



ORIGINAL RESEARCH PAPER

## Estimation of the parameters of Weibull distribution and net premium in the presence of outliers

E. Sohrabi, M. Jabbari Nooghabi\*

Department of Statistics, Faculty of Mathematics of Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

### ARTICLE INFO

#### Article History:

Received 26 August 2023

Revised 01 November 2023

Accepted 15 November 2023

#### Keywords:

Estimation

Net premium

Outliers

Weibull distribution

### ABSTRACT

**BACKGROUND AND OBJECTIVES:** Weibull distribution, which was introduced by a Swedish physicist named Weibull, is today the most common model used in studies of reliability, longevity, quality control and widely in various branches of science, including Insurance, medicine and engineering are used. This distribution has a lot of flexibility for modeling different data. The main goal of this research is to calculate insurance premiums and estimate Weibull distribution parameters using different estimation methods.

**METHODS:** In this article, the parameters of the Weibull distribution and the net premium are estimated using moment estimation, maximum likelihood, least squares of error, weighted least squares, percentage, Cramer-Von- Mises, Mixture of moment and maximum-likelihood are discussed in the presence of outliers. R software was used for simulation and numerical calculations and Easyfit software was used to fit the Weibull distribution to the real example data. At the end, real data examples for obtaining different estimators of the premium in the case of unknown parameters  $\beta$  and  $\theta$  and known  $\alpha$  is provided.

**FINDINGS:** In this research bias, the mean square error of net premium and unknown parameters  $\beta$  and  $\theta$  were obtained by using different estimators for Weibull distribution data as well as the generalized variance of unknown parameters  $\beta$  and  $\theta$ .

**CONCLUSION:** In this part, the evaluation and comparison of the estimators using real and simulated data was discussed, which was obtained by different  $k_{\max}$  for real data. For example, in the moment method,  $k_{\max}$  was equal to 5, based on which the net premium is 3.37657. In the simulated data, according to  $k$  (number of outliers),  $n$  (sample size) and  $\beta$  and  $\theta$  values, bias values, mean squared error and generalized variance of premium and different estimators were obtained. As an example, for  $n=10$ ,  $k=1$ ,  $\beta=1.5$ ,  $\theta=3$  and  $\alpha=70$ , by comparing the bias and generalized variance of the estimators, we come to the conclusion that based on the bias, the percentile estimator has a better performance than the other estimators. That is, it has less bias and according to the generalized variance, the maximum likelihood estimator has a better performance than other estimators and the estimators are consistent (the generalized variance decreases with the increase of the sample size). Based on the mean square error, the moment estimation has a better performance than other estimators.

\*Corresponding Author:

Email: [jobbarinm@um.ac.ir](mailto:jobbarinm@um.ac.ir)

Phone: +9851 38805694

ORCID: [0000-0002-5636-2209](https://orcid.org/0000-0002-5636-2209)

DOI: [10.22056/ijir.2024.01.04](https://doi.org/10.22056/ijir.2024.01.04)

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).





## مقاله علمی

# برآورد پارامترهای توزیع وایبول و حق بیمه خالص در مواجهه داده‌های دورافتاده

ابراهیم سهرابی، مهدی جباری نوقابی\*

گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، مشهد، ایران

### چکیده:

**پیشینه و اهداف:** توزیع وایبول که فیزیک‌دانی سوئدی به نام وایبول معرفی کرد، امروزه رایج‌ترین مدل مورد استفاده در مطالعات قابلیت اطمینان، طول عمر، کنترل کیفیت و به‌طور وسیعی در شاخه‌های مختلف علوم از جمله بیمه، پزشکی و مهندسی استفاده می‌شود. این توزیع برای الگوسازی داده‌های مختلف، انعطاف‌پذیری زیادی دارد. هدف اصلی این پژوهش محاسبه حق بیمه و برآورد پارامترهای توزیع وایبول با استفاده از روش‌های مختلف برآورد است.

**روش‌شناسی:** در این مقاله به برآورد پارامترهای توزیع وایبول و حق بیمه خالص با استفاده از برآوردگرهای گشتاوری، ماکسیمم درست‌نمایی، کمترین توان‌های دوم خطا، کمترین توان‌های دوم وزنی، صدکی، کرامر-فون-میسز، آمیخته گشتاوری و ماکسیمم درست‌نمایی در حضور داده‌های پرت پرداخته شده است. از نرم‌افزار R برای شبیه‌سازی و محاسبات عددی و از نرم‌افزار Easyfit برای برازش توزیع وایبول به داده‌های مثال واقعی استفاده شد. در پایان دو مثال واقعی برای به‌دست آوردن برآوردگرهای متفاوت از حق بیمه در صورتی که دو پارامتر  $\theta$  و  $\beta$  مجهول و  $\alpha$  معلوم است، ارائه شده است.

**یافته‌ها:** در این پژوهش اریبی، میانگین توان دوم خطا، واریانس تعمیم یافته و حق بیمه خالص برای پارامترهای مجهول  $\beta$  و  $\theta$  با استفاده از برآوردگرهای مختلف، برای داده‌های توزیع وایبول به دست آمد.

