوری	$\log(m_{x,t}) = \beta_x + \frac{1}{n_a} k_t + \frac{1}{n_a} \gamma_{t-x} + \epsilon_{x,t}$	مشکل شناسایی‌پذیری دارد و نیاز به یک قید اضافه‌تر نیز دارد.
CBD	$\log it q(x, t) = \beta_x^{(1)} k_t^{(1)} + \beta_x^{(2)} k_t^{(2)}$	مشکل شناسایی‌پذیری ندارد
CMI	$\log it q(x, t) = k_t^{(1)} + k_t^{(2)}(x - \bar{x}) + k_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}$	مشکل شناسایی‌پذیری دارد
پلیت	$\ln m(x, t) = \alpha_x + k_t^{(1)} + k_t^{(2)}(\bar{x} - x) + k_t^{(3)}(\bar{x} - x)^+ + \gamma_{t-x}$	مشکل شناسایی‌پذیری دارد
هیندمن اولا	$y_t(x_i) = f_t(x_i) + \epsilon_t(x_i)\epsilon_{x,t}$	مشکل شناسایی‌پذیری دارد
چند متغیره گام تصادفی با رانش	$\log(m_{x,t}) = \theta_n + \log(m_{x,t-1}) + \epsilon_{x,t}$	مشکل شناسایی‌پذیری ندارد
لی لی	$\ln(m_{x,t}^v) = \alpha_x^v + B_x K_t + b_x^v k_t^v + \epsilon_{x,t}^v$	مشکل شناسایی‌پذیری دارد
مدل توزیع سن در زمان فوت		
ماکسیمم آنتروپی	$y_{n,t} = \log \tilde{\mu}_{n,t} - \log \tilde{\mu}_{n,t-1} $	مشکل شناسایی‌پذیری ندارد

۲) مدل لی کارتر

مدل تصادفی لی کارتر، یکی از مهم‌ترین مدل‌های دو عاملی برای پیش‌بینی مرگ‌ومیر است (Lee and Carter, 1992). این مدل به دلیل عملکرد نسبتاً خوبی که در برآورد نرخ‌های مرگ‌ومیر داشته است، در میان بیم‌سنج‌ها و جمعیت‌شناسان شهرت یافته است. روش لی کارتر به‌عنوان یک روش برون‌یابی، ترکیبی از یک مدل جمعیت‌شناسی غنی (با کمترین پارامتر) و روش‌های سری زمانی است. این مدل به وسیله مولفه‌های اصلی در این بخش تخمین زده می‌شود و با کمک

تجزیه ماتریسی، مولفه‌های مستقل مرگ‌ومیر و همچنین الگوهای سنی و اهمیت آنان در طول زمان شناسایی می‌شوند. اگرچه در این روش هم مانند سایر روش‌های برون‌یابی، اطلاعات پیرامون تأثیرات حاصل از پیشرفت‌های پزشکی، رفتاری، یا اجتماعی در محاسبه مرگ‌ومیرها بی‌تأثیر هستند، اما بنابر دلایلی استفاده از آن بر سایر روش‌های برون‌یابی برتری دارد. اول، پارامترهای مدل قابلیت تفسیرپذیری دارند و اغلب اوقات، مدل قدم زدن تصادفی با رانش برای برون‌یابی روند آتی نرخ مرگ‌ومیر مناسب است. دوم، در کشورهای صنعتی و توسعه‌یافته، بخش زیادی از تغییرات در نرخ مرگ‌ومیر کل کشور به کمک این مدل پوشش داده می‌شود؛ سوم، علاوه بر پیش‌بینی تکی نرخ‌های مرگ‌ومیر قادر به ارائه بازه‌های اطمینان متناظر با آن‌ها نیز است. در کتب جمعیت‌شناسی از این روش به‌عنوان مدل آماری برجسته در پیش‌بینی بلندمدت میزان مرگ‌ومیر کل جمعیت یاد شده است

۳) مدل رینشو- هابرمین

این مدل توسط (Renshaw and Haberman, 2006) ارائه شد. در واقع این مدل با اضافه کردن اثر هم‌گروهی، کلاس مدل لی‌کارت را به کلاس وسیع‌تری از مدل‌های غیرخطی، پارامتری و تعمیم‌یافته گسترش داده است و این امکان را می‌دهد تا مدل‌سازی و برون‌یابی اثرات گروهی در دوره خاص انجام شود. مطالعه هم‌گروهی مطالعه‌ای است بر روی گروهی از افراد با حداقل یک تجربه مشترک خاص در یک دوره زمانی خاص.

۴) مدل کوری

این مدل توسط (Currie, 2006) معرفی شد. این مدل در واقع ساده شده مدل رینشو هابرمین است به طوری که اگر در مدل رینشو هابرمین پارامترها را به صورت $\beta_x^{(1)} = \beta_x^{(2)} = \frac{1}{n_a}$ تعریف کنیم، به مدل APC می‌رسیم. و در این مدل n_a بیانگر تعداد افراد در هر سن است.

۵) مدل CBD^۱

این مدل در سال ۲۰۰۶ توسط کارینز^۲ ارائه شد که در آن به جای استفاده از نرخ فوت از احتمال مرگ استفاده می‌شود (Cairns, Blake and Dowd, 2006).

۶) مدل CMI

این مدل در سال ۲۰۰۷ معرفی شد و یک تعمیم از مدل CBD است.

۷) مدل پلیت

این مدل در سال ۲۰۰۹ معرفی شد. این مدل چهار متغیر تصادفی دارد. $k_t^{(1)}$ بیانگر تغییرات در سطح مرگ‌ومیر برای همه سنین است. عامل $k_t^{(2)}$ اجازه می‌دهد تا تغییرات در مرگ‌ومیر بین سنین متفاوت باشد. همچنین به نظر می‌رسد پویایی مرگ‌ومیر در سنین پایین‌تر (تا سن ۴۰-۵۰ سال) می‌تواند در برخی مواقع متفاوت باشد، ضریب $k_t^{(3)}$ برای ثبت این پویایی‌ها اضافه شده است.

۸) مدل هیندمن اولاً

در این روش مدل‌بندی مرگ‌ومیر با نگاهی تابعی به لگاریتم نرخ فوت خام هر سال است. در این مدل نمودار لگاریتم نرخ فوت به شکل تابعی از سن x_1, \dots, x_p در نظر گرفته می‌شود. این روش نسبت به سال‌های دورافتاده مثل سال‌های جنگ یا شیوع بیماری خاص استوار است و امکان اعمال قیده‌های مختلف را به ما می‌دهد (Hyndman and Shahid Ullah, 2007).

فرض اصلی در این روش بدین صورت است که اگر $y_t(x_i)$ بیانگر لگاریتم نرخ فوت خام مشاهده شده در سن x و زمان t باشد، تابع پیوسته همواری همچون $f_t(x_i)$ در زمان t وجود دارد، به طوری که با کمی خطا در نقاط گسسته سن x_1, \dots, x_p قابل مشاهده است. بنابراین، مشاهده‌ها

1. Cairns-Blake-Dowd
2. Cairns

در این روش به شکل $\{x_i, y_i(x_i)\}$ ، $(i = 1, \dots, p)$ ، $(t=1, \dots, n)$ و $x \in [x_1, x_p]$ نمایش داده می‌شوند که در آن مقدار از نوفه موجود در سری است که با تغییر سن تغییر می‌کند (وجود این مؤلفه در خطای مدل، به دلیل تصحیح فرض هم واریانسی در مدل لی‌کارت است) (ذکایی و آل حسینی، ۱۳۹۳).

۹) مدل چند متغیره گام تصادفی با رانش

این مدل نشان‌دهنده یک برون‌یابی خطی ساده از لگاریتم نرخ فوت خام خاص سن، بر اساس اولین و آخرین مقادیر فوت مشاهده شده در سری زمانی چندمتغیره است (Pascariu, Lenart and Canudas-Romo, 2019).

۱۰) مدل مرگ‌ومیر چند جمعیتی لی‌لی

مدل لی‌لی نیز تعمیمی از مدل کلاسیک لی‌کارت است که در سال ۲۰۰۵ معرفی شد. مدل لی‌لی به جای در نظر گرفتن تنها داده‌های یک کشور، داده‌های چندین کشور را در نظر می‌گیرد. ایده آن‌ها در مورد مدل‌سازی چندجمعیتی ناشی از وجود همبستگی بین گروه‌هایی است که نزدیک به یکدیگر هستند: در چنین شرایطی، بهبود مرگ‌ومیر در یک کشور بر بهبود مرگ‌ومیر در کشور دیگر تأثیر می‌گذارد و منجر به کاهش همبستگی در نرخ مرگ‌ومیر می‌شود.

α_x^T پارامتر خاص سنی است و شکل کلی مرگ‌ومیر را برای جمعیت v نشان می‌دهد. $B_x K_x^T$ برای جلوگیری از انحرافات گروه‌های فرعی در یک جمعیت تعریف می‌شود و می‌توان گفت بیانگر تغییرات مرگ‌ومیر در کل گروه است. این فاکتور مشترک برای همه جمعیت‌ها به کمک روش SVD که در مدل لی‌کارت استفاده شد، به دست می‌آید و پارامتر B_x بیانگر حساسیت سن x به اثر دوره است. پارامتر K_x^T اثر دوره است و برای تعدیل میانگین امید زندگی کل گروه تنظیم شده است. یک روش قدم زدن تصادفی با رانش بر روی پارامتر K_x^T اعمال می‌شود تا روند رایج

در مرگ‌ومیر آینده پیش‌بینی شود. B بیانگر رانش مدل و e_t متغیر تصادفی نرمال مستقل و هم‌توزیع هستند (Li and Lee, 2005).

۱۱) مدل ماکسیمم آنتروپی

آنتروپی مقداری برای اندازه‌گیری میزان عدم قطعیت یک متغیر تصادفی را مشخص می‌کند. این اندازه برای نخستین بار توسط شانون در سال ۱۹۴۸ و سپس توسط (Shannon and Weaver, 1949) معرفی شده است و می‌توان گفت بهترین شکل ریاضی معرفی شده است زیرا ماکسیمم کردن آن با کمک روش لاگرانژ به آسانی انجام می‌شود و به دلیل مقعر بودن تابع، دارای ماکسیمم نسبی است.

با توجه به اصل ماکسیمم آنتروپی، با شرط داشتن اطلاعات بخصوصی از یک متغیر تصادفی و یا یک بردار از متغیرهای تصادفی بایستی توزیعی را برای آن انتخاب کرد که علاوه بر داشتن اطلاعات مورد نظر، دارای ماکسیمم عدم قطعیت باشد. در این بخش از یک روش جدید و نوآورانه برای مدل‌سازی و پیش‌بینی مرگ‌ومیر با استفاده از شکل تابع چگالی و گشتاورهای آماری استفاده شده است و با استفاده از گشتاورهای آماری و تابع چگالی توزیع فوت به پیش‌بینی سطوح مرگ‌ومیر خاص سنی پرداخته شده است. یک گشتاور آماری یکی از مشخصه‌های توزیعی که به آن تعلق دارد را توصیف می‌کند. اگرچه برای تعیین تابع چگالی یک توزیع، مجموع تمام گشتاورها به صورت منحصر به فرد نیاز است اما داشتن اطلاعات در مورد تمام گشتاورها تا بی‌نهایت امری غیرممکن است. روش ارائه شده در این بخش مسئله گشتاور متناهی را در نظر می‌گیرد که در آن چگالی مثبت، $f(x)$ را می‌توان با کمک تعداد محدودی از گشتاورها به دست آورد. ما چند گشتاور مشاهده شده از توزیع سن در زمان فوت را ارزیابی می‌شود تا با استفاده از مدل‌های سری زمانی چندمتغیره به پیش‌بینی آن‌ها پرداخته شود و توزیع پیش‌بینی را با استفاده از رویکرد ماکسیمم آنتروپی ارائه شده

توسط مید و پاپانویکا (۱۹۸۴) که بر میانگین نرخ تولید اطلاعات تصادفی یعنی آنترופی اطلاعات تکیه دارد، بازسازی می‌شود.

