

## Application of Phase Distributions in the Aging Process of Patients with Skin Cancer

Raoufeh Asghari<sup>1</sup>, Amin Hassan Zadeh<sup>2,\*</sup>

<sup>1</sup> Department of Statistics, Shahid Beheshti University, Tehran, Iran

<sup>2</sup> Department of Actuarial Science, Shahid Beheshti University, Tehran, Iran

\* **Corresponding Author:** Amin Hassan Zadeh, Department of Actuarial Science, Shahid Beheshti University, Tehran, Iran. E-mail: am\_hassanzadeh@sbu.ac.ir

**Received:** 30 Jun 2018

**Revised:** 18 Sep 2018

**Accepted:** 20 Oct 2018

**Online Published:** 20 Nov 2018

### Abstract

**Introduction:** Skin cancer is one of the most common types of cancers in the United States of America and can be prevented in most cases. Skin cancer significantly affects the quality of life of people and can cause disorder or even death. A review of the aging process of cancer patients, its modeling with physiological age and comparison with normal people is conducted in this paper by using phase-type distributions.

**Methods:** In this model, it is assumed that the length of any physiological age follows an exponential distribution in a Markov chain environment. In the continuous-time Markov chain, a state is assumed to represent death, and  $n_x$  transient states, where  $x$  is the patient's age at the time of diagnosis of cancer. Each transient state represents a physiological age and aging is a process of change from a physiological age to the next physiological age to reach the end of the process. There is also an absorbing state that the transition from any state to the absorbing state can take place. In this study, using data of skin cancer patients in the United States, the unknown parameters associated with the aging process were estimated.

**Results:** The study was conducted on patients with melanoma-related cancer in the United States during the years 1973 to 2014, with aged 60-65 years old. The registered number of melanoma cases was 1,882, of which 1,251 were male (66.5%) and 631 were female (33.5%). A table of parameters for estimating survival probability and related charts for the whole population in the age group and gender is presented.

**Conclusions:** The fitting results of data modeling are very satisfying. The physiological age parameters were estimated in general that could be useful in estimating the distribution of the phase-type parameters and to calculate the function and the moments. And also the effect of gender on the survival rate of patients was determined which indicates that the survival of males is higher than that of women. On the other hand, the life expectancy of cancer patients has been compared with the entire population of the United States, which was less it was expected.

**Keywords:** Survival Analysis, Aging Process, Continuous Markov Chain of Time, Phase Type Distribution, Skin Cancer

### Citation:

Asghari R, Hassan Zadeh A. Application of phase distributions in the aging process of patients with skin cancer. Iran J Health Insur. 2018;1(3):97-104.

## کاربرد توزیع فازنوع در فرآیند سالخوردگی بیماران با سرطان پوست

رئوفه اصغری<sup>۱</sup>، امین حسن زاده<sup>۲</sup>

<sup>۱</sup> گروه آمار، دانشگاه شهید بهشتی، تهران، ایران

<sup>۲</sup> گروه بیم‌سنجی، دانشگاه شهید بهشتی، تهران، ایران

\* نویسنده مسئول: امین حسن زاده، گروه بیم‌سنجی، دانشگاه شهید بهشتی، تهران، ایران. ایمیل: am\_hassanzadeh@sbu.ac.ir

تاریخ انتشار آنلاین: ۱۳۹۷/۰۸/۲۹

تاریخ پذیرش: ۱۳۹۷/۰۷/۲۸

تاریخ تصحیح: ۱۳۹۷/۰۶/۲۷

تاریخ دریافت: ۱۳۹۷/۰۴/۰۹

### چکیده

**مقدمه:** سرطان پوست یکی از شایع‌ترین نوع سرطان‌ها در ایالات متحده آمریکا است و در بیشتر موارد قابل پیشگیری می‌باشد. سرطان پوست به طور قابل توجهی در کیفیت زندگی افراد تأثیر می‌گذارد و می‌تواند مبتلایان را دچار اختلال و یا حتی مرگ کند. بررسی فرآیند پیری بیماران سرطانی، مدل‌بندی آن با سن فیزیولوژیکی و مقایسه با افراد نرمال جامعه اهمیت خاصی دارد که در این مطالعه با استفاده از توزیع فازنوع به آن‌ها پرداخته خواهد شد.

**روش بررسی:** در این مدل فرض می‌شود که طول هر سن فیزیولوژیکی از توزیع نمایی در یک محیط زنجیر مارکوف پیروی می‌کند. در زنجیر مارکوف پیوسته زمان، یک وضعیت، وضعیت جذب (مرگ) و  $n_X$  وضعیت گذار می‌باشد که  $X$  سن بیمار در زمان تشخیص سرطان است. هر یک از وضعیت‌های گذار بیانگر یک سن فیزیولوژیکی است و سالخوردگی به صورت یک فرآیند متوالی از تغییر وضعیت‌ها از یک سن فیزیولوژیکی به سن فیزیولوژیکی بعدی است تا فرآیند به انتهای خود برسد. همچنین یک وضعیت جذب وجود دارد که انتقال از هر وضعیت به وضعیت جذب می‌تواند صورت بگیرد که این وضعیت جذب همان مرگ است. در این مطالعه با استفاده از داده‌های مربوط به بیماران سرطان پوست در ایالات متحده آمریکا، پارامترهای نامعلوم که به فرآیند سالخوردگی مرتبط هستند، برآورد شد.

**یافته‌ها:** مطالعه بر روی بیماران دارای سرطان ملانوما مربوط به ایالات متحده آمریکا در طی سال‌های ۱۹۷۳ الی ۲۰۱۴ انجام شده است که افراد با گروه سنی ۶۰ تا ۶۵ ساله مورد بررسی قرار گرفتند. تعداد ثبت شده در ابتلا به سرطان ملانوما ۱۸۸۲ نفر بوده است که ۱،۲۵۱ نفر را مرد (۶۶/۵ درصد) و ۶۳۱ نفر را زن (۳۳/۵ درصد) تشکیل می‌دهند. جدول مربوط به برآورد پارامترها برای تخمین احتمال بقا و نمودارهای مربوطه برای کل افراد در گروه سنی مذکور و همچنین به تفکیک جنسیت ارائه شده است. **نتیجه‌گیری:** نتایج برازش داده شده در مدل‌بندی داده‌ها با استفاده از توزیع فازنوع بسیار رضایت بخش هستند. پارامترهای سن فیزیولوژیکی در حالت کلی برآورد شدند که می‌تواند در برآورد توزیع فازنوع برای محاسبه تابع احتمال، تابع توزیع و گشتاورها و ... مفید باشد و همچنین تأثیر جنسیت در میزان بقای بیماران تعیین شد که نشان می‌دهد بقای مردان سرطانی کمتر از زنان است. از طرفی امید به زندگی افراد سرطانی با کل جمعیت آمریکا مورد مقایسه قرار گرفت که مقدار کمتری را اختیار می‌کند.