**نتیجه‌گیری:** در این قسمت به بررسی و مقایسه برآوردگرها با استفاده از داده‌های واقعی و شبیه‌سازی شده پرداخته شد که برای داده‌های واقعی در روش‌های مختلف k-max تفاوت به‌دست آمد. برای نمونه، در روش گشتاوری k-max برابر ۵ شد که براساس آن حق بیمه خالص ۳,۳۷۶۵۷ است. در داده‌های شبیه‌سازی شده نیز با توجه به k (تعداد داده‌های دورافتاده)، n (حجم نمونه) و مقادیر  $\beta$  و  $\theta$ ، مقادیر اریبی، میانگین توان دوم خطا و واریانس تعمیم یافته حق بیمه و برآوردگرهای مختلف به‌دست آمد که برای نمونه برای  $n=10$ ،  $k=1$ ،  $\beta=1.5$ ،  $\theta=3$  و  $\alpha=70$  با مقایسه اریبی و واریانس تعمیم یافته برآوردگرها به این نتیجه می‌رسیم که براساس اریبی، برآوردگر صدکی دارای عملکرد بهتری نسبت به دیگر برآوردگرهاست، یعنی دارای اریبی کمتری است و با توجه به واریانس تعمیم یافته، برآوردگر ماکسیمم درست‌نمایی دارای عملکرد بهتری نسبت به سایر برآوردگرهاست و برآوردگرها سازگارند (با افزایش حجم نمونه واریانس تعمیم یافته کاهش می‌یابد). براساس میانگین توان دوم خطا، برآورد گشتاوری دارای عملکرد بهتری نسبت به سایر برآوردگرهاست.

### اطلاعات مقاله

#### تاریخ‌های مقاله:

تاریخ دریافت: ۰۴ شهریور ۱۴۰۲

تاریخ داوری: ۱۰ آبان ۱۴۰۲

تاریخ پذیرش: ۲۴ آبان ۱۴۰۲

#### کلمات کلیدی:

برآورد

توزیع وایبول

حق بیمه خالص

داده‌های دورافتاده

\*نویسنده مسئول:

ایمیل: [jabbarinm@um.ac.ir](mailto:jabbarinm@um.ac.ir)

تلفن: +۹۸۵۱ ۳۸۸۰۵۶۹۴

ORCID: 0000-0002-5636-2209

DOI: 10.22056/ijir.2024.01.04

توجه: مدت‌زمان بحث و انتقاد برای این مقاله تا ۱ آوریل ۲۰۲۴ در وب‌سایت IJIR در «نمایش مقاله» باز است.

جدید برای برآورد پارامترهای مدل گاما در حضور داده‌های دورافتاده اتخاذ کرد. (Dixit and Jabbari Nooghabi (2011) در مقاله‌ای تحت عنوان «برآورد پارامترهای توزیع پارتو در حضور داده‌های دورافتاده»، به برآورد پارامترهای توزیع پارتو پرداختند. دانشمندان زیادی در مواجهه با مسائل مختلف، تعاریفی از داده‌های دورافتاده ارائه کرده‌اند. در یک نمونه ممکن است، تعدادی از مشاهدات با بقیه هماهنگ نباشند که این فکر را ایجاد کند که روش نمونه‌گیری اشتباه است یا مشاهدات از جامعه دیگری باشند این مقادیر داده‌های دورافتاده هستند (Kendal and Buckland, 1957). (Anscombe (1960 یک مشاهده با باقی‌مانده بزرگ و غیرطبیعی یک مشاهده دورافتاده را تلقی می‌کند. (Hawkins (1980 معتقد بود، مشاهده‌ای که آن قدر از بقیه داده‌ها انحراف داشته باشد، به طوری که این ایده را در ذهن بیاورد که با روش و سازوکار دیگری تولید شده است، داده دورافتاده است. در این مقاله با توجه به ضرورت به برآورد پارامترها در صورتی که توزیع  $k$  و نیز  $n-k$  مشاهده هر دو وایبول باشند به برآورد پارامترهای توزیع وایبول وقتی که هر دو پارامتر  $\beta$  و  $\theta$  مجهول و  $\alpha$  معلوم باشد، پرداخته شده است. در این پژوهش هدف اصلی برآورد حق‌بیمه خالص است و چون برای برآورد آن لازم است که پارامترهای توزیع وایبول در مواجهه با داده‌های دورافتاده مشخص باشد و در اینجا مجهول است، لازم است که ابتدا پارامترهای مجهول و سپس حق‌بیمه خالص برآورد شود. یکی از وجوه تمایز این مقاله با مقالات دیگر، آن است که در اغلب موارد توزیع داده‌های دورافتاده و غیردورافتاده یکسان نیست، از طرف دیگر اگر داده دورافتاده در داده‌ها وجود داشته باشد منطقی است که یکی از پارامترهای توزیع متفاوت باشد که در این پژوهش از این رویکرد استفاده شده است، هرچند که از نظر مبانی نظری مدل‌سازی و برآورد پارامترها در این حالت پیچیده‌تر است. ادامه مقاله به صورت ذیل است: در بخش دوم روش پژوهش، توزیع وایبول، داده دورافتاده و معرفی مدل دیکشیت برای مدل‌سازی داده‌های دورافتاده پرداخته شده است. در بخش سوم به بررسی تابع چگالی در حضور داده‌های پرت، در بخش چهارم به بررسی برآوردگرهای گشتاوری (Moment Method)، کمترین ماکسیمم درست‌نمایی (Maximum Likelihood)، کمترین توان‌های دوم خطا (Least Squares Error)، کمترین توان‌های دوم وزنی (Weighted Least Squares Error)، صدک (Percentile)، برآوردگر کرامر-فون-میسز (Cramer-von-Mises)، اندرسن-دارلینگ (Anderson-Darling) و آمیخته گشتاوری و ماکسیمم درست‌نمایی در برآورد پارامترهای مجهول پرداخته شده است، در بخش پنجم حق‌بیمه خالص را به دست می‌آوریم. در بخش ششم، برآورد حق‌بیمه با استفاده از روش‌های مختلف به دست آمده و به مقایسه برآوردگرهای مختلف حق‌بیمه پرداخته شده است. در بخش هفتم دو مثال واقعی برای محاسبه حق‌بیمه ارائه و در بخش هشتم به نتیجه‌گیری و تحلیل یافته‌ها پرداخته شده است.