۱۲) مشکل گشتاور متناهی

بازسازی چگالی احتمال $f(x)$ ، که تنها تعداد محدودی از گشتاورهای آن مشخص است، کار ساده‌ای نیست و در ادبیات ریاضی به‌عنوان مسئله گشتاور متناهی شناخته می‌شود. اخیراً روش‌های زیادی برای حل این مشکل پیشنهاد شده است، اما در حالت کلی هدف تمام روش‌ها ساخت دنباله‌های خاصی از توابع $f_N(x)$ است که در نهایت به توزیع واقعی $f(x)$ همگرا شوند، زیرا تعداد گشتاورهای N به بی‌نهایت میل می‌کنند به عبارتی تابعی را باید یافت که $N + 1$ گشتاور اول آن با گشتاورهای واقعی $j = 0, 1, \dots, N$ و μ_j برابر باشد.

در ابتدا قیود مورد نیاز برای چگالی را بیان می‌شود:

$$f(x) \geq 0 \quad \forall x \in s,$$

$$\int_a^w f(x) dx = 1.$$

با توجه به تابع گشتاور آماری داریم:

$$\widehat{\mu}_n = \int_a^w x^n f_N(x) dx, \quad n = 0, 1, 2, \dots, N. \quad (8)$$

در رابطه (8)، N تعداد محدودیت‌های گشتاوری، x^n گشتاورهای مشخصه و $\widehat{\mu}_n$ برآورد نمونه‌ای n امین گشتاور خام است. رابطه (8) شامل $N + 1$ معادله است که در آن $\widehat{\mu}_0, \dots, \widehat{\mu}_N$ از تابع چگالی $f_N(x)$ به دست می‌آیند.

معمولاً از ۳ تا ۶ گشتاور برای پیدا کردن تناسب خوبی از چگالی واقعی استفاده می‌شود. در این قسمت از بازسازی ماکسیم آنترופی و الگوریتم‌های توسعه یافته به‌عنوان یک روش مشخص برای ساخت دنباله‌ای از تقریب‌ها نزدیک به چگالی واقعی پیروی می‌کنیم.

این روش براساس آنتروپی اطلاعات داده شده وابسته به تابع چگالی است. به طوری که تابع H ، ماکسیمم شود.

$$H = - \int_a^w f(x) \log f(x) dx$$

w و a به ترتیب حداقل و حداکثر سن هستند. لازم به ذکر است در این بخش از گشتاورهای خام استفاده شده است.

آنتروپی تابع چگالی f یک تابع مقعر بر حسب f بوده و دارای ماکسیمم است. حال از روش لاگرانژ اویلر برای m گشتاور می‌توان نوشت:

$$\begin{aligned} L &= \int_a^w -f(x) \log f(x) dx + \lambda_0 \left(\int_a^w f(x) dx - 1 \right) + \sum_{n=1}^N \lambda_n \int_a^w f(x) x^n dx - \mu_n \\ &= \int_a^w \left[-f(x) \log f(x) + \lambda_0 f(x) + f(x) \sum_{n=1}^N \lambda_n x^n \right] dx - \left(\lambda_0 + \sum_{n=1}^N \lambda_n \mu_n \right). \end{aligned}$$

با استفاده از معادله اویلر

$$\frac{\partial L}{\partial f} - f \log f + (\lambda_0) f + \left(\sum_{n=1}^N \lambda_n x^n \right) f = 0,$$

یا

$$f(x) = f_N(x) = \exp \left\{ - \sum_{n=0}^N \lambda_n x^n \right\} = \exp \left(-\lambda_0 - \sum_{n=1}^N \lambda_n x^n \right). \quad (9)$$

رابطه (9)، فرم کلی تابع چگالی احتمال ماکسیمم آنتروپی (ME) است.

برای پیدا کردن یک شکل عمومی از چگالی، ME معادله (9) را در تابع لاگرانژ جایگزین کرده و مجددا بررسی می‌کنیم

$$\begin{aligned} L &= E(-\log f(x)) = \int_{-\infty}^{+\infty} f(x)(\lambda_0 + \lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_n x^n) dx \\ &= \lambda_0 + \lambda_1 E(x^1) + \lambda_2 E(x^2) + \dots + \lambda_n E(x^n) \\ &= \lambda_0 + \lambda_1 \mu_1 + \lambda_2 \mu_2 + \dots + \lambda_n \mu_n = \lambda_0 + \sum_{n=1}^N \lambda_n \mu_n \end{aligned}$$

از طرفی n امین گشتاور خام را می‌توان به صورت زیر بازنویسی کرد:

$$\mu_n = \int_a^{\omega} x^n \exp\left(-\lambda_0 - \sum_{n=1}^N \lambda_n x^n\right) dx \quad (10)$$

با توجه به دانستن اولین گشتاور از بین $N+1$ تا گشتاور موجود، معادلات (9) و (10) را باید به عنوان یک معادله غیرخطی از $N+1$ معادله برای ضرایب لاگرانژ $\lambda_0, \lambda_1, \dots, \lambda_N$ در نظر گرفت. با توجه به اینکه در تابع چگالی نرمال شده $f(x)$ ، اولین گشتاور همواره برابر با ۱ است، با توجه به رابطه (10) خواهیم داشت:

$$\mu_0 = \int_a^{\omega} x^0 f_N(x) dx = \int_a^{\omega} \exp\left(-\lambda_0 - \sum_{n=1}^N \lambda_n x^n\right) dx = 1, \quad (11)$$

با توجه به رابطه (11) اولین ضریب لاگرانژ (λ_0) بر حسب ضرایب باقیمانده لاگرانژ به دست می‌آید.

$$\int_a^{\omega} \exp\left(-\sum_{n=1}^N \lambda_n x^n\right) dx = e^{\lambda_0} = Z, \quad (12)$$

و با توجه به روابط (9) و (10) خواهیم داشت:

$$\mu_n = \frac{\int_a^{\omega} x^n \exp(-\sum_{n=1}^N \lambda_n x^n)}{\int_a^{\omega} \exp(-\sum_{n=1}^N \lambda_n x^n)}, \quad n = 0, 1, 2, \dots, N \quad (13)$$

برای حل ضرایب لاگرانژ $\lambda = [\lambda_1, \dots, \lambda_N]$ از روش نیوتن رافسون استفاده می‌شود

$$L = \log Z + \sum_{n=1}^N \mu_n \lambda_n \quad (14)$$

که در آن مقدار عددی گشتاور واقعی است.

برای حل ضرایب لاگرانژ $\lambda = [\lambda_1, \dots, \lambda_N]^+$ از روش نیوتن رافسون، از روش به‌هنگام‌سازی

متوالی استفاده می‌شود

$$\lambda_{(i)} = \lambda_{(i-1)} - H^{-1} \frac{\partial L}{\partial \lambda_n}$$

که شیب

$$\frac{\partial L}{\partial \lambda_n} = \mu_n - \frac{\int_a^\omega x^n \exp(-\sum_{n=1}^N \lambda_n x^n) dx}{\int_a^\omega \exp(-\sum_{n=1}^N \lambda_n x^n) dx} = \mu_n - \mu_n(\lambda), \quad n = 0, 1, 2, \dots, N \quad (15)$$

و ماتریس هسین

$$H_{nm} = \frac{\partial^2 L}{\partial \lambda_n \partial \lambda_m} = \mu_{n+m}(\lambda) - \mu_n(\lambda) \mu_m(\lambda),$$

$$\mu_{n+m}(\lambda) = \frac{\int_a^\omega x^{n+m} \exp(-\sum_{n=1}^N \lambda_n x^n) dx}{\int_a^\omega \exp(-\sum_{n=1}^N \lambda_n x^n) dx}, \quad n, m = 0, 1, \dots, N$$

به دلیل اینکه ماتریس هسین H در همه جا محدب و مثبت معین است، جوابی منحصر به فرد

برای معادله فوق وجود دارد. مزیت این روش به این دلیل است که موضوع حداکثرسازی اندکی

از محدودیت‌های گشتاوری را می‌توان به آسانی روش الگوریتم نیوتن رافسون انجام داد.

۱۳) ایده ماکسیمم آنتروپی و الگوریتم مدل‌سازی مرگ و میر

ایده این مدل برای پیش‌بینی ساده است. سطوح مرگ و میر خاص سن برای یک جمعیت در زمان آینده، با برون‌یابی تعداد محدودی از گشتاورهای آماری ارائه شده توسط شاخص توزیع سن در زمان فوت جدول مرگ و میر (d_x) تعیین می‌شود. برای برون‌یابی، از مدل‌های سری زمانی چندمتغیره استفاده می‌شود و توزیع سن در زمان فوت در هر زمان دلخواه از گشتاورهای پیش‌بینی شده با استفاده از الگوریتم ME به دست می‌آید. در ابتدا گشتاورهای ترتیبی سوم به بعد نرمال می‌شوند و یک تبدیل لگاریتمی به مقادیر مطلق تمام گشتاورهای مشاهده شده اعمال می‌شود. استفاده از مقادیر مطلق گشتاورهای مشاهده شده، تضمینی برای مثبت بودن معیارهای شکل مربوطه در هر افق پیش‌بینی است. شاخص دوره مورد استفاده در پیش‌بینی، با تفاوت مرتبه اول گشتاورهای مرکزی مشاهده شده، $\tilde{\mu}_{n,t}$ به این صورت است

$$y_{n,t} = \log |\tilde{\mu}_{n,t}| - \log |\tilde{\mu}_{n,t-1}|. \quad (16)$$

از یک قدم زدن چندمتغیره به همراه بردار پارامترهای رانش θ_n ، برای سنجش پویایی شاخص‌های دوره چندگانه استفاده می‌کنیم.

$$y_{n,t} = \theta_n + y_{n,t-1} + \varepsilon_{n,t}, \quad t = 1, \dots, \tau, \quad n = 1, 2, \dots, N \quad (17)$$

که $\varepsilon_{n,t} \sim N(0, \Omega)$ و $\Omega = CC'$ همچنین C نشان دهنده ماتریس فاکتورسازی چولسکی از ماتریس واریانس کوواریانس Ω است. پارامترهای θ_n به روش کمترین مربعات معمولی برآورد می‌شوند.