**واژگان کلیدی:** آنالیز بقا، فرآیند سالخوردگی، زنجیر مارکوف پیوسته زمان، توزیع فازنوع، سرطان پوست

### مقدمه

تولید نکرده و به رنگ صورتی، قرمز یا بنفش ظاهر می‌شود. معمولاً عوامل ژنتیکی، تابش اشعه ماورای بنفش و تماس‌های محیطی در ایجاد ملانوما دخیل هستند. این سرطان بیشتر در افراد بالغ دیده شده اما ممکن است برای کودکان نیز رخ دهد. در مردان، ملانوما بیشتر در ناحیه تنه (بین شانه‌ها و لگن)، سر یا گردن ایجاد می‌شود، ولی در زنان ملانوما بیشتر روی بازوها و پاها رخ می‌دهد. اگر ملانوما زود تشخیص داده شود و در مراحل اولیه بیماری درمان شود، تقریباً همیشه قابل درمان است. در غیر اینصورت سرطان گسترش و در نواحی دیگر از بدن رشد می‌یابد و در این حالت درمان به سختی صورت می‌گیرد و می‌تواند منجر به مرگ شود. تحلیل بقا یکی از عمده‌ترین تحلیل‌های آماری در حوزه تحقیقات بالینی است و برای تحلیل داده‌هایی کاربرد دارد که زمان تا وقوع رخدادی موردنظر باشد. عبارت داده‌های بقا برای توصیف داده‌هایی اطلاق می‌شود که زمان تا وقوع پیشامد خاصی را اندازه‌گیری می‌کنند. معمولاً این مقادیر نامنفی زمان شکست یک مولفه فیزیکی یا زمان مرگ یک واحد زنده است و همچنین می‌تواند زمان بقا یا زمان

سرطان پوست یکی از شایع‌ترین نوع سرطان در ایالات متحده آمریکا است و در بیشتر موارد قابل پیشگیری می‌باشد. سرطان پوست به طور قابل توجهی در کیفیت زندگی افراد تأثیر می‌گذارد و می‌تواند دچار اختلال و یا حتی مرگ شود. در چهار دهه اخیر تعداد آمریکایی‌هایی که سرطان پوست را در برخی از نقاط بدنشان تجربه کرده‌اند بیشتر از تعداد دیگر سرطان‌هاست و میزان بروز سرطان پوست در سال‌های اخیر همچنان افزایش یافته است. سرطان پوست در واقع بر روی پوست به وجود می‌آید و با گسترش سلول‌های غیر طبیعی می‌تواند به قسمت‌های دیگر بدن آسیب برساند. سه نوع سرطان پوست وجود دارد: سرطان پوست غیر ملانوما (سلول پایه‌ای (BCC)، کارسینوم سلول سنگفرشی (SCC)) و ملانوما. ملانوما بدخیم‌ترین نوع سرطان پوستی است و مسبب اکثر مرگ‌های بر اثر سرطان پوست می‌باشد. ملانوما از ملانوسیت‌های پوست منشأ می‌گیرد. ملانوسیت‌ها، رنگدانه تیره پوست، مو، چشم و خال‌های بدن را تولید می‌کنند. از این رو تومورهای ملانوما اکثراً قهوه‌ای یا سیاه هستند. ولی در موارد معدودی نیز ملانوما رنگدانه

ارزش‌های فعلی بیمسجی را برای محاسبه مستمری و بیمه عمر سالیانه مورد استفاده قرار دادند [۹]. حسن‌زاده و همکارانش برای مدل‌بندی بیمه از کارافتادگی از توزیع فانونوع استفاده کردند. این مدل، فرآیند سالخوردگی را به صورت عبور از تعدادی فاز که با کاهش حیات همراه است نشان می‌دهد. زمانیکه فردی از کارافتاده می‌شود، از چندین فاز عبور می‌کند که مدت از کارافتادگی را نشان می‌دهد [۱۰]. یک فرآیند مارکوف پیوسته زمان را در نظر بگیرید. در این مطالعه مدلی را بر اساس این فرآیند معرفی می‌شود که دارای یک وضعیت جذب (مرگ) می‌باشد و فرآیند سالخوردگی را برای بیماران توصیف می‌کند. سالخوردگی فرآیندی است که فرد پیرتر می‌شود. این فرآیند یک فرآیند پیوسته، پیچیده و پویا است که با تولد شروع شده و با مرگ به پایان می‌رسد. هر کدام از وضعیت‌ها در فرآیند مارکوف یک فاز را نشان می‌دهند که به هر فاز یک سن فیزیولوژیک گویند. نوع مدلی که در این مطالعه بیان شده است به ما اجازه می‌دهد تا از یک مدل نیمه مارکوفی استفاده کنیم. در فرآیند نیم مارکوف علاوه بر تغییر وضعیت‌های بین فازها، مدت زمانی که در هر وضعیت سپری می‌شود و زمان‌های بین تغییر وضعیت‌ها هم مورد نظر قرار می‌گیرند که این زمان‌ها خود یک متغیر تصادفی هستند که در فرآیند مارکوف چنین نیست. بنابراین تحت فرضیانی بیان شده، زمانی که هر فرد در وضعیت سلامتی و بیماری قرار دارد از توزیع فانونوع پیروی می‌کند. خانواده توزیع فانونوع در مجموعه تمام توزیع‌هایی که مقدار حقیقی و نامنفی را اختیار می‌کنند چگال است، و می‌توان از آن برای تقریب هر توزیع با مقدار مثبت استفاده کرد [۱۱]. از طرفی برای توزیع فانونوع فرم بسته برای توابع توزیع و چگالی وجود دارد و علاوه بر آن می‌توان مقدار میانگین و تمام گشتاورهای غیرمرکزی را با گرفتن مشتق از تبدیل لاپلاس-استیلتس به دست آورد. از آنجایی که توزیع فانونوع بر اساس فرآیند مارکوف ذکر شده تعریف شده است، پارامترهای (نماینده) آن بردار احتمال اولیه و ماتریس زیرشدت می‌باشند. بنابراین نماینده‌های یک توزیع فاز نوع ممکن است قابل تفسیر باشد و از این ویژگی توزیع فانونوع برای این مدل استفاده شده است.