یکی از مهم‌ترین موضوعات در علوم بیمه، حق‌بیمه است. حق‌بیمه  $P$  مقدار پولی است که بیمه‌گر به‌منظور پوشاندن زیان‌های مورد انتظار از بیمه‌گذار دریافت می‌کند. (Young (2006 روش‌های متفاوت برای تعیین حق‌بیمه پیشنهاد داده است. فرض کنید که  $X$  مقدار ادعاهای خسارت که در گذشته رخ داده است، آنگاه حق‌بیمه خالص از فرمول  $p = E(X)$  به دست می‌آید. توزیع وایبول به‌طور فراوان در آزمایش طول عمر، تجزیه و تحلیل بقا استفاده شده است. (Weibull (1951 از این توزیع برای به دست آوردن نرخ خرابی لوله‌ی خلأ استفاده کرد، (Lieblein and Zelen (1956 از این توزیع برای بررسی نرخ خرابی بلبرینگ استفاده کردند. (Singh et al. (2013 به تجزیه و تحلیل مدل سیستم موازی با اشتراک بار  $k$  مؤلفه پرداختند با فرض اینکه توزیع زمان شکست از مؤلفه‌ها به صورت وایبول باشد. (Chaturvedi et al. (2014 به تحلیل بیزی استوار از مدل شکست وایبول تحت مدل شکست بیزی از کلاس سیگماهای آلوده شده از پیشین‌ها برای پارامترها پرداختند. (Fang et al. (2016 در مقاله‌ای تحت عنوان «طول عمر سیستم‌های سری با ترتیب نامنظم (پراکنده)» از مدل وایبول با داده‌های دورافتاده چندگانه پرداختند. در سال (Gupta and Singh (2017 در مقاله‌ای به برآورد کلاسیک و بیزی از توزیع وایبول در حضور داده‌های دورافتاده پرداختند که در آن داده‌های دورافتاده از توزیع نمایی تبعیت می‌کنند. (Strzelecki (2021 یک نسخه سه پارامتری به‌روزشده از توزیع وایبول برای تعیین طول عمر برای سطوح مختلف استرس معرفی کرد. (Sindhu and Atangana (2021 توزیع وایبول نمایی معکوس برای تجزیه و تحلیل قابلیت اطمینان به کار بردند. (Abubakar and Muhammad Sabri (2021 یک شکل تعدیل‌شده از مدل وایبول را برای تجزیه و تحلیل بازده سرمایه‌گذاری در نظر گرفتند. (Liu et al. (2021 از نسخه جدیدی از توزیع وایبول برای تحلیل پدیده کووید-۱۹ استفاده کردند. (Alyami et al. (2022 از یک توزیع وایبول اصلاح‌شده برای تحلیل پزشکی و مهندسی مجموعه داده‌ها استفاده کردند. در آزمایش‌های مربوط به طول عمر گاهی، یک آزمایشگر با موقعیتی مواجه می‌شود که برخی مشاهدات در یک نمونه در مقایسه با مجموعه مشاهدات باقی‌مانده بسیار کوچک یا خیلی بزرگ‌اند. به عبارت دیگر، در طول یک آزمایش، با شرایطی مواجه می‌شویم که ممکن است یک یا چند مشاهدات با بقیه مشاهدات همگن نباشند و از این رو می‌توان آن‌ها را به عنوان داده دورافتاده در نظر گرفت (Barnett and Lewis, 1984). تشخیص نقاط دورافتاده هم می‌تواند مهم باشد، زیرا با تشخیص مشاهدات دورافتاده، فرد می‌تواند از آلودگی تخمین‌های بعدی جلوگیری کند. (Kale and Sinha (1971 برآورد طول عمر مورد انتظار را در حضور داده‌های دورافتاده، در زمینه آزمایش‌های طول عمر با در نظر گرفتن توزیع نمایی، به عنوان یک مدل طول عمر در نظر گرفتند. (Pettit (1988 رویکرد بیزی را برای مدل‌سازی داده دورافتاده اتخاذ کرد و از آن در رابطه با اعضای خانواده نمایی استفاده کرد. (Dixit (1989 یک روش