الگوریتم مدل‌سازی مرگ‌ومیر

1. در ابتدا داده‌ها شامل تعداد فوتی‌ها و جمعیت کشور ایران به ترتیب از سازمان ثبت احوال و مرکز آمار ایران برای ۹ سال متوالی اخذ می‌شود.
2. داده‌ها به دو پنجره پسر و پیشرو تقسیم می‌شوند.
3. حدود ۷۰ درصد از داده‌ها مربوط به پنجره پسر و مابقی مربوط به پنجره پیشرو است.
4. تمامی مدل‌های مرگ‌ومیر تصادفی مبتنی بر نرخ فوت و توزیع سن در زمان فوت بر روی پنجره پسر و پیاده‌سازی می‌شوند و برای تجزیه و تحلیل برازش مدل‌ها از آن‌ها استفاده می‌شود. برازش مدل‌ها شامل:
مدل لی کارتر مدل رینشو هابرمن مدل APC مدل CBD مدل CMI مدل پلینت
مدل هیندمن اولاً مدل لی لی ماکسیمم آنتروپی با ۴، ۵ و ۶ گشتاور مدل قدم زدن تصادفی
5. در پنجره پیشرو به پیش‌بینی نرخ‌های فوت خام براساس روش نمونه‌گیری خارج از نمونه پرداخته می‌شود.
6. در پنجره پیشرو نرخ‌های فوت خام مشاهده شده با نرخ‌های فوت خام پیش‌بینی شده برای هر یک از مدل‌ها با هم مقایسه می‌شوند.
7. با توجه به معیارهای خطای معرفی شده در قبل خطای گزارش شده هر مدل ارائه می‌شود.
8. با توجه به خطای گزارش شده، قدرت پیش‌گویی مدل‌ها با هم مقایسه می‌شوند.
9. در پایان به تحلیل نتایج و ارائه جدول پرداخته می‌شود.

روش تحقیق و داده‌ها

یکی از روش‌های جدید مبتنی بر توزیع سن در زمان فوت، استفاده از شکل تابع چگالی و گشتاورهای آماری است که با استفاده از گشتاورهای آماری و تابع چگالی توزیع فوت به پیش‌بینی سطوح مرگ‌ومیر خاص سنی می‌پردازد (Pascariu, Lenart and Canudas-Romo, 2019). روش سری زمانی برای برون‌یابی تعداد محدودی از گشتاورها استفاده می‌شود و به کمک گشتاورهای برون‌یابی شده به بازسازی توزیع سن در زمان فوت با کمک روش ماکسیمم

آنتروپی پرداخته می‌شود. این روش بر میانگین نرخ تولید اطلاعات توسط یک منبع تصادفی داده یا تابع چگالی، یعنی آنتروپی اطلاعات تکیه دارد.

در این مقاله از داده‌های مرگ و میر کشور ایران برای مدل‌سازی و پیش‌بینی نرخ فوت استفاده شده است. داده‌های واقعی شامل تعداد فوتی‌ها و جمعیت کشور (مرگ و میر) ایران، برای ۹ سال متوالی (۱۳۹۰-۱۳۹۸) و برای بازه سنی ۰ تا ۱۰۰ سال و به تفکیک هر دو جنسیت است که به ترتیب از سازمان ثبت احوال کشور و مرکز آمار ایران اخذ شده است. در ابتدا داده‌ها به دو بازه تقسیم می‌شوند. پنجره پسر و شامل داده‌ها از سال ۱۳۹۰ تا ۱۳۹۵ است و پنجره پیشرو شامل داده‌های سه سال پایانی ۱۳۹۷ تا ۱۳۹۸ است. در ادامه تمامی مدل‌های مرگ و میر تصادفی مبتنی بر نرخ فوت و توزیع سن در زمان فوت بر روی پنجره پسر و پیاده‌سازی می‌شوند و برای تجزیه و تحلیل برازش مدل‌ها از آن‌ها استفاده می‌شود. سپس در پنجره پیشرو به پیش‌بینی نرخ‌های فوت خام براساس روش نمونه‌گیری خارج از نمونه پرداخته می‌شود و نرخ‌های فوت خام مشاهده شده با نرخ‌های فوت خام پیش‌بینی شده برای هر یک از مدل‌ها با هم مقایسه می‌شوند و با توجه به معیارهای خطای معرفی شده در این بخش، خطای گزارش شده هر مدل در قالب جدول ارائه می‌شود. در نهایت با توجه به خطای گزارش شده، قدرت پیش‌گویی مدل‌ها با هم مقایسه می‌شوند و بهترین مدل تصادفی برای پیش‌بینی نرخ فوت کشور ایران برحسب جنسیت معرفی می‌گردد.

برای مقایسه دقت پیش‌بینی مدل‌ها از معیارهای ارزیابی زیر استفاده شده است. شایان ذکر است که در این مقاله

$$\delta(x, t) = m_{x,t} - \widehat{m}_{x,t}, \quad x = 0, 1, \dots, w, \quad t = 1, 2, \dots, T$$

که w حداکثر سن مورد بررسی و t تعداد سال‌های پیش‌بینی شده است که در واقع فاصله بین نرخ فوت خام مشاهده شده و پیش‌بینی شده است. همچنین حداکثر سن مورد بررسی برای تمامی افراد ۱۰۰ سال است و در سال‌های ۱۳۹۶ تا ۱۳۹۸ اختلاف نرخ فوت پیش‌بینی

شده با نرخ فوت واقعی محاسبه می‌شود. قابل توجه است خطاهای معرفی شده وابسته به داده مورد بررسی هستند و همچنین معیارهای ارزیابی دقت بر مبنای $\delta(x, t)$ هستند و نمی‌توان از آن‌ها برای مقایسه سری‌هایی که در مقیاس‌های مختلف هستند استفاده کرد. در این مقاله ۷ معیارهای ارزیابی خطا در سه طبقه (خطاهای وابسته به مقیاس^۱، درصد خطا^۲ و خطاهای مقیاس شده^۳) تعریف شده است. زیرا ارزیابی دقت پیش‌بینی را تنها می‌توان با در نظر گرفتن عملکرد یک مدل بر روی داده‌های جدیدی که هنگام برازش مدل استفاده نشده است، تعیین کرد و توجه به اینکه یک مدل چقدر با داده‌های تاریخی مطابقت دارد، کافی نیست. هنگام ارزیابی مدل‌ها، استفاده از بخشی از داده‌های موجود برای آزمایش (برازش)^۴ و استفاده از بقیه داده‌ها برای تخمین یا آموزش (پیش‌بینی)^۵ مدل رایج است. سپس داده‌های آزمایشی را می‌توان برای اندازه‌گیری میزان احتمال پیش‌بینی مدل بر روی داده‌های جدید مورد استفاده قرار داد. اندازه مجموعه داده‌های استفاده شده برای برازش مدل معمولاً حدود ۲۰٪ از کل نمونه است، اگرچه این مقدار به مدت زمان نمونه مورد بررسی و افق پیش‌بینی بستگی دارد. تعداد سال‌های در نظر گرفته شده برای برازش مدل، در حالت ایده‌آل باید حداقل به اندازه حداکثر افق پیش‌بینی مورد نیاز باشد. شایان ذکر است که برازش معقول یک مدل با داده‌ها به معنی پیش‌بینی خوب آن مدل نیست و به کمک مدلی با تعداد پارامترهای کافی می‌توان به برازش خوبی از مدل دست پیدا کرد.

الف- میانگین قدرمطلق خطا: در آمار، میانگین قدرمطلق خطا (MAE) معیاری از خطاهای بین مشاهدات زوجی است که یک پدیده را بیان می‌کنند. این تابع زیان بدون در نظر گرفتن

-
1. Scale-dependent Errors
 2. Percentage Errors
 3. Scaled Errors
 4. Testing
 5. Training

علامت، از فاصله بین مقادیر پیش‌بینی شده و واقعی به‌عنوان معیار استفاده می‌کند. زمانی که داده پرت یا دورافتاده در مشاهدات وجود داشته باشد، استفاده از تابع زیان میانگین قدر مطلق خطا کارایی بالایی دارد. همچنین درک و محاسبه این معیار برای مقایسه پیش‌بینی‌های صورت گرفته در یک مجموعه داده آسان است (شعاعی و کشمرزی، ۱۴۰۱ و Chen, Z., Yang, 2004).

$$MAE = \frac{\sum_{x=0}^w \sum_{t=1}^T |\delta(x, t)|}{n}$$

ب- ریشه میانگین مربعات خطا: میانگین تفاوت بین مقادیر پیش‌بینی شده توسط یک معیار و مقادیر واقعی را اندازه‌گیری می‌کند و تخمینی از اینکه مدل مورد نظر چقدر قادر به پیش‌بینی مقدار هدف (دقت) است را ارائه می‌کند.

$$RMSE = \sqrt{\frac{\sum_{x=0}^w \sum_{t=1}^T \delta(x, t)^2}{n}}$$

ج- میانگین مربعات خطا: تابع زیان میانگین مربعات خطا، به‌عنوان میانگین اختلاف بین مقادیر واقعی و تخمینی تعریف می‌شود.

$$MSE = \frac{1}{n} \sum_{x=0}^w \sum_{t=1}^T (\delta(x, t))^2$$

د- خطای مطلق نسبی:

$$RAE = \sum_{x=0}^w \sum_{t=1}^T \frac{|\delta(x, t)|}{|\bar{m}_{x,t} - m_{x,t}|}$$

ه- خطای استاندارد باقی مانده:

$$RSE = \sqrt{\frac{1}{n-2} \sum_{x=0}^w \sum_{t=1}^T (\delta(x, t))^2}$$

و- میانگین درصد قدرمطلق خطای متقارن: یک معیار اندازه‌گیری دقت براساس درصد خطا است. در این تابع زیان نیز بدون در نظر گرفتن علامت، تفاوت بین مقادیر پیش‌بینی شده و واقعی بر نصف مجموع مقادیر مطلق واقعی و پیش‌بینی شده تقسیم می‌شود.

$$sMAPE = \frac{100}{n} \sum_{x=0}^w \sum_{t=1}^T \frac{|\delta(x, t)|}{(|m_{x,t}| + |\hat{m}_{x,t}|)/2}$$

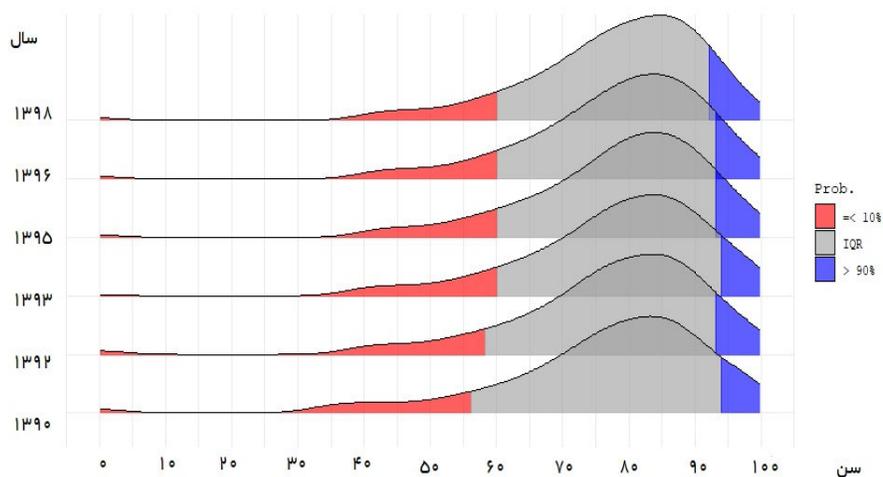
ز- میانگین درصد قدرمطلق خطا: که انحراف میانگین درصد مطلق نیز نامیده می‌شود، دقت یک سیستم پیش‌بینی را اندازه می‌گیرد. این دقت را به صورت درصد اندازه‌گیری می‌کند و می‌تواند به عنوان میانگین درصد خطای مطلق برای هر دوره زمانی منهای مقادیر واقعی تقسیم بر مقادیر واقعی محاسبه شود.