### روش بررسی

سالخوردگی یک پدیده‌ای طبیعی است که به وسیله تغییرات بیولوژی، فیزیولوژی، بیوشیمی و آناتومی در سلول‌های بدن ایجاد می‌شود. از لحاظ تئوری، سالخوردگی از زمان تولد شروع شده و در سراسر طول زندگی ادامه می‌یابد و به مرور زمان این تغییرات بر عملکرد سلول‌ها اثر می‌گذارد. تغییرات فیزیولوژی در انسان غالباً با تغییرات روانی، رفتاری، اجتماعی و اقتصادی همراه است. این فرآیند از لحاظ ژنتیکی تعیین شده است و عوامل محیطی بر آن تأثیر دارند. ابعاد زیستی-فیزیولوژیکی سالخوردگی شامل هر دو جنبه عوامل زمینه ساز سالخوردگی و وضعیت سلامت عمومی بدن است. سالخوردگی اغلب با افزایش احتمال مرگ و میر همراه است اما لزوماً علت مرگ نیست. از آنجاییکه با افزایش سن احتمال مرگ به سرعت افزایش می‌یابد، پس قطعاً تغییراتی در بدن رخ می‌دهد که فرد را نسبت به بیماری‌ها آسیب پذیر می‌سازد و عملکرد بسیاری از ارگان‌ها را کاهش می‌دهد. منظور از سن همان سن شناسنامه‌ای یا تقویمی است و تعداد سال‌هایی محسوب می‌شود که یک فرد زندگی می‌کند و منظور از سال همان

وقوع رخدادی باشد. روش‌های مختلفی تحلیل بقای داده‌ها تا وقوع یک پیشامد وجود دارد. روش‌های ناپارامتری که شامل جدول عمر و روش کاپلان مایر می‌باشد که توصیف زمان بقاء بدون در نظر گرفتن متغیرهای کمکی است. روش‌های پارامتری که هدف، بررسی اثرات متغیرهای کمکی بر روی زمان بقاء است که به دو دسته اصلی مدل رگرسیون خطرات متناسب و مدل شتابدار زمان شکست تقسیم می‌شوند. در مدل خطرات متناسب اثر متغیرهای کمکی بر روی تابع مخاطره به دست می‌آید. در این حالت اگر مخاطره پایه به صورت پارامتری در نظر گرفته شود یکی از مدل‌های نمایی، وایبل و گومپرتز به دست می‌آید و در بعضی مواقع ممکن است توزیعی برای مخاطره پایه نتوان در نظر گرفت و توزیع آن نامعلوم باشد، در این صورت این مدل یک مدل نیمه پارامتری است که به مدل کاکس معروف است. در مدل رگرسیونی شتابدار، زمان شکست اثر متغیرهای کمکی بر روی لگاریتم زمان بقا به دست می‌آید. در این حالت مانند مدل رگرسیون معمولی، توزیع باقی مانده اهمیت به سزایی دارد و نوع مدل را مشخص می‌کند. مدل‌های به دست آمده در این حالت بصورت نمایی، وایبل، لگ نرمال، لگ لجستیک و گامای تعمیم یافته است [۱]. با توجه به روش‌های تحلیل بقا، مقالات مختلفی در خصوص سرطان‌ها و بقای بیماران سرطانی کار شده است. اکرام و همکارانش در مطالعه خود از روش مدل‌بندی ناپارامتری و پارامتری برای آنالیز بقای بیماران سرطانی استفاده کردند و تابع بقا را با استفاده از روش کاپلان مایر و مدل وایبل برآورد کردند و نیز یافته‌ها به تفکیک جنسیت مورد بررسی قرار گرفت [۲]. همچنین رویستون استفاده از توزیع لگ نرمال در آنالیز زمان‌های بقای بیماران سرطانی سینه و تخمدان را نشان می‌دهد و بیان می‌کند که این مدل یک مدل مفید برای آنالیز داده‌های بقای سانسور شده می‌باشد [۳]. هوی از مدل وایبل برای آنالیز عوامل پیش‌بینی‌کننده بیماری در بیماران با سرطان معده استفاده می‌کند و این مدل را با مدل کاکس مقایسه می‌کند که با توجه به نتایج به دست آمده، مدل وایبل با ارزش تر از مدل کاکس است [۴]. سیمسک از روش مدل خطرات متناسب کاکس برای مدل‌بندی بقای پنج ساله بیماران سرطانی سینه با در نظر گرفتن متغیر کمکی سن، نوع درمان و مرحله تشخیص بیماری استفاده می‌کند [۵]. وان و همکارانش از روش کاپلان مایر برای محاسبه نرخ‌های بقا در یک نوع سرطان خاص با در نظر گرفتن سن، جنسیت، نژاد و غیره از سال ۱۹۹۰ تا ۲۰۰۹ و از مدل‌های خطرات متناسب کاکس برای محاسبه نسبت‌های خطر برای بهبود در بقا و رهایی از سرطان استفاده کرده است [۶].

مدل در نظر گرفته شده در این مقاله یک مدل پارامتریک است که فرض بر این است که زمان تا مرگ از توزیعی به نام توزیع فاز نوع (PH) پیروی می‌کند. اخیراً از توزیع فانونوع در آنالیز بقا نیز استفاده شده است. آلن از توزیع فانونوع برای مدل کردن فاصله زمانی بین اولین و دومین تولد فرزند در خانواده کمک گرفت [۷]. توزیع فانونوع کاکسی یک حالت خاصی از توزیع فانونوع است که می‌تواند برای بیان زمان‌های بقا استفاده شود و در آن فازهایی تعریف می‌شود که پیشرفت یک فرد را تا ترک سیستم نشان می‌دهد. به عنوان مثال بستری شدن در بیمارستان یا زمان تا مرگ که مارشال و زنگا از مدل کاکسی برای بیماران که در بیمارستان بستری شده‌اند استفاده کردند [۸]. لین و لیو از توزیع فانونوع برای آنالیز مرگ و میر استفاده کردند و احتمالات شرطی بقا و

اگر متغیر  $Y$  را به صورت  $Y = \inf\{t; J(t) \in 0\}$  تعریف کنیم، آنگاه می‌توان گفت که  $Y$  یک متغیر تصادفی فازنوع با پارامتر  $(\alpha, T)$  است. برای یک متغیر تصادفی پیوسته  $Y$ ، پیشامد  $Y > y$  نشاندهنده آن است که فرآیند از هر وضعیت گذرا شروع شده و تا زمان  $y$  به وضعیت جاذب نرسیده باشد. بنابراین تابع توزیع  $Y$  را می‌توان به صورت زیر به دست آورد:

$$\begin{aligned} F(y) &= 1 - P(Y > y) = P(J(y) = 0) \\ &= \sum_{j=0}^n P(J(y) = 0 | J(0) = j) \\ &= P(J(y) = 0 | J(0) = 0)P(J(0) = 0) \\ &+ \sum_{j=1}^n P(J(y) = 0 | J(0) = j)P(J(0) = j) \\ &= \alpha[1 - e^{yT}] = 1 - \alpha e^{yT} \quad y \geq 0 \end{aligned}$$

با گرفتن مشتق از تابع توزیع و در نظر گرفتن  $T1 = -t_0$  تابع چگالی احتمال  $Y$  برابر است با:

$$f_Y(y) = \alpha e^{yT} t_0, \quad y \geq 0$$

و همچنین تبدیل لاپلاس و گشتاور مرتبه  $k$ -ام آن برابر است با

$$\Phi(s) = \alpha(sI - T)^{-1} t_0$$

$$E(Y^k) = k! \alpha (-T^{-1})^k t_0, \quad k = 0, 1, 2, \dots$$

برای به دست آوردن اطلاعات کامل در خصوص توزیع فازنوع به [۶] مراجعه کنید.