## روش‌شناسی پژوهش

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i) \sum_{(i_1, i_2, \dots, i_k)} \prod_{j=1}^k \frac{g(x_{i_j})}{f(x_{i_j})} [C(n, k)]^{-1}, \quad (3)$$

که در آن:

$$\sum_{(i_1, i_2, \dots, i_k)} = \sum_{i_1=1}^{n-k+1} \sum_{i_2=i_1+1}^{n-k+2} \dots \sum_{i_k=i_{k-1}+1}^n. \quad (4)$$

همچنین برای تک‌مشاهده  $X$ ، تابع چگالی و توزیع متغیر تصادفی به شکل زیر نوشته می‌شود:

$$h(x) = bg(x) + \bar{b}f(x), \quad (5)$$

و

$$H(x) = bG(x) + \bar{b}F(x). \quad (6)$$

که در آن  $b = \frac{k}{n}$  و  $\bar{b} = 1 - b$ . شایان ذکر است که در مدل دیکشیت، متغیرهای تصادفی  $(X_1, X_2, \dots, X_n)$  مستقل نیستند.

تابع چگالی توأم  $(X_1, X_2, \dots, X_n)$  در حضور داده‌های دورافتاده فرض کنید، متغیرهای تصادفی  $(X_1, X_2, \dots, X_n)$  به صورتی باشند که هر  $k$  تا از آن‌ها مستقل و هم توزیع وایبول با پارامترهای  $\alpha$ ،  $\beta$  و  $\theta$  به صورت معادله (۷) باشد.

$$g(x; \alpha, \beta, \theta) = \alpha (\beta \theta)^\alpha x^{\alpha-1} e^{-(\beta \theta x)^\alpha}, \quad x > 0, \alpha > 0, \beta > 1, \theta > 0, \quad (7)$$

و  $n - k$  تای باقی‌مانده نیز مستقل و هم توزیع وایبول و از معادله (۱) تبعیت می‌کند، آنگاه تابع چگالی توأم  $(X_1, X_2, \dots, X_n)$  در حضور داده‌های دورافتاده بر اساس معادله (۳) به صورت معادله (۸) خواهد بود.

$$f(x_1, x_2, \dots, x_n; \alpha, \beta, \theta) = \frac{\alpha^n \theta^{n\alpha} e^{-\theta^\alpha \sum_{i=1}^n x_i^\alpha} \beta^{k\alpha}}{C(n, k)} \prod_{i=1}^n (x_i)^{\alpha-1} \sum_{(i_1, i_2, \dots, i_k)} \prod_{j=1}^k e^{-(\theta^\alpha)(\beta^\alpha - 1)(x_{i_j})^\alpha}, \quad (8)$$

و تابع چگالی حاشیه‌ای  $X_i$  بر اساس معادله (۵) به صورت معادله (۹) خواهد بود.

$$f(x; \alpha, \beta, \theta) = b\alpha(\beta\theta)^\alpha x_i^{\alpha-1} e^{-(\beta\theta x_i)^\alpha} + \bar{b}\alpha(\theta)^\alpha x_i^{\alpha-1} e^{-(\theta x_i)^\alpha} = \alpha\theta^\alpha x_i^{\alpha-1} (b\beta^\alpha e^{-(\theta x_i)^\alpha} + \bar{b}e^{-(\theta x_i)^\alpha}), \quad (9)$$

از آن جا که توزیع وایبول کاربرد فراوانی در مدل سازی داده‌های چوله به راست به خصوص متغیرهای اقتصادی دارد و نیز در مدل سازی میزان و ادعای خسارت در بیمه به کار برده می‌شود و از طرفی، در مواجهه با داده‌های دور افتاده مورد توجه محققین قرار نگرفته است، بنابراین در این پژوهش توزیع وایبول در مواجهه با داده‌های دور افتاده انتخاب شده است. لذا ابتدا پارامترهای توزیع وایبول را با توجه به روش‌های مختلف برآورد می‌کنیم و کارایی آن‌ها مقایسه می‌شود. سپس حقیقیه خالص براساس روش‌های گشتاوری، ماکسیمم درست‌نمایی، کمترین توان‌های دوم خطا، کمترین توان‌های دوم خطای وزنی، صدکی، کرامر\_فون\_میسز، اندرسن دارلینگ و آمیخته گشتاوری و ماکسیمم درست‌نمایی برآورد می‌کنیم. در انتها با استفاده از دو مثال کاربردی نتایج عملی آن را گزارش می‌کنیم.