$$MAPE = \frac{100 \times \sum_{x=0}^w \sum_{t=1}^T \left| \frac{\delta(x, t)}{m_{x,t}} \right|}{n}$$

در معیارهای ارزیابی مبتنی بر درصد خطا اگر مقدار مشاهده شده داده مورد بررسی، صفر باشد آنگاه نتیجه در مجموعه تست بی‌نهایت یا تعریف نشده است، و زمانی که هر مقدار واقعی داده نزدیک به صفر باشد، مقادیر بزرگی گزارش می‌شود (Chen, Yang, 2004 and Canudas, 2010)

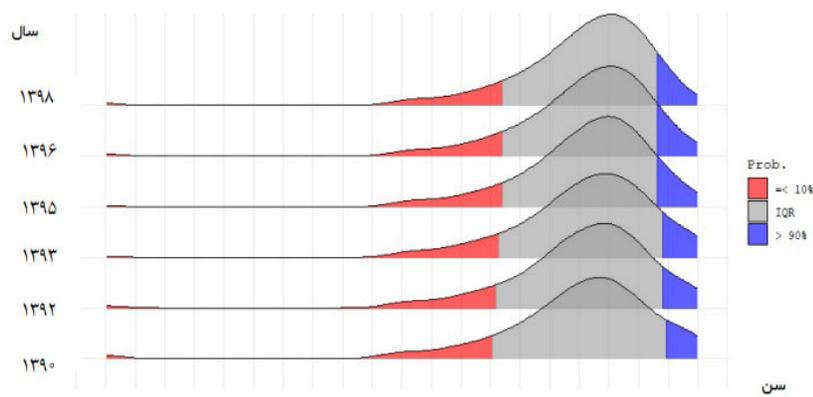
یافته‌ها

باتوجه به نمودار ۱ جمعیت نوزادان (۰ تا ۱) ساله در سال ۱۳۹۰ مرگ‌ومیر بیشتری را نسبت به سال ۱۳۹۸ تجربه کرده‌اند. همچنین در سال ۱۳۹۰، ۱۰ درصد از جمعیت مردان تا قبل از ۵۶ سالگی فوت کردند درحالی‌که این عدد در سال ۱۳۹۸ به ۶۰ سال افزایش پیدا کرده است. بیشترین فراوانی فوت در سال ۱۳۹۰ مربوط به مردان ۸۶ ساله است و این آمار برای سال ۱۳۹۸ مربوط به ۸۷ سالگی است. واضح است با در دست داشتن داده‌ها با بازه‌ی طولانی‌تر درک ما از روند الگوی مرگ‌ومیر بیشتر خواهد بود. قابل توجه است در تمامی ۹ سال داده مشاهده شده تنها ۱۰ درصد از جمعیت مردان ایران شانس زندگی کردن پس از حدود ۹۲ سالگی را داشته‌اند. حداکثر سن در سال ۱۳۹۰ نسبت به سال ۱۳۹۸ کمی بیشتر است که می‌توان علت آن را شیوع بیماری کرونا در نظر گرفت. همچنین ۸۰ درصد از جمعیت در تمامی ۹ سال مورد بررسی در حدود بازه سنی (۶۰-۹۲) سال زندگی کرده‌اند و بنابراین طبق الگوی سنی مشاهده شده می‌توان گفت بیشترین تعداد فوتی مربوط به گروه سنی میان‌سالان و سالمندان است.



نمودار ۱: توزیع سن مشاهده شده در زمان فوت برای جمعیت مردان ایران در سال‌های ۱۳۹۰ تا ۱۳۹۸

نمودار ۲، نشان می‌دهد در سال ۱۳۹۰، ۱۰ درصد از جمعیت زنان تا قبل از ۶۵ سالگی فوت می‌کردند و این عدد در سال ۱۳۹۸ به حدود ۶۷ سال افزایش یافته است. همچنین ۱۰ درصد از جمعیت سالمندان در سال ۱۳۹۰ شانس زندگی تا ۹۵ سال را داشتند و این عدد برای سال ۱۳۹۸ در حدود ۹۳ سال است. بیشترین تعداد فوت مشاهده شده در سال ۱۳۹۰ مربوط به سن ۸۴ سال است که در سال ۱۳۹۸ به سن ۸۷ سال افزایش پیدا کرده است. همچنین با مقایسه نمودار ۱ و ۲ نتیجه‌گیری می‌شود که طول عمر زنان ایران از مردان ایران بیشتر است.



نمودار ۲: توزیع سن مشاهده شده در زمان فوت برای جمعیت زنان ایران در سال‌های ۱۳۹۰ تا ۱۳۹۸

در این پژوهش، عملکرد روش مدل ماکسیمم آنتروپی با (چهار، پنج و شش) گشتاور که براساس مدل‌سازی فراوانی مرگ‌ومیر در توزیع سن در زمان فوت است، با ۹ مدل مرگ‌ومیر تصادفی که بر مدل‌سازی نرخ لگاریتم فوت/مرگ‌ومیر تأکید دارند، مقایسه می‌شود. از آنجایی‌که همیشه می‌توان، برازش کامل داده‌ها را با استفاده از مدلی با پارامترهای کافی به‌دست آورد و با توجه به اینکه برازش خوب لزوماً به معنای عملکرد خوب پیش‌بینی نیست (هیندمن و اولاه، ۲۰۰۷)، مدل‌ها براساس توانایی آن‌ها در برازش داده‌های تاریخی ارزیابی نخواهند شد و مدل‌ها براساس عملکرد پیش‌بینی خارج از نمونه بر روی داده‌های مشاهده شده آزمایش می‌شوند. بدین

صورت که نرخ فوت خام پیش‌بینی شده در پنجره پیشرو برای سال‌های ۱۳۹۸-۱۳۹۷ با نرخ فوت خام مشاهده شده در همین دوره مقایسه می‌شوند و با توجه به معیارهای ارزیابی خطا نتایج آن‌ها در جداول ۲ و ۳ ارائه می‌شود.

در جدول (۲) اعداد مربوط به نرخ فوت خام برای جمعیت مردان ایران، تحت ۷ معیار ارزیابی برازش مدل مختلف برای مدل‌های مرگ و میر تصادفی معرفی شده در این مطالعه قابل مشاهده است. شایان ذکر است تمامی اعداد در ۱۰۰ ضرب شده‌اند. برای تمامی معیارهای خطا، مقدار کوچک‌تر بر مقدار بزرگ‌تر برتری دارد. برای محاسبه رتبه نهایی مدل مجموع رتبه‌ها در هر یک از معیارهای ارزیابی با هم جمع می‌شوند و در نهایت از کوچک به بزرگ با شماره ۱ تا ۱۲ مشخص می‌شوند. لذا همان‌طور که واضح است، مدل لی‌کارتر تقریباً عملکرد بهتری را نسبت به سایر مدل‌های تصادفی در میان جمعیت مردان داشته است. براساس دو معیار میانگین درصد قدرمطلق خطای متقارن و ریشه میانگین مربعات خطا مدل لی‌کارتر رتبه دوم را در بین ۱۲ مدل کسب کرده است و با اختلافی جزئی به ترتیب مدل هیندمن اولا و مدل پلیت رتبه نخست را در دو معیار خطای ذکر شده دارند. براساس ۵ معیار ارزیابی برازش مدل دیگر یعنی (میانگین قدرمطلق خطا، میانگین درصد قدرمطلق خطا، میانگین مربعات، خطای مطلق نسبی و خطای استاندارد باقی‌مانده) مدل لی‌کارتر رتبه نخست را دارد. به صورت کلی می‌توان نتیجه گرفت مدل لی‌کارتر رتبه نخست و مدل هیندمن اولا، رتبه دوم را در بین مدل‌های مورد مقایسه دارد. شایان ذکر است به دلیل کوچک بودن اعداد جداول ۲ و ۳ برای سهولت مقایسه ظاهری، تمامی اعداد در ۱۰۰ ضرب شده‌اند.

در حین پیاده‌سازی هریک از مدل‌ها باید به ۳ نکته دقت داشت که ۱- آیا مدل توانایی تولید مسیرهای نمونه را دارد یا خیر؟ عدم اطمینان به فرآیندی که مسیرهای نمونه تصادفی را ایجاد می‌کند، برای کارهایی نظیر قیمت‌گذاری ابزارهای مالی مرتبط با طول عمر و تدوین استراتژی‌های مربوط به حمایت از امنیت ضروری است، که تمامی مدل‌ها در این قسمت خوب عمل کرده‌اند.

۲- مهم است که اگر اثرات گروهی وجود دارد و باید تمام محدودیت‌های آن را بررسی شود که در بخش قبل توضیح داده شده است. ۳- ساختار همبستگی با اهمیت: در حقیقت این ویژگی توانایی تولید یک ساختار همبستگی غیرارادی بین تغییرات سال به سال در میزان مرگ و میر در سنین مختلف را بررسی می‌کند. تجزیه و تحلیل آماری از میزان مرگ و میر به تغییرات در $m_{x,t}$ در سنین مختلف که میزان در ارتباط هستند، اشاره دارد. ساختار همبستگی زمانی بی اهمیت است، که ارتباط کاملی بین تغییرات در میزان مرگ و میر در سنین مختلف از یک سال به سال دیگر وجود داشته باشد. به عنوان مثال در مورد مدل لی کارتر، در جایی که یک سری فرآیند سری kt وجود دارد، که این مورد را نشان می‌دهد. برای مدل‌های لی و کارتر و رینشو و هابرم ما در تمام سنین به جز در سنین جوانی ارتباط کاملی بین سنین داریم، جایی که اثر تصادفی اضافه‌ای ناشی از ورود یک گروه جدید با اثر گروهی ناشناخته است. مدل‌های خانواده CBD امکان ایجاد یک همبستگی غیرساختاری را فراهم می‌آورند، زیرا همه آن‌ها بیش از یک عامل اساسی مخاطره دوره‌ای را دارند.

در جدول ۳ نرخ فوت خام مربوط به جمعیت زنان ایران، تحت ۷ معیار سنجش خطا ارائه شده است. همان‌طور که مشخص است مدل کوری رتبه نخست را در بین مدل‌های معرفی شده برای پیش‌بینی نرخ فوت خام دارد و پس از آن مدل پلیت رتبه دوم را اخذ می‌کند. براساس دو شاخص (میانگین قدرمطلق خطا و میانگین درصد قدرمطلق خطای متقارن) مدل پلیت و براساس معیار خطای استاندارد باقی مانده، مدل رینشو هابرم عملکرد نسبتاً بهتری نسبت به سایر مدل‌های مرگ و میر دارند.