هدف ما در این مطالعه مدل‌بندی فرآیند سالخوردگی براساس سن فیزیولوژیک برای بیماران سرطانی است. سن فیزیولوژیک همانطور که به آن اشاره شد به عنوان یک توصیف ذهنی از تجربه یک شخص با استفاده از مشخصه‌های غیر فیزیکی تعریف شده است و زمانیکه وضعیت فیزیکی فردی را با فرد دیگر با همان سن شناسنامه‌ای مقایسه می‌شود، سن فیزیولوژیک یک سن نسبی است و می‌توان گفت که میزان خطر ابتلا به بیماری بیشتر با سن فیزیولوژیک فرد در ارتباط است تا سن شناسنامه‌ای‌اش. سن فیزیولوژیک توسط بسیاری از روانشناسان در قالب نظریه‌های مرحله تعریف شده‌اند که این مراحل، مراحل تکوینی هستند که هر شخص از آن عبور می‌کند. سالخوردگی بر تمامی فرآیندهای فیزیولوژیکی تأثیر می‌گذارد و یک روندی است که در بدو تولد شروع و تا پایان زندگی ادامه دارد.

روش بررسی در این مطالعه برای مدل‌بندی فرآیند سالخوردگی، یک روش بهبودیافته از روشی است که توسط لین و لیو [۹] ارائه شده است. همانطور که در بخش قبل اشاره شد، فرآیند مارکوف پیوسته زمان با وضعیت متناهی به تعداد  $n_x + 1$  وضعیت ( $x$  سن شناسنامه‌ای فرد در زمان تشخیص بیماری است) که  $n_x$  وضعیت گذرا و یک وضعیت جاذب است که مرگ نامیده می‌شود. هر یک از وضعیت‌های گذرا نمایانگر یک سن فیزیولوژیک است و انتقال از یک سن فیزیولوژیک به سن فیزیولوژیک بعدی سالخوردگی را بیان می‌کند. برای هر سن فیزیولوژیک  $i$ ، دو پارامتر به کار رفته است که یکی پیشرفت فرآیند سالخوردگی و دیگری مرگ را نشان می‌دهند. از طرفی با توجه به مطالب گفته شده، زمان تا مرگ (زمان جذب) از توزیع فازنوع پیروی می‌کند. بنابراین با کمی تغییر در روش استفاده شده در مقاله، ماتریس نرخ انتقال فرآیند مارکوف را به صورت زیر در نظر می‌گیریم:

تعداد گردش زمین به دور خورشید است. اما ممکن است که این سن شناسنامه‌ای کمی با سن واقعی بدن و توانمندی سلول‌ها متفاوت باشد یعنی کمتر، برابر و یا بیشتر باشد. به این سن واقعی سن فیزیولوژیک گویند و هر کدام از این وضعیت‌ها بیانگر شرایط سلامتی خاصی خواهد بود. بنابراین سن فیزیولوژیک شاخص سلامتی عمومی است و برای بررسی‌های پزشکی و بهداشتی فاکتور معتبرتری نسبت به سن تقویمی محسوب می‌شود. دو فرد همسن به لحاظ تقویمی می‌توانند شرایط جسمانی و بهداشتی کاملاً متفاوتی داشته باشند پس سن فیزیولوژیک در واقع بیانگر سن سلامت افراد است.

روش ما در این مقاله برای مدل‌بندی فرآیند سالخوردگی یک روش پیشرفته است که در مطالعه لین و لیو [۹] استفاده شده که توسط لین و لیو بیان شده است. ما از فرآیند مارکوف پیوسته زمان با تعداد متناهی وضعیت به تعداد  $n_x + 1$  وضعیت ( $x$  سن تقویمی فرد در زمان تشخیص بیماری است) که  $n_x$  تعداد وضعیت‌های گذرا و یک وضعیت جاذب است که مرگ در آن رخ می‌دهد و بقیه وضعیت‌ها مربوط به وضعیت سلامتی و بیماری است. هر کدام از وضعیت‌های گذرا یک سن فیزیولوژیک را نشان می‌دهند و سالخوردگی در واقع انتقال از یک سن فیزیولوژیک به سن فیزیولوژیک بعدی است و این فرآیند زمانی به پایان می‌رسد که انتقال از هر وضعیتی به وضعیت مرگ اتفاق بیفتد. فرآیند مارکوف پیوسته زمان را به صورت  $\{J(t), t \geq 0\}$  که فضای وضعیت آن  $\Gamma = \{0, 1, \dots, n\}$ ،  $n \in \mathbb{N}$  می‌باشد، در نظر بگیرید که وضعیت 0 یک وضعیت جذب و بقیه وضعیت‌ها گذرا هستند. با توجه به این فرآیند فرض کنید که ماتریس شدت که نرخ بی‌درنگ تغییر وضعیت‌ها در میان وضعیت‌ها را نشان می‌دهد با  $Q$  و بردار احتمال اولیه که احتمال شروع فرآیند از هر وضعیت دلخواه است را  $\beta$  نمایش دهیم.

ماتریس شدت  $Q$  به صورت زیر تعریف می‌شود:

$$Q = \begin{bmatrix} 0 & 0' \\ t_0 & T \end{bmatrix}$$

که در آن ماتریس  $T$ ، ماتریس زیر شدت  $n \times n$  بعدی است (ماتریس  $B = (b_{ij})_{i,j=1,\dots,n}$  یک ماتریس زیر شدت نامیده می‌شود اگر  $\sum_{k=1}^n b_{ik} \leq 0$ ،  $b_{ij} \geq 0$ ،  $b_{ii} \leq 0$ ،  $i \neq j$  و  $i, j = 1, \dots, n$ ،  $t_0$  یک بردار ستونی  $n$  بعدی است که نرخ‌های انتقال از وضعیت‌های گذرا به وضعیت جاذب صفر را نشان می‌دهند. بنابراین  $T1 + t_0 = 0$  را خواهیم داشت (که 0 و 1 بردارهای ستونی  $n$  بعدی از صفرها و یک هاست). از طرفی با توجه به ماتریس شدت می‌توان به راحتی ثابت کرد که

$$e^{Qt} = \begin{bmatrix} 1 & 0' \\ 1 - e^{tT} & e^{tT} \end{bmatrix}$$

(ماتریس نمایی ماتریس مربعی  $B$  با توجه به سری‌های توانی به صورت

$$e^B = \sum_{k=0}^n \frac{B^k}{k!} \text{ تعریف می‌شود.})$$

همچنین فرض بر این است که احتمال شروع از وضعیت جاذب صفر  $P(J(0) \in \{0\}) = \beta_0 = 0$  باشد. در نتیجه بردار احتمال اولیه را به صورت  $\beta = (\beta_0, \alpha)$  در نظر می‌گیریم که  $\alpha = (\beta_1, \beta_2, \dots, \beta_n)$ . بنابراین با توجه به یک بودن مجموع احتمال‌ها خواهیم داشت:  $\alpha 1 = 1$

میزان ابتلا در مردان دو برابر بیشتر از زنان است. بنابراین سن افراد، جنسیت و نوع و مرحله سرطان و ... می‌تواند در میزان رخداد سرطان تأثیر داشته باشد.