## توزیع وایبول

توزیع وایبول که به وسیله یک فیزیکدان سوئدی به نام Weibull (1951) معرفی شد، امروزه متداول‌ترین مدل مورد استفاده در مطالعات قابلیت اطمینان، طول عمر، کنترل کیفیت است و به‌طور وسیعی در شاخه‌های مختلف علوم، از جمله بیمه، پزشکی و مهندسی استفاده می‌شود. این توزیع برای الگوسازی داده‌های مختلف، دارای انعطاف‌پذیری زیادی است. فرض کنید متغیرهای تصادفی  $(X_1, X_2, \dots, X_n)$  دارای توزیع وایبول باشند، تابع چگالی و توزیع آن به ترتیب به صورت معادلات (۱) و (۲) نوشته می‌شود.

$$f(x) = \alpha \theta^\alpha x^{\alpha-1} e^{-(\theta x)^\alpha}, \quad x > 0, \theta > 0, \alpha > 0, \quad (1)$$

و

$$F(x) = 1 - e^{-(\theta x)^\alpha}, \quad x > 0, \alpha > 0, \theta > 0. \quad (2)$$

## داده دورافتاده

داده دورافتاده، مشاهده‌ای است که به‌طور غیرعادی یا اتفاقی از وضعیت عمومی داده‌های تحت آزمایش و نسبت به قاعده‌ای که براساس آن آنالیز می‌شود، انحراف داشته باشد.

برای مدل‌سازی داده‌های دورافتاده می‌توان از مدل‌های مختلفی استفاده کرد. حال اگر بخواهیم به‌طور کلی مدلی بیان کنیم، تا براساس آن توزیع و چگالی یک متغیر تصادفی در مواجهه با  $k$  داده دورافتاده را به‌دست آورد، می‌توانیم از مدل زیر که به مدل Dixit (1987) معروف است، استفاده کنیم.

## مدل دیکشیت

فرض کنید  $X_i$  ( $i \geq 1$ ) یک‌سری از متغیرهای تصادفی نامنفی باشد. آنگاه تابع چگالی توأم  $(X_1, X_2, \dots, X_n)$  به صورت زیر خواهد بود:

روش‌های برآورد

بنابراین برآورد گشتاوری پارامتر  $\theta$  به صورت معادله (۱۵) است.

$$\hat{\theta}_{mm} = \frac{-A_2 + \sqrt{\Delta}}{2A_1}, \quad (15)$$

و در نهایت با توجه به معادله (۱۴) و جای گذاری  $\hat{\theta}_{mm}$  در آن، برآورد گشتاوری پارامتر  $\beta$  به صورت معادله (۱۶) نوشته می‌شود.

$$\hat{\beta}_{mm} = \frac{b\Gamma\left(1+\frac{1}{\alpha}\right)}{\bar{x}\hat{\theta}_{mm} - \bar{b}\Gamma\left(1+\frac{1}{\alpha}\right)}, \quad (16)$$

که در آن باید  $\bar{x}\hat{\theta}_{mm} > \bar{b}\Gamma\left(1+\frac{1}{\alpha}\right)$  باشد. با توجه به معلوم بودن  $n$  و  $k$  در مدل دیکشیت و با توجه به روش‌های عددی و شبیه‌سازی برای نمونه‌های مختلف، مقدار  $\hat{\beta}_{mm}$  به دست می‌آید.

برآوردگر ماکسیمم درست‌نمایی

فرض کنید  $X_1, X_2, \dots, X_n$  متغیرهای تصادفی با حجم  $n$  از تابع چگالی وایبول در حضور داده‌های دورافتاده باشند و فرض کنیم:

$$L = L(\beta, \theta | (x_1, x_2, \dots, x_n, \alpha)) = f(x_1, x_2, \dots, x_n, \alpha, \beta, \theta), \quad (17)$$

بنابراین، با توجه به معادله (۸) و گرفتن لگاریتم طبیعی از طرفین آن داریم:

$$\ln(L) = n \ln(\alpha) + n\alpha \ln(\theta) + (\alpha - 1) \sum_{i=1}^n \ln(x_i) - \theta^\alpha \sum_{i=1}^n (x_i)^\alpha + \alpha k \ln(\beta) + \ln\left(\sum_{(i_1, i_2, \dots, i_k)} Z\right), \quad (18)$$

که در آن،  $Z$  به صورت معادله (۱۹) است.

$$Z = \prod_{j=1}^k e^{-(\theta^\alpha)(\beta^\alpha - 1)(x_{i_j})^\alpha} = e^{-(\theta^\alpha)(\beta^\alpha - 1)\sum_{j=1}^k (x_{i_j})^\alpha}, \quad (19)$$

لذا، برآوردگر ماکسیمم درست‌نمایی  $\beta$  به صورت معادله (۲۰) به دست می‌آید.

$$\frac{\partial \ln(L)}{\partial \beta} = \frac{K\alpha}{\beta} + \frac{\left(\sum_{(i_1, i_2, \dots, i_k)} Z\right)}{\left(\sum_{(i_1, i_2, \dots, i_k)} Z\right)} = 0, \quad (20)$$

که در آن

در این قسمت با استفاده از روش‌های مختلف، به برآورد پارامترهای  $\alpha$ ،  $\beta$  و  $\theta$  می‌پردازیم که در آن معلوم  $\alpha$  و  $\beta$  و  $\theta$  مجهول است.