با مقایسه جدول ۲ با جدول ارائه شده در مقاله (شعاعی و کشمرزی، ۲۰۲۱) درمی‌یابیم که جدول معرفی شده در این مقاله، خطاهای کوچک‌تری را برای مدل‌های مشابه با معیار ارزیابی خطاهای یکسان گزارش می‌کند.

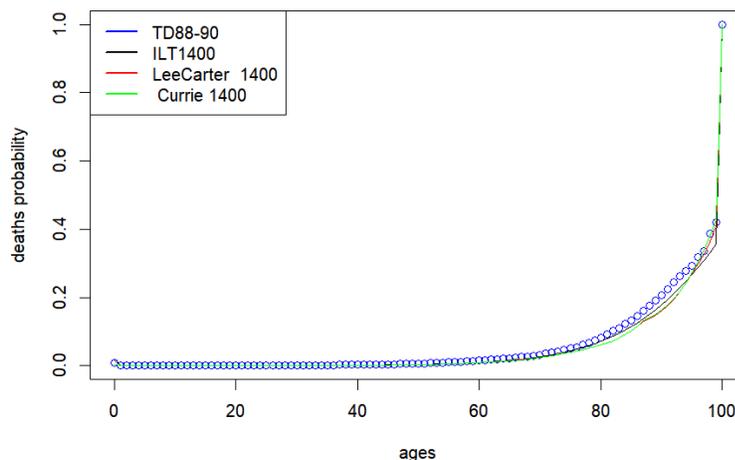
جدول ۲: معیارهای ارزیابی برآزش مدل‌های مرگومیر تصادفی برای جمعیت مردان ایران، تا ۱۰۰ ساله X ۱۰۰

ردیف	خطای استاندارد باقی مانده	خطای مطلق نسبی	میانگین مربعات خطا	میانگین درصد خطا	میانگین مربعات خطا	ریشه میانگین مربعات خطا	میانگین قدرمطلق مقارن	میانگین خطا	مدل
۳	۳/۵۸۶۲(۴)	۶/۷۳۱۲(۳)	۰/۰۲۴۰(۵)	۴/۸۷(۳)	۱/۸۴۶۲(۵)	۰/۶۵۷۲(۳)	۰/۴۱۱(۳)	مدل قدم زدن تصادفی	
۷	۴/۷۵۰۳(۵)	۶/۹۱۱۱(۴)	۰/۰۴۵۷(۷)	۵/۸۷(۹)	۲/۱۴۰۵(۷)	۵/۵۱۲۴(۵)	۰/۴۵۸۹(۵)	مدل لی لی	
۱	۲/۷۴۵(۱)	۵/۶۵۶(۱)	۰/۰۲۶۸(۱)	۴/۶۵۸۳(۱)	۱/۶۵۶۴(۲)	۴/۵۰۴۱(۲)	۰/۳۷۸۱(۱)	مدل لی کارتر	
۲	۲/۳۲۲(۳)	۵/۷۳۸۵(۲)	۰/۰۲۸۱(۳)	۴/۶۸۴۵(۲)	۱/۶۷۸۶(۳)	۴/۴۷۹۱(۱)	۰/۳۸۲۹(۲)	مدل هیندمن اولا	
۱۰	۵/۳۰۷(۸)	۸/۳۵۰۲(۹)	۰/۰۵۲۴(۹)	۶/۷۷۷۹(۱۰)	۲/۲۹۰۰(۹)	۵/۹۲۵(۱۰)	۰/۵۳۲۴(۹)	مدل ریشو هابرمن	
۹	۶/۷۷۸۱(۸)	۸/۹۰۰۹(۱۰)	۰/۰۴۲۷(۱۰)	۵/۵۰۲۷(۵)	۲/۲۹۷۱(۱۰)	۵/۲۰۰۹(۵)	۰/۵۶۵۶(۱۰)	مدل ماکسیم آنروبی با ۴ گشتاور	
۷	۵/۴۶۳(۶)	۷/۳۴۱۱(۷)	۰/۰۴۸۶(۷)	۵/۱۱۲(۶)	۲/۲۰۵۰(۸)	۵/۲۲۰۷(۶)	۰/۵۲۱۵(۸)	مدل ماکسیم آنروبی با ۵ گشتاور	
۵	۰/۵۵۰۵(۸)	۷/۴۰۵۵(۸)	۰/۰۴۳۹(۶)	۵/۳۲۴(۴)	۲/۰۹۵۷(۶)	۵/۱۰۱(۴)	۰/۴۷۷۷(۷)	مدل ماکسیم آنروبی با ۶ گشتاور	
۶	۷/۲۴۵۸(۹)	۷/۲۴۸۵(۶)	۰/۰۲۸۶(۴)	۵/۳۶۷(۸)	۱/۶۹۲۸(۴)	۵/۵۱۲۰(۸)	۰/۴۷۱۲(۶)	مدل APC	
۱۲	۳۲/۹۱۵۵(۱۲)	۳۲/۹۱۵۵(۱۲)	۲/۳۰۹(۱۲)	۱۴۲/۲۸۰۴(۱۲)	۴/۸۰۵۲(۱۲)	۴۰/۹۰۵۲(۱۲)	۱/۵۸۷(۱۲)	مدل CBD	
۱۱	۸/۴۸۶۸(۱۱)	۱۶/۴۳۳۹(۱۱)	۰/۱۲۶۷(۱۱)	۲۵/۳۲۵۴(۱۱)	۳/۵۶۰۱(۱۱)	۲۰/۳۷۱۶(۱۱)	۰/۹۵۹۱(۱۱)	مدل CMI	
۴	۲/۷۶۲۴(۲)	۶/۹۱۵۹(۵)	۰/۰۲۷۴(۲)	۵/۸۲۳(۷)	۱/۶۳۸۹(۱)	۵/۳۳۳(۷)	۰/۴۵۲۱(۴)	مدل PLAT	

جدول ۳. معیارهای ارزیابی برازش مدل‌های مرگ‌ومیر تصادفی برای جمعیت زنان ایران ۰ تا ۱۰۰ ساله × ۱۰۰

ردیف	خطای استاندارد باقی مانده	خطای مطلق نسبی	میانگین مربعات خطا	میانگین درصد خطا	ریشه میانگین مربعات خطا	میانگین قدرمطلق در صد خطای متقارن	میانگین قدرمطلق خطا	مدل
۵	۱/۱۶۱۶(۴)	۵/۳۴۹۹(۶)	۰/۰۱۴۹(۶)	۷/۶۳۶۵(۶)	۱/۲۲۳۳(۶)	۷/۳۶۶۵(۶)	۰/۳۵۶۷(۶)	مدل قلم زدن تصادفی
۳	۱/۰۶۲۵(۳)	۴/۵۹۰۶(۳)	۰/۰۱۰۳(۴)	۶/۵۷۵۵(۲)	۱/۰۱۵۷(۴)	۶/۹۰۵۸(۲)	۰/۳۱۱۷(۳)	مدل لی لی
۵	۲/۰۴۴۸(۵)	۵/۷۴۳۰(۹)	۰/۰۱۹۴(۷)	۷/۳۹۶۶(۵)	۱/۳۹۳۴(۷)	۷/۱۳۹۷(۵)	۰/۳۸۸۴(۹)	مدل لی کارتر
۷	۲/۰۴۰۷(۶)	۵/۷۴۰۰(۷)	۰/۰۱۹۴(۷)	۷/۹۴۸۸(۴)	۱/۳۹۳۷(۸)	۷/۱۳۶۲(۴)	۰/۳۸۸۱(۸)	مدل هیندمن اولا
۴	۰/۰۱۹۱۷(۰)	۴/۶۹۴۸(۴)	۰/۰۰۷۹(۲)	۷/۸۰۲۵(۸)	۰/۸۹۲۲(۲)	۷/۶۷۲۴(۸)	۰/۳۱۴۳(۴)	مدل رینشو هارمن
۱۰	۳/۱۲۸۸(۷)	۶/۹۸۶۳(۱۰)	۰/۰۲۶۸(۱۰)	۸/۷۴۶۱(۹)	۱/۶۳۸۱(۱۰)	۸/۱۱۰۸(۹)	۰/۴۵۳۷(۱۰)	مدل ماکسیمم آتروبی با ۴ گشتاور
۹	۲/۰۴۶۷(۷)	۵/۷۱۰۵(۷)	۰/۰۱۹۶(۹)	۸/۵۵۱۹(۱۰)	۱/۴۰۲۹(۹)	۸/۳۶۴۷(۱۰)	۳/۳۸۱۰(۷)	مدل ماکسیمم آتروبی با ۵ گشتاور
۶	۵/۰۹۷۸(۱۰)	۵/۰۹۷۶(۵)	۰/۰۱۴۵(۵)	۷/۰۸۴۷(۷)	۱/۲۰۵۸(۵)	۷/۵۸۷۷(۷)	۰/۳۴۶۳(۵)	مدل ماکسیمم آتروبی با ۶ گشتاور
۱	۰/۰۹۸۷(۲)	۴/۳۴۰۰(۱)	۰/۰۰۵۷(۱)	۶/۴۵۹۵(۳)	۰/۸۶۷۷(۱)	۶/۹۲۸۱(۳)	۰/۹۲۸۱(۱)	مدل APC
۱۲	۱۸/۳۳۴(۱۲)	۱۸/۳۳۴(۱۲)	۰/۰۶۲۴(۱۲)	۴۲۰/۵۴۷۱(۱۲)	۲/۴۹۹۴(۱۲)	۵۲/۶۱۱(۱۲)	۱/۰۱۱۹(۱۲)	مدل CBD
۱۱	۷/۱۲۴(۱۱)	۱۰/۷۷۶۳(۱۱)	۰/۰۵۲۷(۱۱)	۲۱۲۰/۴۹(۱۱)	۲/۲۹۷۲(۱۱)	۱۹/۴۹۸۶(۱۱)	۰/۶۵۷۱(۱۱)	مدل GMI
۲	۴/۳۴۰۰(۹)	۴/۵۸۵۱(۲)	۰/۰۰۸۵(۳)	۶/۶۶۲۲(۱)	۰/۹۲۳۲(۳)	۶/۶۵۷۵(۱)	۰/۳۰۵۴(۲)	مدل PLAT

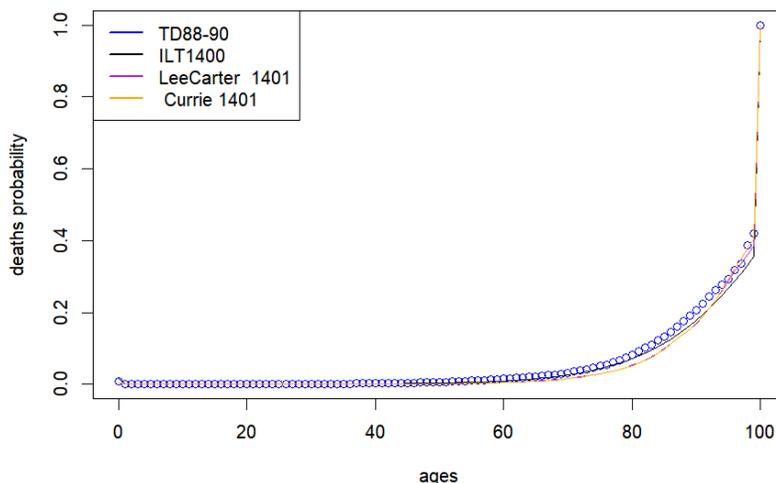
در ادامه همچنین به مقایسه احتمال مرگ و میر جداول (۸۸-۹۰ TD فرانسه) و (۱۴۰۰ ILT ایران) که در شرکت‌های بیمه مورد استفاده قرار می‌گیرند با نرخ مرگ و میر پیش‌بینی شده در این مقاله برای سال‌های ۱۴۰۰ و ۱۴۰۱ پرداخته شده است.