در این بخش مدل پیشنهادی بر روی مجموعه داده‌ها برازش داده می‌شود که این داده‌ها از برنامه SEER از موسسه ملی سرطان (SEER9) در طی سال‌های (۱۹۷۳-۲۰۱۴) مربوط به ایالات متحده آمریکا از سایت [www.seer.cancer.gov] جمع آوری شده‌اند. هر بیمار دارای یک شناسه یا کد شخصی می‌باشد که از دیگر بیماران متمایز می‌شود. متغیرهایی مانند سن بیمار در زمان تشخیص بیماری، سال تشخیص، جنسیت، وضعیت حیات، زمان تا پایان مطالعه و ... نیز از بانک اطلاعاتی SEER توسط نرم افزار SEER\*Stat 8.3.4 استخراج شده است. از سال ۱۹۷۳ هر فرد با تشخیص سرطان ملانوما وارد مطالعه شده و طی دوره ای بعد از گذشت زمان (بر اساس ماه) از مطالعه به دللی خارج شده و یا در مطالعه زنده باقی می‌ماند. اگر مطالعه پایان یابد و بیمار همچنان زنده باشد، مدت زمانی که در مطالعه سپری کرده است سانسور محسوب می‌شود. هر بیمار ممکن است در طول مطالعه این بیماری را بیش از یک بار تجربه کند. بنابراین افراد با تشخیص ملانوما با ثبت جنسیت و سن وارد مطالعه می‌شوند و بدون توجه به تعداد عودها تا پایان مطالعه، مدت زمانی که سپری می‌کنند را مورد نظر قرار می‌دهیم. مدت زمانی هر شخص در مطالعه سپری می‌کند را طول عمر نامیده می‌شود و متغیر مطلوب را زمان تا مرگ در نظر می‌گیریم. با جمع آوری اطلاعات از نرم افزار ذکر شده در بالا، تعداد افراد در گروه سنی ۳۰ تا ۹۰ سال که با ابتلا شدن به سرطان پوست از نوع ملانوما در دوره مطالعاتی ۱۹۷۳ تا ۲۰۱۴ در ایالات متحده آمریکا وارد مطالعه شده‌اند، ۱۷،۹۶۲ نفر ثبت شده است که شامل ۱۱،۷۶۷ نفر مرد (۶۵/۵ درصد) و ۶،۱۹۵ نفر زن (۳۴/۵) می‌باشند. در این میان، تعداد ۸،۳۳۱ نفر بر اثر این بیماری فوت شدند. نرخ وقوع ملانوما در قبل از ۴۰ سالگی در زنان بیشتر از مردان به دست آمده است اما بعد از سن ۶۰ سالگی نرخ وقوع در مردان دو برابر زنان می‌شود.

همانطور که فرض کردیم متغیر  $Y$  زمان تا وقوع مرگ است. با استفاده از روش کاپلان-مایر، احتمال بقای  $Y$  را می‌توان برآورد کرد. بنابراین احتمال بقا را برای گروه سنی مختلف به دست می‌آوریم. همانطور که در بخش قبل گفته شد،  $J(x)$  سن فیزیولوژیک هر فرد را با سن شناسنامه‌ای  $x$  نشان می‌دهد. از آنجایی که سن فیزیولوژیک یک فرد  $x$  ساله نامعلوم است، می‌توان این مقدار را با استفاده از میانگین شرطی سن فیزیولوژیک به شرط سن شناسنامه‌ای  $x$ ،  $E(i_0|x)$ ، برآورد کرد:

$$E(i_0|x) = \sum_{i=1}^n iP(J_x = i | T > x) = \sum_{i=1}^n i \left[ \frac{ae^{Tx}}{ae^{Tx} + 1} \right]_i \quad (1)$$

در واقع شروع از سن فیزیولوژیک ۱ غیر منطقی است زیرا هر شخص در سن معلوم  $x$  بعد از تشخیص سرطان وارد مطالعه می‌شود که مقدار آن بیشتر از ۳۰ سال است. بنابراین برای هر گروه سنی معلوم، سن فیزیولوژیک از مقدار مورد انتظارش یعنی  $E(i_0|x)$  شروع می‌شود. از طرفی دیگر با توجه به مقاله، بیشترین و بهترین سن فیزیولوژیک مناسب برابر ۲۰۰ در نظر گرفته شده است، بنابراین تعداد سنین فیزیولوژیک در مدل ما  $200 - E(i_0|x)$  خواهد بود. برای به دست آورد مقدار  $E(i_0|x)$  با توجه به رابطه (۱)، باید ماتریس  $T$  را داشته باشیم.

$$T = \begin{bmatrix} -(q_1 + \lambda) & \lambda & 0 & \dots & 0 \\ 0 & -(q_2 + \lambda) & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -(q_n + \lambda) \end{bmatrix}$$

که پارامتر  $\lambda$  نرخ تغییر وضعیت از یک سن فیزیولوژیک به سن فیزیولوژیک بعدی است که اندازه پیشرفت سالخوردگی را نشان می‌دهد. بنابراین مدت زمان ماندن در هر وضعیت قبل از ورود به وضعیت بعدی دارای توزیع نمایی با میانگین  $\frac{1}{q_i + \lambda}$  است. پارامتر  $q_i$  خطر مرگ در سن فیزیولوژیک  $i$  را نشان می‌دهد. در هر سن،  $q_i$  تحت تأثیر دو عامل است، یکی مرگ مستقل از سالخوردگی،  $h_1(i)$  است مانند آسیب‌ها و تصادفات که به صورت زیر تعریف می‌شود:

$$h_i(i) = \begin{cases} a + b & \text{for } i_1 < i \leq i_2 \\ b & \text{otherwise} \end{cases}$$

که مقدار ثابت  $b$  یک نرخ پایه به عنوان بازتاب کلی از محیط زندگی است و مقدار ثابت  $a$  به عنوان عاملی است که نرخ مرگ بالاتری را در یک بازه زمانی مشخص بعد از تشخیص سرطان نشان می‌دهد. از آنجایی که پارامتر  $a$  به سن مرتبط است، بین سن فیزیولوژیک  $i_1$  و  $i_2$  ظاهر می‌شود. پس ما می‌توانیم بازه زمانی بین پارامترهای  $i_1$  و  $i_2$  را به عنوان زمانی که نرخ مرگ در آن افزایش می‌یابد در نظر بگیریم که بر اساس بهترین برازش توسط مدل پیشنهادی در این مقاله انتخاب می‌شوند. عامل دیگر  $h_2(i)$  تابعی صعودی نسبت به  $i$  است و این صعودی بودن به دلیل افزایش خطر مرگ ناشی از تابع فیزیولوژیکی سالخوردگی است. همانطور که فرض کردیم  $h_2(i)$  یک تابع توانی به شکل  $h_2(i) = qi^p$  است که  $q$  پارامتر مقیاس و  $p$  اندازه نسبی وقوع سالخوردگی است.