روش گشتاوری

فرض کنید  $(X_1, X_2, \dots, X_n)$  متغیرهای تصادفی با حجم  $n$  از تابع چگالی وایبول در حضور داده‌های دورافتاده باشند، داریم:

$$E(X) = b\frac{1}{\beta\theta}\Gamma\left(1+\frac{1}{\alpha}\right) + \bar{b}\frac{1}{\theta}\Gamma\left(1+\frac{1}{\alpha}\right), \quad (10)$$

و

$$E(X)^2 = b\frac{1}{(\beta\theta)^2}\Gamma\left(1+\frac{2}{\alpha}\right) + \bar{b}\frac{1}{\theta^2}\Gamma\left(1+\frac{2}{\alpha}\right). \quad (11)$$

برای به دست آوردن برآورد گشتاوری دو پارامتر مجهول  $\beta$  و  $\theta$  لازم است که گشتاورهای جامعه را با گشتاورهای متناظر در نمونه برابر قرار دهیم، لذا:

$$E(X) = b\frac{1}{\beta\theta}\Gamma\left(1+\frac{1}{\alpha}\right) + \bar{b}\frac{1}{\theta}\Gamma\left(1+\frac{1}{\alpha}\right) = \bar{x}, \quad (12)$$

$$E(X)^2 = b\frac{1}{(\beta\theta)^2}\Gamma\left(1+\frac{2}{\alpha}\right) + \bar{b}\frac{1}{\theta^2}\Gamma\left(1+\frac{2}{\alpha}\right) = \bar{x}^2. \quad (13)$$

بنابراین با توجه به معادله (۱۲) با محاسبه  $\beta$  داریم:

$$\beta = \frac{b\frac{1}{\theta}\Gamma\left(1+\frac{1}{\alpha}\right)}{\bar{x} - \bar{b}\frac{1}{\theta}\Gamma\left(1+\frac{1}{\alpha}\right)}, \quad (14)$$

و همچنین با توجه به معادله (۱۳) و به توان ۲ رساندن معادله (۱۴) و برابر قرار دادن این دو معادله و بعد از محاسبات جبری نتیجه می‌گیریم:

$$\left(\bar{x}^2 b \Gamma^2\left(1+\frac{1}{\alpha}\right) \Gamma\left(1+\frac{2}{\alpha}\right) - (\bar{x})^2 b \Gamma\left(1+\frac{2}{\alpha}\right)\right) \theta^2 + \left(2\bar{x}\bar{b}b \Gamma\left(1+\frac{1}{\alpha}\right) \Gamma\left(1+\frac{2}{\alpha}\right)\right) \theta - b(\bar{b})^2 \Gamma^2\left(1+\frac{1}{\alpha}\right) \Gamma\left(1+\frac{2}{\alpha}\right) = 0$$

$$A_1 \theta^2 + A_2 \theta + A_3 = 0, \quad \Delta = (A_2)^2 - 4A_1 A_3$$

مشاهدات توزیع  $G(X)$  و  $(n-k)$  تایی باقی مانده از توزیع  $G(X)$  پیروی کند، براساس مدل دیکشیت تابع توزیع حاشیه‌ای آن  $H(X)$  است.

$$H(x) = bG(x) + \bar{b}F(x) = b\left(1 - e^{-(\theta x)^\alpha}\right) + \bar{b}\left(1 - e^{-(\theta x)^\alpha}\right) = 1 - be^{-(\theta x)^\alpha} - \bar{b}e^{-(\theta x)^\alpha}. \quad (27)$$

بنابراین، مینیمم کردن معادله (28)، برآورد کمترین توان‌های دوم خطا  $\beta$  و  $\theta$  را می‌دهد.

$$Z = \sum_{j=1}^n \left( H(X_{(j)}; \alpha, \beta, \theta) - \frac{j}{n+1} \right)^2 = \sum_{j=1}^n \left( -be^{-(\theta x_{(j)})^\alpha} - \bar{b}e^{-(\theta x_{(j)})^\alpha} + 1 - \frac{j}{n+1} \right)^2. \quad (28)$$

لذا، برآورد کمترین توان‌های دوم خطا  $\theta$  به صورت (29) به دست می‌آید.

$$\frac{\partial Z}{\partial \beta} = \sum_{j=1}^n \left( 2b\alpha(\theta x_{(j)})^\alpha \beta^{\alpha-1} e^{-(\theta x_{(j)})^\alpha} \right) \left( -\bar{b}e^{-(\theta x_{(j)})^\alpha} - be^{-(\theta x_{(j)})^\alpha} + 1 - \frac{j}{n+1} \right) = 0, \quad (29)$$

که دارای فرم بسته‌ای نیست و همچنین برآورد کمترین توان‌های دوم خطا  $\theta$  به صورت معادله (30) است.