شکل ۳. مقایسه نرخ مرگ و میر جداول عمر مختلف با نرخ مرگ و میر پیش‌بینی شده در سال ۱۴۰۰ براساس بهترین مدل انتخابی

رنگ آبی مربوط به جدول TD88-90، مشکی مربوط به جدول ILT1400، قرمز مربوط به پیش‌بینی مدل انتخابی لی کارتر برای جمعیت مردان در سال ۱۴۰۰ و رنگ سبز مربوط به پیش‌بینی مدل انتخابی کوری برای جمعیت زنان در سال ۱۴۰۰ است.

همان‌طور که مشخص است جدول TD88-90 نرخ مرگ و میر بیشتری را نسبت به بقیه جداول گزارش می‌کند. همچنین نرخ مرگ و میر گزارش شده از مدل لی کارتر نسبت به جدول ILT1400 مقدار کمتری را به جز در دو بازه سنی (۱۶-۲۶) و (۹۰-۹۹) سال نشان می‌دهد. همچنین با مقایسه نرخ مرگ و میر به دست آمده از پیش‌بینی مدل کوری و جدول ILT1400 در می‌یابیم که جدول مرگ و میر حاصل از بهترین مدل انتخابی برای زنان، مدل کوری در سال ۱۴۰۰ همواره عدد کوچک‌تری را نسبت به جدول کنونی ایران به جز در بازه ۷ ساله (۹۲-۹۹) سال گزارش می‌کند.



شکل ۴. مقایسه نرخ مرگ‌ومیر جداول موجود در ایران با نرخ مرگ‌ومیر پیش‌بینی شده در سال ۱۴۰۱ براساس بهترین مدل انتخابی

رنگ آبی مربوط به جدول TD88-90، مشکی مربوط به جدول ILT1400، نارنجی مربوط به پیش‌بینی بهترین مدل انتخابی لی‌کارتی برای جمعیت مردان در سال ۱۴۰۱ و بنفش مربوط به پیش‌بینی بهترین مدل انتخابی کوری برای جمعیت زنان در سال ۱۴۰۱ است. با مقایسه نرخ‌های مرگ‌ومیر پیش‌بینی شده در سال ۱۴۰۱ براساس بهترین مدل انتخابی با جدول مرگ‌ومیر کنونی (ILT1400) در می‌یابیم که نتایج پیش‌بینی شده در تمامی بازه‌ها به جز (۹۲-۹۹ سال) برای مدل انتخابی کوری، برای جمعیت زنان در سال ۱۴۰۱ و بازه سنی (۷۹-۹۹ سال) برای مدل لی‌کارتی، برای جمعیت مردان در سال ۱۴۰۱ مقادیر کوچکتری را گزارش می‌کنند. اگرچه جداول مرگ‌ومیر در تمام دنیا برای هرجنسیت به صورت جداگانه تدوین می‌شود اما هیچ یک از جداول موجود در ایران علی‌رغم مزایای زیادی که به همراه دارند به تفکیک جنسیت ارائه نشده است. انتظار می‌رود با استفاده از جداول معرفی شده در این مقاله که براساس عملکرد بهترین مدل تصادفی مرگ‌ومیر در برآزش و پیش‌بینی نرخ فوت به تفکیک جنسیت انتخاب شده‌اند، محاسبات دقیق‌تری در بیمه‌عمر و مستمری به عمل آید.

جدول ۴: جدول مرگ و میر سال ۱۴۰۰، برای جمعیت زنان ایران براساس پیش‌بینی مدل کوری

x	m_x	q_x	l_x	d_x	L_x	T_x	e_x
1	0.00175	0.00175	99,443	173	99,356	7,949,312	79
2	0.00083	0.00083	99,269	81	99,228	7,849,955	79
3	0.00049	0.00049	99,187	48	99,163	7,750,726	78
4	0.00034	0.00034	99,139	33	99,122	7,651,563	77
5	0.00026	0.00026	99,105	25	99,092	7,552,440	76
6	0.00025	0.00025	99,079	25	99,067	7,453,348	75
7	0.00027	0.00027	99,054	26	99,041	7,354,281	74
8	0.00026	0.00026	99,027	25	99,014	7,255,240	73
9	0.00029	0.00029	99,001	28	98,987	7,156,225	72
10	0.00030	0.00030	98,973	29	98,958	7,057,237	71
11	0.00032	0.00032	98,943	31	98,928	6,958,279	70
12	0.00034	0.00034	98,912	33	98,895	6,859,350	69
13	0.00036	0.00036	98,879	35	98,861	6,760,454	68
14	0.00039	0.00039	98,843	38	98,824	6,661,593	67
15	0.00042	0.00042	98,805	41	98,784	6,562,768	66
16	0.00046	0.00046	98,763	45	98,741	6,463,984	65
17	0.00049	0.00049	98,718	48	98,694	6,365,242	64
18	0.00049	0.00049	98,670	48	98,646	6,266,548	63
19	0.00049	0.00049	98,621	48	98,597	6,167,901	62
20	0.00048	0.00048	98,573	47	98,549	6,069,304	61
21	0.00047	0.00047	98,525	46	98,502	5,970,754	60
22	0.00045	0.00045	98,479	44	98,457	5,872,251	59
23	0.00044	0.00044	98,434	43	98,413	5,773,794	58
24	0.00043	0.00043	98,391	42	98,370	5,675,381	57
25	0.00042	0.00042	98,348	41	98,328	5,577,011	56
26	0.00041	0.00041	98,307	40	98,287	5,478,683	55
27	0.00041	0.00041	98,267	40	98,246	5,380,395	54
28	0.00041	0.00041	98,226	40	98,206	5,282,148	53
29	0.00042	0.00042	98,186	40	98,166	5,183,942	52
30	0.00043	0.00043	98,145	42	98,124	5,085,776	51
31	0.00045	0.00045	98,103	44	98,081	4,987,651	50
32	0.00048	0.00048	98,058	47	98,035	4,889,570	49
33	0.00051	0.00051	98,011	50	97,986	4,791,535	48
34	0.00055	0.00055	97,961	54	97,934	4,693,548	47

x	m_x	q_x	L_x	d_x	L_x	T_x	e_x
35	0.00060	0.00060	97,907	58	97,877	4,595,614	46
36	0.00065	0.00065	97,848	63	97,817	4,497,736	45
37	0.00070	0.00069	97,785	67	97,751	4,399,918	44
38	0.00075	0.00075	97,717	73	97,680	4,302,167	44
39	0.00081	0.00080	97,644	78	97,605	4,204,486	43
40	0.00087	0.00087	97,565	84	97,523	4,106,881	42
41	0.00093	0.00093	97,481	90	97,436	4,009,357	41
42	0.00100	0.00100	97,390	96	97,342	3,911,921	40
43	0.00107	0.00107	97,293	103	97,242	3,814,579	39
44	0.00114	0.00114	97,190	111	97,134	3,717,337	38
45	0.00123	0.00123	97,079	119	97,019	3,620,202	37
46	0.00132	0.00132	96,960	128	96,896	3,523,182	36
47	0.00143	0.00143	96,832	138	96,762	3,426,286	35
48	0.00156	0.00155	96,693	150	96,618	3,329,523	34
49	0.00170	0.00169	96,543	163	96,461	3,232,905	33
50	0.00185	0.00185	96,379	178	96,290	3,136,443	32
51	0.00202	0.00201	96,201	193	96,104	3,040,152	31
52	0.00220	0.00220	96,008	210	95,902	2,944,047	30
53	0.00241	0.00240	95,797	230	95,682	2,848,145	29
54	0.00264	0.00264	95,566	252	95,440	2,752,463	28
55	0.00292	0.00291	95,314	277	95,175	2,657,022	27
56	0.00323	0.00323	95,037	306	94,883	2,561,846	26
57	0.00360	0.00360	94,730	340	94,559	2,466,962	26
58	0.00405	0.00404	94,389	381	94,198	2,372,402	25
59	0.00454	0.00453	94,008	425	93,795	2,278,203	24
60	0.00509	0.00508	93,582	475	93,344	2,184,408	23
61	0.00571	0.00569	93,107	529	92,842	2,091,063	22
62	0.00641	0.00639	92,577	591	92,281	1,998,221	21
63	0.00713	0.00711	91,986	653	91,659	1,905,940	20
64	0.00797	0.00794	91,332	725	90,969	1,814,281	19
65	0.00895	0.00891	90,607	807	90,203	1,723,311	19
66	0.01007	0.01002	89,800	899	89,349	1,633,108	18
67	0.01137	0.01130	88,900	1,004	88,397	1,543,759	17
68	0.01273	0.01265	87,895	1,111	87,338	1,455,361	16
69	0.01431	0.01420	86,783	1,232	86,166	1,368,022	15

انتخاب مدل مرگ و میر تصادفی برای جمعیت زنان و مردان ایرانی ۴۷

x	m_x	q_x	l_x	d_x	L_x	T_x	e_x
70	0.01609	0.01596	85,551	1,365	84,866	1,281,856	14
71	0.01811	0.01795	84,186	1,510	83,428	1,196,990	14
72	0.02036	0.02016	82,675	1,666	81,839	1,113,561	13
73	0.02333	0.02306	81,008	1,868	80,071	1,031,722	12
74	0.02660	0.02625	79,140	2,077	78,097	951,651	12
75	0.03022	0.02977	77,063	2,293	75,910	873,553	11
76	0.03421	0.03363	74,769	2,514	73,505	797,643	10
77	0.03849	0.03775	72,255	2,727	70,882	724,137	10
78	0.04347	0.04254	69,527	2,957	68,037	653,255	9
79	0.04897	0.04779	66,569	3,181	64,965	585,217	8
80	0.05516	0.05367	63,388	3,401	61,671	520,251	8
81	0.06233	0.06042	59,986	3,624	58,155	458,580	7
82	0.07083	0.06838	56,361	3,853	54,412	400,424	7
83	0.08097	0.07778	52,508	4,083	50,438	346,012	6
84	0.09238	0.08824	48,424	4,272	46,254	295,573	6
85	0.10533	0.09997	44,151	4,413	41,905	249,319	5
86	0.12015	0.11321	39,737	4,498	37,443	207,413	5
87	0.13795	0.12886	35,238	4,540	32,916	169,970	4
88	0.15177	0.14081	30,698	4,322	28,482	137,053	4
89	0.16771	0.15440	26,375	4,072	24,282	108,571	4
90	0.18618	0.16987	22,303	3,788	20,349	84,289	3
91	0.20929	0.18884	18,514	3,496	16,705	63,939	3
92	0.23678	0.21084	15,018	3,166	13,372	47,234	3
93	0.26602	0.23357	11,851	2,768	10,406	33,861	2
94	0.30106	0.25996	9,083	2,361	7,843	23,455	2
95	0.33997	0.28821	6,722	1,937	5,698	15,611	2
96	0.37847	0.31509	4,784	1,507	3,983	9,912	2
97	0.40911	0.33576	3,277	1,100	2,689	5,929	1
98	0.45117	0.36312	2,176	790	1,751	3,239	1
99	0.48808	0.38620	1,386	535	1,096	1,487	1
100	0.51046	1.00000	850	850	390	390	0