بنابراین پارامتر  $q_i$  به صورت

$$q_i = \begin{cases} a + b + q(i + E(i_0|x) - 1)^p & \text{for } i_1 < i \leq i_2 \\ b + q(i + E(i_0|x) - 1)^p & \text{otherwise} \end{cases}$$

خواهد بود که  $E(i_0|x)$  مقدار موردانتظار سن فیزیولوژیک به شرط معلوم بودن سن شناسنامه‌ای است که در بخش بعد توضیح داده می‌شود.

## یافته‌ها

سرطان پوست شایع‌ترین سرطان در ایالات متحده آمریکا و بسیاری از کشورهای جهان است و بر اساس آمارها تخمین زده می‌شود که از هر پنج آمریکایی یک نفر تا سن ۷۰ سالگی گرفتار سرطان پوست می‌شود. روزانه تقریباً تعداد ۹۵۰۰ نفر در ایالات متحده دچار سرطان پوست می‌شوند. سرطان پوست ملانوما نوع خطرناک این سرطان محسوب می‌شود به طوری که در هر یک ساعت، یک نفر جان خود را از دست می‌دهد. در دهه گذشته تعداد موارد جدید ملانوما که سالانه تشخیص داده شده‌اند، ۵۳ درصد افزایش یافته است. برآورد نرخ بقای پنج ساله برای بیماران مبتلا به ملانوم که در اوایل تشخیص داده می‌شود، حدود ۹۹ درصد است. هنگامی که بیماری به گره‌های لنفاوی برسد، میزان بقای آن به ۶۳ درصد کاهش می‌یابد و زمانی که بیماری به اندام‌های دور متاستاز می‌رسد، ۲۰ درصد کاهش پیدا می‌کند. تنها ۲۰ تا ۳۰٪ از ملانوم‌ها در حال‌های سیاه موجود، یافت می‌شوند، در حالی که ۷۰٪ تا ۸۰ درصد در پوست ظاهراً طبیعی ظاهر می‌شوند. قبل از سن ۵۰ سالگی، میزان ابتلا به ملانوم در زنان بالاتر از مردان و در سن ۶۵ سالگی

برآورد کرد. روش برآورد، روش کمترین مربعات است که رابطه (۲) را مینیمم می‌کند. نمودار احتمال بقا و جدول برآورد پارامترهای ماتریس  $T$  برای همه گروه سنی برای کل افراد آمریکایی در سال ۱۹۹۴ در تصویر ۱ و جدول ۱ آمده است.

$a$	$b$	$q$	$p$	$\lambda$	$(i_1, i_2]$
$1.00e-6$	$9.18e-5$	$6.06e-18$	6.65	0.18	(130,175]

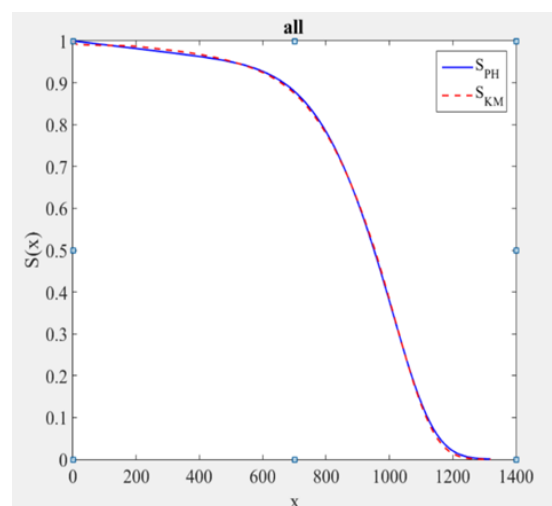
از مردان و ۲۲۳ نفر زن که ۳۵ درصد از زنان را تشکیل می‌دهند. برای گروه سنی ۶۰ تا ۶۵ سال، مقدار مورد انتظار سن فیزیولوژیک به کمک داده‌های مرگ و میر کل جمعیت آمریکایی در سال ۱۹۹۴ با توجه به جدول ۱ برابر ۱۳۶ به دست آمده است. بنابراین تعداد سن فیزیولوژیک برای این گروه سنی با توجه به ۱۳۶-۲۰۰ برابر ۶۴ خواهد بود. در این ۶۴ فاز، اگر سن فیزیولوژیک بین ۳۸ و ۵۵ باشد که شامل پارامتر  $a$  است، خطر مرگ برای بیماران بیشتر از سایر فازهاست. در نتیجه نمودار احتمال بقا و جدول برآورد پارامترها به ترتیب در تصویر ۲ و جدول ۲ آمده است:

با استفاده از تعداد سن فیزیولوژیک و پارامترهای برآورد شده برای گروه سنی مورد نظر، می‌توان مقدار امید به زندگی،  $e_x$  را به دست آورد. در این مطالعه امید به زندگی، میانگین مدت زمانی است که انتظار می‌رود که یک بیمار مبتلا به سرطان پوست بر اساس سال تولد و سن تشخیص بیماری زنده بماند. مقدار امید به زندگی با استفاده از تابع احتمال بقای فانونوع در رابطه (۳) به دست می‌آید که برای افراد ۶۰ تا ۶۵ ساله برابر با ۱۷,۲ خواهد بود. از طرفی به کمک داده‌های مرگ و میر آمریکا در سال ۱۹۹۴ مقدار امید به زندگی برای کل جمعیت برابر ۲۱,۱ سال می‌باشد. بنابراین می‌توان نتیجه گرفت که مقدار امید به زندگی در افراد سرطانی کمتر از امید به زندگی کل جمعیت آمریکا می‌باشد. علاوه بر گروه سنی و سال تشخیص بیماری، جنسیت نیز در نرخ وقوع بیماری نقش دارد. بنابراین با تفکیک داده‌ها به دو گروه مرد و زن، می‌توان به طور مجزا احتمال بقا را محاسبه و مقایسه کرد. جنسیت به عنوان یک نشانگر عمل می‌کند که مقدار صفر را برای زنان و مقدار یک را برای مردان اختیار می‌کند. با اضافه کردن متغیر جنسیت به مدل، این مدل کمی متفاوت خواهد شد، بنابراین در برآورد پارامترها نیز تغییر به وجود می‌آید. نرخ جدید مرگ و میر به صورت زیر تعریف می‌شود:

$$q_i = \begin{cases} a + b + q(i + E(i_0|x) - 1)^{p+\beta z} & \text{for } i_1 < i \leq i_2 \\ b + q(i + E(i_0|x) - 1)^{p+\beta z} & \text{otherwise} \end{cases} \quad (4)$$

که پارامتر  $\beta$  اثر متغیر جنسیت (مرد) در مدل را نشان می‌دهد.