$$\frac{\partial Z}{\partial \theta} = \sum_{j=1}^n \left( 2\bar{b}\alpha(x_{(j)})^\alpha \theta^{\alpha-1} e^{-(\theta x_{(j)})^\alpha} + 2b\alpha(\beta(x_{(j)}))^\alpha \theta^{\alpha-1} e^{-(\theta \beta x_{(j)})^\alpha} \right) \left( -\bar{b}e^{-(\theta x_{(j)})^\alpha} - be^{-(\theta \beta x_{(j)})^\alpha} + 1 - \frac{j}{n+1} \right) = 0. \quad (30)$$

برآوردگر کمترین توان‌های دوم خطای وزنی با توجه به مطالب ذکر شده در برآوردگر کمترین توان‌های دوم خطا، برآوردگر کمترین توان‌های دوم خطای وزنی با مینیمم کردن معادله (31) به دست می‌آید.

$$\sum_{j=1}^n w_j \left[ F(X_{(j)}; \alpha, \beta, \theta) - \frac{j}{n+1} \right]^2, \quad (31)$$

که در آن

$$w_j = \frac{1}{\text{Var}(F(X_j; \alpha, \theta))}, \quad (32)$$

و نیز،  $j = 1, 2, \dots, n$ . بنابراین، مینیمم‌سازی

$$Z' = \frac{\partial Z}{\partial \beta} = - \left( \alpha \beta^{\alpha-1} \theta^\alpha \sum_{j=1}^k X_{ij}^\alpha \right) \left( e^{-\theta^\alpha \beta^{(\alpha-1)} \sum_{j=1}^k X_{ij}^\alpha} \right), \quad (21)$$

با جای‌گذاری (21) در معادله (20)، برآوردگر ماکسیمم درست‌نمایی  $\beta$  به دست می‌آید که دارای فرم بسته‌ای نیست و برآوردگر ماکسیمم درست‌نمایی  $\theta$  به صورت معادله (22) محاسبه می‌شود.

$$\frac{\partial \ln(L)}{\partial \theta} = \frac{n\alpha}{\theta} - \alpha \theta^{\alpha-1} \left( \sum_{i=1}^n X_i^\alpha \right) + \frac{\left( \sum_{(i_1, i_2, \dots, i_k)} Z \right)'}{\left( \sum_{(i_1, i_2, \dots, i_k)} Z \right)} = 0, \quad (22)$$

که در آن

$$Z' = \frac{\partial Z}{\partial \theta} = - \left( \alpha \theta^{(\alpha-1)} \beta^{(\alpha-1)} \sum_{j=1}^k X_{ij}^\alpha \right) e^{-\theta^\alpha \beta^{(\alpha-1)} \left( \sum_{j=1}^k X_{ij}^\alpha \right)}. \quad (23)$$

با جای‌گذاری (23) در معادله (22) برآوردگر ماکسیمم درست‌نمایی  $\theta$  به دست می‌آید که دارای فرم بسته‌ای نیست. معادله‌های (20) و (22) برای برآورد پارامترهای مجهول  $\beta$  و  $\theta$  می‌تواند با استفاده از هر روش تکرار عددی مانند نیوتن-رافسن به‌طور هم‌زمان حل شود.

برآوردگر کمترین توان‌های دوم خطا

اگر  $(X_1, X_2, \dots, X_n)$  یک نمونه تصادفی با تابع توزیع  $F(\cdot)$  و  $X_1 \leq X_2 \leq \dots \leq X_n$  نشان‌دهنده آماره‌های مرتب باشد، با توجه به [Johnson et al. \(1995\)](#) امید ریاضی، واریانس و کوواریانس تابع توزیع به صورت معادلات (24) تا (26) نوشته می‌شوند.

$$E(F(X_{(j)}; \alpha, \beta, \theta)) = \frac{j}{n+1}, \quad (24)$$

$$\text{Var}(F(X_{(j)}; \alpha, \beta, \theta)) = \frac{j(n+1-j)}{(n+1)^2(n+2)}, \quad (25)$$

و

$$\text{Cov}(F(X_{(j)}; \alpha, \beta, \theta), F(X_{(i)}; \alpha, \beta, \theta)) = \frac{j(n+1-i)}{(n+1)^2(n+2)}, \quad (26)$$

که در آن  $i, j = 1, 2, \dots, n, i \neq j$  با مینیمم کردن تابع  $\sum_{j=1}^n \left( F(X_{(j)}; \alpha, \beta, \theta) - \frac{j}{n+1} \right)^2$  نسبت به پارامترهای مجهول، برآوردگرهای کمترین توان‌های دوم خطا به دست می‌آید. فرض کنید  $k$  تا از

(1974) Stephens مراجعه کنید).

$$\sum_{j=1}^n w_j \left[ -be^{-(\theta\beta x_{(j)})^\alpha} - \bar{b}e^{-(\theta x_{(j)})^\alpha} + 1 - \frac{j}{n+1} \right]^2, \quad (33)$$

آنگاه مینیمم‌سازی معادله (۳۷) نسبت به  $\theta$  و  $\beta$  به ترتیب  $\hat{\theta}_{ad}$  و  $\hat{\beta}_{ad}$  را می‌دهد.