جدول ۵: جدول مرگ و میر سال ۱۴۰۰، برای جمعیت مردان ایران براساس پیش‌بینی مدل لی-کارتر

x	m_x	q_x	l_x	d_x	L_x	T_x	e_x
1	0.00441	0.00440	100,000	439	99,779	7,712,147	77
2	0.00119	0.00119	99,560	118	99,501	7,612,367	76
3	0.00064	0.00064	99,441	64	99,409	7,512,866	75
4	0.00046	0.00046	99,377	45	99,354	7,413,456	74
5	0.00039	0.00039	99,332	38	99,312	7,314,101	73
6	0.00035	0.00035	99,293	34	99,276	7,214,788	72
7	0.00034	0.00034	99,259	33	99,242	7,115,512	71
8	0.00037	0.00037	99,225	36	99,206	7,016,269	70
9	0.00035	0.00035	99,188	34	99,171	6,917,063	69
10	0.00034	0.00034	99,154	33	99,137	6,817,891	68
11	0.00034	0.00034	99,120	33	99,103	6,718,754	67
12	0.00036	0.00036	99,086	35	99,068	6,619,650	66
13	0.00039	0.00039	99,051	38	99,031	6,520,582	65
14	0.00049	0.00049	99,012	48	98,988	6,421,550	64
15	0.00060	0.00060	98,964	59	98,934	6,322,561	63
16	0.00072	0.00072	98,905	71	98,869	6,223,626	62
17	0.00085	0.00085	98,834	83	98,792	6,124,757	61
18	0.00098	0.00098	98,750	97	98,701	6,025,965	61
19	0.00104	0.00104	98,653	102	98,601	5,927,263	60
20	0.00109	0.00109	98,550	107	98,496	5,828,661	59
21	0.00112	0.00112	98,443	110	98,388	5,730,164	58
22	0.00114	0.00114	98,332	112	98,276	5,631,776	57
23	0.00115	0.00115	98,220	112	98,164	5,533,499	56
24	0.00115	0.00115	98,107	112	98,051	5,435,335	55
25	0.00114	0.00114	97,995	111	97,939	5,337,284	54
26	0.00113	0.00113	97,883	110	97,827	5,239,345	53
27	0.00112	0.00112	97,772	109	97,717	5,141,517	52
28	0.00111	0.00111	97,662	108	97,608	5,043,799	51
29	0.00109	0.00109	97,554	106	97,501	4,946,190	50
30	0.00108	0.00108	97,448	105	97,395	4,848,689	49
31	0.00107	0.00107	97,343	104	97,290	4,751,293	48
32	0.00107	0.00107	97,238	103	97,187	4,654,002	47
33	0.00107	0.00107	97,135	103	97,083	4,556,815	46
34	0.00111	0.00111	97,031	107	96,978	4,459,732	45

x	m_x	q_x	l_x	d_x	L_x	T_x	e_x
35	0.00116	0.00115	96,924	111	96,868	4,362,754	45
36	0.00121	0.00121	96,812	117	96,753	4,265,885	44
37	0.00129	0.00129	96,695	124	96,632	4,169,132	43
38	0.00138	0.00138	96,570	132	96,504	4,072,499	42
39	0.00144	0.00144	96,437	138	96,368	3,975,994	41
40	0.00152	0.00152	96,298	146	96,225	3,879,626	40
41	0.00161	0.00161	96,152	154	96,075	3,783,400	39
42	0.00171	0.00171	95,998	164	95,916	3,687,324	38
43	0.00183	0.00183	95,834	175	95,746	3,591,408	37
44	0.00198	0.00197	95,659	188	95,564	3,495,661	36
45	0.00213	0.00213	95,470	203	95,368	3,400,097	35
46	0.00230	0.00230	95,267	218	95,157	3,304,728	34
47	0.00247	0.00247	95,048	234	94,930	3,209,570	33
48	0.00266	0.00265	94,813	251	94,687	3,114,639	32
49	0.00281	0.00281	94,562	265	94,429	3,019,952	31
50	0.00302	0.00301	94,296	284	94,154	2,925,523	31
51	0.00327	0.00327	94,011	307	93,858	2,831,369	30
52	0.00358	0.00358	93,704	335	93,537	2,737,510	29
53	0.00395	0.00395	93,369	368	93,185	2,643,973	28
54	0.00449	0.00448	93,001	417	92,792	2,550,788	27
55	0.00508	0.00506	92,584	468	92,349	2,457,996	26
56	0.00570	0.00568	92,115	523	91,853	2,365,646	25
57	0.00636	0.00634	91,591	580	91,300	2,273,793	24
58	0.00706	0.00704	91,011	641	90,690	2,182,493	24
59	0.00780	0.00777	90,370	703	90,019	2,091,803	23
60	0.00856	0.00853	89,668	765	89,285	2,001,784	22
61	0.00935	0.00931	88,903	828	88,489	1,912,499	22
62	0.01018	0.01013	88,075	892	87,628	1,824,011	21
63	0.01107	0.01101	87,183	960	86,702	1,736,382	20
64	0.01202	0.01195	86,223	103	85,707	1,649,680	19
65	0.01308	0.01299	85,193	111	84,638	1,563,973	18
66	0.01428	0.01418	84,086	119	83,488	1,479,334	18
67	0.01570	0.01557	82,894	129	82,246	1,395,846	17
68	0.01741	0.01726	81,603	141	80,896	13,135,100	16
69	0.01916	0.01898	80,194	152	79,430	1,232,703	15

x	m_x	q_x	l_x	d_x	L_x	T_x	e_x
70	0.02118	0.02096	78,672	165	77,845	1,153,273	15
71	0.02350	0.02323	77,023	179	76,125	1,075,428	14
72	0.02615	0.02581	75,234	194	74,259	999,303	13
74	0.03160	0.03111	71,189	221	70,076	852,809	12
75	0.03468	0.03409	68,975	235	67,792	782,733	11
76	0.03827	0.03755	66,623	2,501	65,364	714,940	10
77	0.04225	0.04137	64,121	2,652	62,786	649,575	10
78	0.04646	0.04539	61,469	2,790	60,063	586,789	9
79	0.05334	0.05194	58,678	3,047	57,141	526,726	8
80	0.06016	0.05839	55,631	3,248	53,990	469,584	8
81	0.06689	0.06470	52,382	3,389	50,669	415,594	7
82	0.07348	0.07085	48,993	3,471	47,236	364,924	7
83	0.07985	0.07674	45,522	3,493	43,752	317,687	6
84	0.08626	0.08264	42,028	3,473	40,267	273,935	6
85	0.09362	0.08937	38,555	3,445	36,805	233,667	6
86	0.10259	0.09751	35,109	3,423	33,368	196,862	5
87	0.11427	0.10798	31,686	3,421	29,942	163,493	5
88	0.13104	0.12282	28,264	3,471	26,491	133,550	4
89	0.15146	0.14055	24,793	3,484	23,007	107,059	4
90	0.17485	0.16042	21,308	3,418	19,549	84,052	3
91	0.20117	0.18223	17,890	3,260	16,205	64,502	3
92	0.22940	0.20499	14,630	2,999	13,073	48,296	3
93	0.25715	0.22674	11,631	2,637	10,256	35,223	3
94	0.28283	0.24635	8,993	2,215	7,833	24,967	2
95	0.31435	0.26973	6,778	1,828	5,816	17,133	2
96	0.35316	0.29753	4,949	1,472	4,170	11,316	2
97	0.38559	0.31995	3,477	1,112	2,885	7,146	2
98	0.40626	0.33386	2,364	789	1,943	4,261	1
99	0.46589	0.37242	1,575	586	1,259	2,318	1
100	0.49111	0.38805	988	383	781	1,058	1

بحث و نتیجه گیری

اگرچه افزایش طول عمر انسان بدون شک یکی از برجسته‌ترین دستاوردهای جوامع مدرن است با این حال بهبود در بقا و همچنین کاهش باروری باعث ایجاد روند سالخوردگی در جمعیت سراسر جهان شده است. در نتیجه، بخش‌های دولتی و خصوصی برای ارائه محصولات بازنشستگی و مراقبت‌های بهداشتی سالمندان با چالش‌های زیادی مواجه شدند. عواقب این دو معضل به‌گونه‌ای است که صنعت مالی برای اشاره به این موضوع اصطلاح ریسک طول عمر را به کار برده است. به این معنا که اگر بازنشستگان (به‌طور متوسط) بیشتر از حد انتظار زندگی کنند. مؤسسه با خروج پولی بیشتر از ذخایر برنامه‌ریزی شده خود مواجه خواهد شد. تفاوت‌های کوچک بین طول عمر محقق شده و پیش‌بینی شده بازنشستگان، در بازار مالی ضررهای بسیار بزرگی را به وجود می‌آورد. به این ترتیب، نیاز به مدل‌های نوآورانه‌ای که بتواند سیر مرگ و میر آینده را با دقت بیشتری نسبت به رویکردهای قبلی پیش‌بینی کند، مشهود است. به‌منظور پیش‌بینی مناسب نرخ فوت جمعیت و با توجه به اینکه مرگ‌ومیر یک فرآیند پویا است که با گذشت زمان دستخوش تغییرات می‌شود، در این مقاله علاوه بر مدل‌های تصادفی کلاسیک از خانواده مدل‌های مربوط به توزیع سن در زمان فوت نیز برای پیش‌بینی نرخ فوت استفاده شده است.