همانطور که در بالا ذکر شد داده‌ها از سال ۱۹۷۳ تا ۲۰۱۴ جمع آوری شده‌اند که مقدار میانگین آنها برابر ۱۹۹۴ خواهد بود. با استفاده از داده‌های مرگ و میر ایالات متحده آمریکا در سایت [www.mortality.org] در سال ۱۹۹۴ می‌توان پارامترهای نامعلوم ماتریس  $T$  و مقدار مورد انتظار رابطه (۱) را برای گروه سنی مختلف جدول ۱: پارامترهای فانونوع برای جمعیت آمریکا در سال ۱۹۹۴



تصویر ۱: منحنی برازش داده شده احتمال بقا برای کل جمعیت آمریکا در سال ۱۹۹۴

حال با مشخص شدن سن فیزیولوژیک برای هر گروه سنی، پارامترهای نامعلوم  $a, b, q, p, \lambda$  را می‌توان برای بیماران سرطان پوست ملانوما برآورد کرد. با استفاده از اصل کمترین توان های دوم، مقادیر پارامترها را با مینیمم کردن مجموع وزنی توان‌های دوم احتمالات بقای توزیع فانونوع و کاپلان مایر یافتیم:

$$E = \sum_{x=0}^k (S(x) - \hat{S}(x))^2 S(x) \quad (2)$$

که  $S(x) = P(T_0 > x)$  احتمال بقا در زمان  $x$  به روش کاپلان-مایر است و  $\hat{S}(x)$  تابع بقای فانونوع در زمان مرگ  $x$  می‌باشد. همانطور که بخش قبل ذکر شد، از آنجایی که زمان مرگ دارای توزیع فانونوع با  $(\alpha, T)$  است، تابع بقا به صورت زیر بیان می‌شود:

$$\hat{S}(x) = ae^{Tx} 1 \quad (3)$$

با مینیمم کردن رابطه (۲)، ما می‌توانیم پارامترهای مجهول ماتریس  $T$  را برای هر گروه سنی برآورد کنیم.

برای محدود کردن کار، بر روی افراد با گروه سنی ۶۰ تا ۶۵ متمرکز می‌شویم. تعداد ثبت شده در ابتلا به سرطان ملانوما ۱,۸۸۲ نفر بوده است که ۱,۲۵۱ نفر را مرد (۶۶/۵ درصد) و ۶۳۱ نفر را زن (۳۳/۵ درصد) تشکیل می‌دهند. تعداد فوت شده‌ها در این دوره مطالعاتی ۷۰۲ نفر بوده که برابر ۳۷ درصد می‌باشد که در بین آنها ۴۷۹ نفر مرد که ۳۸ درصد

جدول ۲: مقادیر پارامترها برای گروه سنی ۶۰-۶۵ سال

$a$	$b$	$q$	$p$	$\lambda$	$E(i_0 x)$	$(i_1, i_2]$
0.0044	0.0028	$3.1099e-20$	3.3593	0.1745	136	(174,191]

جدول ۳. مقادیر پارامترها برای گروه سنی ۶۰-۶۵ سال به تفکیک جنسیت

$a$	$b$	$q$	$p$	$\lambda$	$\beta$	مردان ( $t_1, t_2$ )	زنان ( $t_1, t_2$ )
0.0037	0.0026	$7.70e - 20$	5.0266	0.1753	2.1073	(176,186]	(176,186]

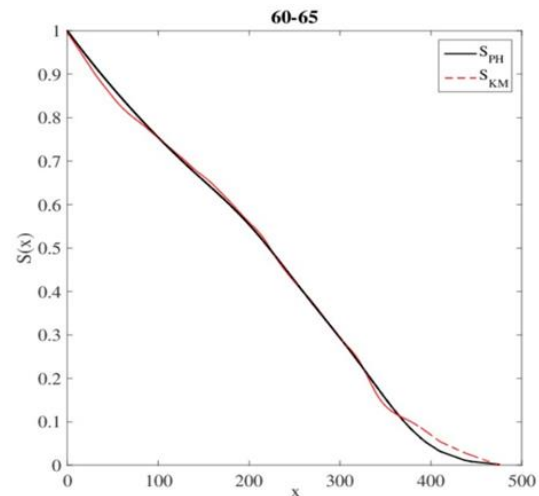
### بحث

بیماری‌های سرطانی یکی از مهمترین مشکلات بهداشتی در جهان و یکی از شایع‌ترین سرطان‌ها، سرطان پوست می‌باشد. این نوع سرطان با قرار گرفتن در معرض پرتوهای ماورابنفش که عمدتاً در نور آفتاب وجود دارد، سابقه خانوادگی، زندگی در ارتفاعات بالا، استفاده از دستگاه‌های برنزکننده پوست و لامپ‌های خورشیدی، وجود بعضی از خال‌های مادرزادی و غیره به وجود می‌آید و ممکن است گسترش پیدا کند و به قسمت‌های دیگر بدن توسعه کند و باعث مرگ فرد شود. در ایالات متحده آمریکا، سرطان پوست یکی از شایع‌ترین نوع سرطان‌ها محسوب می‌شود. در این مقاله، مدل پیشنهادی برای فرآیند سالخوردگی بیماران مبتلا به سرطان پوست توسط داده‌های واقعی انجام شد و برای گروه سنی ۶۰ تا ۶۵ ساله، احتمال بقا توسط توزیع فازنوع به دست آمد و پارامترهای سالخوردگی تحت فرآیند مارکوف برآورد شدند. در این مطالعه از لحاظ توزیع جنسی ۱,۲۵۱ نفر را مرد و ۶۲۱ نفر را زن تشکیل می‌دهند. بنابراین نسبت تعداد مردان به زنان ۱,۹۸ به دست آمده است که تقریباً دو برابر تعداد زنان است و این به دلیل حضور بیشتر مردان در زیر نور آفتاب نسبت به زنان است و همچنین کمتر از پوست خود محافظت می‌کنند مانند استفاده از کرم ضدآفتاب. در این میان تعداد ۴۷۹ نفر از مردان و ۲۲۳ نفر از زنان جان خود را از دست داده‌اند که نسبت مردان به زنان برابر ۲,۱۴ می‌باشد و این به دلیل این است که مردان به انجام آزمایش‌های پزشکی اهمیت کمتری می‌دهند و علائم اولیه بیماری را نادیده می‌گیرند. چه بسا که کشف و درمان در مراحل اولیه بیماری راحت تر و سریعتر صورت می‌گیرد [۱۲, ۱۳].