در این پژوهش به انتخاب بهترین مدل مرگ‌ومیر تصادفی برای پیش‌بینی نرخ فوت ایران به تفکیک جنسیت و براساس داده‌های واقعی پرداخته شده است. در واقع برای انجام این کار از ۹ مدل تصادفی مبتنی بر پیش‌بینی نرخ فوت/نرخ مرگ‌ومیر و مدل ماکسیمم آنتروپی مبتنی بر پیش‌بینی فراوانی فوت استفاده شده است. برای ارزیابی قدرت پیش‌گویی هر یک مدل‌ها از ۷ معیار ارزیابی مدل استفاده شده است. براساس بازه مورد نظر (۰-۱۰۰) سال مشخص شد که از بین ۱۰ مدل مورد بررسی، مدل تصادفی لی‌کارتر برای جمعیت مردان و مدل تصادفی کوری برای جمعیت زنان بهترین عملکرد را دارند، یعنی مدل دو عاملی و سه عاملی مبتنی بر خانواده مدل‌های لی‌کارتر نتایج دقیق‌تری را ارائه کرده‌اند و بعد از آن‌ها با اندکی اختلاف به ترتیب برای

جمعیت مردان و زنان ایران مدل هیندمن اولاً و مدل پلینت رتبه دوم را اخذ می‌کنند. در ادامه جداول عمر ایران باتوجه برای سال ۱۴۰۰ باتوجه به عملکرد پیش‌بینی مدل کوری و مدل لی‌کارت به ترتیب برای جمعیت زنان و مردان ایران ارائه شد. همان‌طور که از جدول ۴ و ۵ مشخص است امید زندگی زنان به‌طور متوسط ۳ سال در سنین کودکی، ۲ سال در سنین میان‌سالی و یکسال در سنین پیری از امید زندگی مردان بیشتر است، لذا الگوی کلی مرگ‌ومیر زنان با مردان متفاوت است. بنابراین بدیهی است که یک مدل برای هر دو جنسیت، عملکرد یکسانی نداشته باشد. همچنین در ادامه مقایسه نرخ مرگ‌ومیر پیش‌بینی شده در سال‌های ۱۴۰۰ و ۱۴۰۱ با نرخ مرگ‌ومیر جداول عمر موجود در ایران انجام شد. از نتایج مقایسه این‌گونه دریافت شد که در اکثر بازه‌های سنی نرخ مرگ‌ومیر پیش‌بینی شده (احتمال فوت) با توجه به عملکرد بهترین مدل، اعداد کوچکتری را نسبت به جداول عمر موجود در ایران گزارش می‌کنند. این موضوع برای صنعت بیمه و صندوق‌های بازنشستگی از اهمیت خاصی برخوردار است. زیرا هرچه احتمال فرد بیمه‌شده کوچکتر باشد، فرد بیشتر عمر می‌کند و در بیمه‌نامه‌های عمر مختلط و بیمه عمر به شرط فوت می‌توانیم حق بیمه کمتری را از فرد بیمه‌شده دریافت کنیم. کاهش حق بیمه می‌تواند یک عاملی برای افزایش تقاضای این محصول بیمه‌ای باشد. همچنین در صدور بیمه‌نامه‌های عمر به شرط حیات به دلیل کاهش احتمال فوت و طولانی‌تر شدن مدت زمان برقراری بیمه‌نامه حق بیمه بیشتری باید از فرد بیمه‌شده اخذ شود.

در این پژوهش، به‌منظور مقایسه دقت پیش‌بینی مدل‌های مرگ‌ومیر تصادفی، از شاخص نرخ فوت خام استفاده شده است. مشخص است که با تغییر شاخص مرگ‌ومیر و همچنین معیارهای ارزیابی مدل، نتایج متفاوتی حاصل می‌شود. سن معین در هنگام مرگ یکی از شاخص‌هایی است که اخیراً شناسایی شده است. سن معین در هنگام مرگ را به‌عنوان اصلی‌ترین و طبیعی‌ترین ویژگی طول عمر انسان معرفی می‌کنند و برخلاف شاخص امید زندگی در بدو تولد، تنها تحت تأثیر مرگ‌ومیر بزرگسالان است و در نتیجه نسبت به تغییراتی که در میان جمعیت سالمند رخ می‌دهد حساس‌تر است. در کشورهای با مرگ‌ومیر پایین که بیشتر فوت‌ها

در سنین بالا اتفاق می‌افتد، سن معین در هنگام مرگ به یک شاخص قوی تبدیل شده است. در نظر گرفتن این شاخص و یا دیگر شاخص‌های جداول مرگ‌ومیر (امید زندگی در بدو تولد، نرخ مرگ‌ومیر، تعداد بازماندگان در سن x و ...) می‌تواند برای پژوهش‌های آتی مفید باشد. همچنین به دلیل تاثیر کیفیت زندگی بر طول عمر انسان می‌توان ضرایب تعدیلی را برای جدول عمر را به تفکیک محل سکونت (شهر و روستا) و استان محل سکونت در نظر گرفت.

تشکر و قدردانی

این مقاله برگرفته از پایان نامه کارشناسی ارشد نویسنده دوم مقاله در رشته بیم‌سنجی است که در دانشکده علوم ریاضی دانشگاه شهید بهشتی به انجام رسیده است. نویسندگان از حمایت مالی پارک فناوری اطلاعات و ارتباطات وزارت علوم، تحقیقات و فناوری در قالب کد اعتباری ۰۰۰۵۵۳ - ۰۱ - ۰۰ - ۰۲ قدردانی می‌نمایند.

منابع

ذکایی، محمد، و آل‌حسینی، فاطمه‌السادات. (۱۳۹۳). مقایسه روش تحلیل مجموعه مقادیر تکین در پیش‌بینی نرخ مرگ‌ومیر با روش‌هایی از خانواده لی-کارتز، پژوهشنامه بیمه، (۱۳)، ۲۸-۱۳.

<https://doi.org/10.22056/ijir.2014.01.02>

شعاعی، شیرین و قلی‌کشمیزی، محمدمهدی (۱۴۰۱). تحلیل خانواده مدل‌های مرگ‌ومیر تصادفی برپایه مدل لی-کارتز در پیش‌بینی نرخ مرگ‌ومیر در بیمه‌های زندگی و درمان، بیمه سلامت ایران، (۱۴)، ۶۸-۷۹.

<http://journal.ihio.gov.ir/article-1-187-fa.html>

شورای عالی بیمه، (۱۳۹۰). آئین نامه شماره ۶۸ درخصوص بیمه‌های زندگی و مستمری. تهران.

عزیزی‌شاکر، عاطفه (۱۳۹۸). ساخت جدول‌های عمر استان‌های ایران در دوره ۱۳۹۰-۱۳۹۵ و انطباق آن با جدول‌های عمر مدل، پایان‌نامه کارشناسی ارشد جمعیت‌شناسی، دانشگاه بوعلی سینا، دانشکده علوم اقتصادی

و اجتماعی، گروه علوم اجتماعی

کمیجانی، اکبر، کوششی، مجید و نیاکان، لیلی (۱۳۹۲)، برآورد و پیش‌بین نرخ مرگ‌ومیر در ایران با استفاده از مدل لی کارتر پژوهشنامه بیمه، ۲(۴)، ۳۱۰-۲۹۵. <https://doi.org/10.22056/ijir.2013.04.01>

Azizi-Shaker, A. (2019). Constructing the life tables of Iran's provinces in the period 2015-2016 and its compliance with the model life tables, Master Thesis in Demography, Bu-Ali Sina University, Faculty of Economic and Social Sciences, Department of Social Sciences. [In Persian].

Basellini, U., & Camarda, C. G. (2019). Modelling and forecasting adult age-at-death distributions. *Population Studies*, 73(1), 119-138. <https://doi.org/10.1080/00324728.2018.1545918>

Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, 3(1-2), 3-43. <https://doi.org/10.1017/s1748499500000440>

Cairns, A. J. G., Blake, D. P., Dowd, K., Coughlan, G., & Epstein, D. (.), A quantitative comparison of stochastic mortality models using data from England & wales and the United States. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1340389>

Cairns, A. J. G., Blake, D., & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *The Journal of Risk and Insurance*, 73(4), 687-718. <https://doi.org/10.1111/j.1539-6975.2006.00195.x>

Canudas-Romo, V. (2010). Three measures of longevity: time trends and record values. *Demography*, 47(2), 299-312. <https://doi.org/10.1353/dem.0.0098>

Chen, Z., Yang, Y. (2004). Assessing forecast accuracy measures. *Preprint Series, 2010* :2004-10. <https://api.semanticscholar.org/CorpusID:15173693>

Currie, I. D. (2006), *Smoothing and forecasting mortality rates with P-splines*, Talk given at the Institute of Actuaries.

De Moivre, A. (1725). *Annuities on Lives: Or, the Valuation of Annuities Upon Any Number of Lives as also of Reversions*, William Person, London.

Dellaportas, P., Smith, A. F. M., & Stavropoulos, P. (2001). Bayesian analysis of mortality data. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)*, 164(2), 275-291. <https://doi.org/10.1111/1467-985x.00202>

- Graunt, J. (1662). *Natural and Political Observations Made upon the Bills of Mortality*, John Hopkins Press.
- Halley, E. (1693). VI. An estimate of the degrees of the mortality of mankind; drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. *Philosophical Transactions of the Royal Society of London*, 17(196), 596–610. <https://doi.org/10.1098/rstl.1693.0007>
- Hyndman, R.J., & Shahid Ullah, M. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics & Data Analysis*, 51(10), 4942–4956. <https://doi.org/10.1016/j.csda.2006.07.028>
- Iran Supreme Council of Insurance (2011), Regulation No. 68 of the Supreme Council of Insurance: Life and Pension Insurances. Tehran [In Persian].
- Komijani, A., Koosheshi, M., & Niakan, L. (2013). Iran mortality rates using Lee-Carter model: Estimation and forecasting. *Iranian Journal of Insurance Research*, 2(4), 295-310. [In Persian]. <https://doi.org/10.22056/ijir.2013.04.01>
- Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association*, 87(419), 659-671. <https://doi.org/10.2307/2290201>
- Levantesi, S., & Pizzorusso, V. (2019). Application of machine learning to mortality modeling and forecasting. *Risks*, 7(1), 26. <https://doi.org/10.3390/risks7010026>
- 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. <https://doi.org/10.1353/dem.2005.0021>
- Makeham, W. M. (1867). On the law of mortality. *Journal of the Institute of Actuaries* (1866), 13(06), 325–358. <https://doi.org/10.1017/s2046166600003238>
- Mazzuco, S., Scarpa, B., & Zanotto, L. (2018). A mortality model based on a mixture distribution function. *Population Studies*, 72(2), 191–200. <https://doi.org/10.1080/00324728.2018.1439519>
- Pascariu, M. D., Lenart, A., & Canudas-Romo, V. (2019). The maximum entropy mortality model: forecasting mortality using statistical moments. *Scandinavian Actuarial Journal*, 2019(8), 661–685. <https://doi.org/10.1080/03461238.2019.1596974>
- Pearson, K. (1897), *The Chances of Death, and Other Studies in Evolution*, London, E. Arnold, <https://doi.org/10.5962/bhl.title.18666>

- Renshaw, A. E., & Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance, Mathematics & Economics*, 38(3), 556-570. <https://doi.org/10.1016/j.insmatheco.2005.12.001>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: The University of Illinois Press. https://pure.mpg.de/rest/items/item_2383164_3/component/file_2383163/content
- Shoaei, S., Gholi Keshmarzi, M.M. (2021) Analysis of the Stochastic Mortality Models Based on Lee-Carter Model in Predicting Mortality Rates in Life and Health Insurance. *Iranian Journal of Health Insurance*, 4(1), 68-79. [In Persian]. <http://journal.ihio.gov.ir/article-1-187-en.html>
- Waldron, H. (2005). Literature review of long-term mortality projections. *Social Security Bulletin*, 66(1), 16-30. <https://www.ssa.gov/policy/docs/ssb/v66n1/v66n1p16.pdf>
- Yang, S. S., Yue, J. C., & Huang, H.-C. (2010). Modeling longevity risks using a principal component approach: A comparison with existing stochastic mortality models. *Insurance, Mathematics & Economics*, 46(1), 254-270. <https://doi.org/10.1016/j.insmatheco.2009.09.013>
- Zokaei, M., & Alehoseini, F. (2013). Comparison of singular value set analysis method in predicting mortality rate with Lee-Carter family methods. *Iranian Journal of Insurance Research*, 3(1), 13-28. [In Persian]. <https://doi.org/10.22056/ijir.2014.01.02>