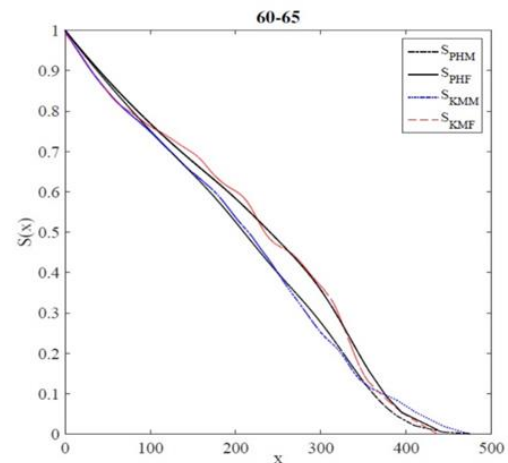
در این مطالعه احتمال بقای فازنوع با منحنی کاپلان-مایر که یک روش ناپارامتری هست مقایسه شد. ویژگی‌های تحلیلی مفیدی برای تجزیه و تحلیل داده‌های بقا در این مدل وجود دارد. فرم بسته‌ای برای تابع توزیع و تابع چگالی با برآورد پارامترها به دست آمد. با اضافه شدن متغیر جنسیت به مدل، تأثیر آن در بقای بیماران سرطانی تعیین شد که بقا در زنان بیشتر از مردان است. همانطور که نمودارها نشان می‌دهند، نتایج مربوط به برازش‌ها رضایتبخش بودند.

### نتیجه‌گیری

در بسیاری از مطالعات طولی بالینی هر فرد امکان تجربه چندین حادثه پی در پی از یک نوع را دارد. در این مطالعه داده‌های مورد استفاده شده مربوط به بیماران سرطان پوست ملانوما بوده که یک یا بیشتر از یکبار این بیماری را تجربه کرده‌اند و همانطور که نمودارها نشان می‌دهند مدل‌بندی به کمک توزیع فازنوع به خوبی انجام شده است. بنابراین آنالیز داده‌های مربوط به حوادث بازگشتی در این مطالعات بالینی برای ارزیابی اثر درمان و مطالعه روند پیشرفت بیماری‌ها بسیار حائز اهمیت است. می‌توان از توزیع فازنوع در این زمینه کمک گرفت و نتایج را بر روی بیماران به دست آورد که بیش از یک بار بیماری خاصی را تجربه



تصویر ۲: منحنی احتمال بقای بیماران سرطانی ۶۰ تا ۶۵ ساله



تصویر ۳: احتمال بقا برای مردان (-) و زنان (-) ۶۰ تا ۶۵ ساله

حال برای برآورد پارامترهای مجهول  $\lambda, p, q, b, a$  و دوباره از مینیمم کردن مجموع وزنی توان‌های دوم خطاها استفاده می‌کنیم که به صورت زیر بیان می‌شود:

$$F = \sum_{z=0}^1 \sum_{x=0}^k (S_z(x) - \hat{S}_z(x))^2 S_z(x) \quad (5)$$

که  $S_z(x)$  همان تابع بقا با اضافه شدن متغیر تصادفی جدید  $Z$  جنسیت است. برای  $z = 1$  و  $S_1(x)$  و  $\hat{S}_1(x)$  به ترتیب احتمال بقای کاپلان-مایر و تابع احتمال بقای فازنوع برای مردان می‌باشد. در نتیجه نمودار احتمال بقا و برآورد پارامترها در تصویر ۳ و جدول ۳ مشاهده می‌شود.

همانطور که در تصویر ۳ مشاهده می‌شود تابع احتمال فازنوع به خوبی به احتمال بقای کاپلان-مایر فیت شده است. احتمال بقای کاپلان مایر برای مردان با رنگ آبی و برای زنان با رنگ قرمز نشان داده شده است. می‌توان نتیجه گرفت که احتمال بقای مردان مبتلا به سرطان از زنان کمتر است که با توجه به داده‌های به دست آمده منطقی می‌باشد.

می‌کنند؛ یعنی بازگشت بیماری در آنها اتفاق می‌افتد که در مطالعات زیستی و پزشکی بسیار مفید است.

## References

1. Kleinbaum DG, Klein M. Survival Analysis. A Self-Learning Textbook: Springer; 2012.
2. Akram M, Ullah MA, Taj R. Survival analysis of cancer patients using parametric and non-parametric approaches. Pak Vet J 2007;27(4):194.
3. Royston P. The Lognormal Distribution as a Model for Survival Time in Cancer, With an Emphasis on Prognostic Factors. Statistica Neerland. 2001;55(1):89-104. doi: 10.1111/1467-9574.00158
4. Zhu HP, Xia X, Yu CH, Adnan A, Liu SF, Du YK. Application of Weibull model for survival of patients with gastric cancer. BMC Gastroenterol. 2011;11:1. doi: 10.1186/1471-230X-11-1 pmid: 21211058
5. Simsek F. Five-year survival analysis of patients with clinical stages and breast cancer who received initial treatment at North Carolina hospitals. North Carolina: North Carolina department of health and human services, 2000.
6. Wan JF, Yang LF, Shen YZ, Jia HX, Zhu J, Li GC, et al. Sex, Race, and Age Disparities in the Improvement of Survival for Gastrointestinal Cancer over Time. Sci Rep. 2016;6:29655. doi: 10.1038/srep29655 pmid: 27406065
7. Aalen OO. Phase type distributions in survival analysis. Scand J Statistics. 1995;447-63.
8. Marshall AH, Zenga M. Simulating Coxian phase-type distributions for patient survival. Int Transact Operat Res. 2009;16(2):213-26.
9. Lin XS, Liu X. Markov aging process and phase-type law of mortality. North Am Actuar J. 2007;11(4):92-109.
10. Hassan Zadeh A, Jones BL, Stanford DA. The use of phase-type models for disability insurance calculations. Scand Actuar J 2014;2014(8):714-28.
11. Asmussen S, Nerman O, Olsson M. Fitting phase-type distributions via the EM algorithm. Scand J Statistics. 1996:419-41.
12. Skin Cancer. Centers for Disease Control and Prevention [updated 2018; cited 2017]. Available from: <http://www.cdc.gov/cancer/skin>.
13. Skin cancer. American Academy of Dermatology [updated 2018; cited 2017]. Available from: <http://www.aad.org/media-resources/stats-and-facts/conditions/skin-cancer/skin-cancer>.