نسبت به  $\theta$  و  $\beta$  به ترتیب  $\hat{\theta}_{wls}$  و  $\hat{\beta}_{wls}$  را می‌دهد که در آن:

$$AD(\alpha, \theta) = -n - \frac{1}{n} \sum_{j=1}^n (2j-1) \quad (37)$$

$$\left[ \ln(F(X_{(j)}; \alpha, \beta, \theta)) + \ln(\bar{F}(X_j; \alpha, \beta, \theta)) \right]^2 =$$

$$-n - \frac{1}{n} \sum_{j=1}^n (2j-1) \left[ \ln \left( -be^{-(\theta\beta x_{(j)})^\alpha} - \bar{b}e^{-(\theta x_{(j)})^\alpha} + 1 \right) + \ln \left( be^{-(\theta\beta x_{(n+1-j)})^\alpha} + \bar{b}e^{-(\theta x_{(n+1-j)})^\alpha} \right) \right]^2,$$

$$w_j = \frac{(n+1)^2(n+2)}{j(n+1-j)}, \quad j=1, 2, \dots, n.$$

برآوردگر صدکی

روش دیگری که سعی می‌کنیم پارامترها را براساس آن برآورد کنیم، برآوردگر صدکی است. برآوردگر صدکی را (Kao 1958) و (Kao 1959) ارائه کردند. در این روش براساس معکوس تابع توزیع تجمعی برآوردگرها به دست آمده است. بنابراین مینیمم‌سازی

$$\sum_{j=1}^n [F(X_j; \alpha, \beta, \theta) - p_j]^2, \quad (34)$$

که در آن

$$\bar{F}(\cdot) = 1 - F(\cdot), \quad (38)$$

نسبت به پارامترهای مجهول، برآوردگر صدکی از پارامترها را می‌دهد که  $p_j$  صدک  $j$ ام و  $j = 1, 2, \dots, n$  است. بنابراین،  $\hat{\theta}_p$  و  $\hat{\beta}_p$  می‌تواند با مینیمم‌سازی معادله (۳۵) به دست آید.

$$\sum_{j=1}^n \left[ -be^{-(\theta\beta x_{(j)})^\alpha} - \bar{b}e^{-(\theta x_{(j)})^\alpha} + 1 - p_j \right]^2, \quad (35)$$

برآوردگر کرامر\_فون\_میسز

در اینجا، برآوردگر کرامر\_فون\_میسز پارامترهای مجهول بحث شده است.  $\hat{\theta}_{cvm}$  و  $\hat{\beta}_{cvm}$  با مینیمم کردن معادله (۳۶) نسبت به  $\theta$  و  $\beta$  به ترتیب به دست آمده است.

$$C(\alpha, \theta) = \frac{1}{12n} + \sum_{j=1}^n \left[ F(X_j; \alpha, \theta) - \frac{2j-1}{2n} \right]^2 = \quad (36)$$

$$\frac{1}{12n} + \sum_{j=1}^n \left[ -be^{-(\theta\beta x_{(j)})^\alpha} - \bar{b}e^{-(\theta x_{(j)})^\alpha} + 1 - \frac{2j-1}{2n} \right]^2,$$

برآوردگر اندرسن\_دارلینگ

در نهایت، برآوردگر اندرسن\_دارلینگ پارامترهای مجهول توزیع وایبول در حضور داده‌های دورافتاده در این بخش به دست آمده است. یکی از آزمون‌های نیکویی برازش برای بررسی نرمال بودن، آماره آزمون اندرسن\_دارلینگ است که اندرسن و دارلینگ در سال ۱۹۵۲ پیشنهاد کردند. این روش، آزمون همگرایی سریع در شرایط مجانبی است (برای اطلاعات بیشتر به Anderson and Darling (1954) و Pettitt (1976) و

مدل آمیخته گشتاوری و ماکسیمم درست‌نمایی (1981) Read روشی را ارائه کرد که از مشکل معادلات پیچیده جلوگیری می‌کند. براساس این روش، جایگزینی برخی اما نه همه معادلات سیستم، ممکن است آن را بیشتر کنترل کند. با توجه به معادله (۱۵) برآوردگر آمیخته پارامتر  $\theta$  به صورت معادله (۳۹) است.

$$\hat{\theta}_{mix} = \hat{\theta}_{mm} = \frac{-A_2 + \sqrt{\Delta}}{2A_1} = \frac{-A_2 + \sqrt{A_2^2 - 4A_3A_1}}{2A_1}, \quad (39)$$

و همچنین برای به دست آوردن برآوردگر آمیخته  $\beta$  (یعنی  $\hat{\beta}_{mix}$ )، کافی است که در معادله (۲۰) برای برآورد ماکسیمم درست‌نمایی  $\beta$ ، به جای پارامتر  $\theta$ ، برآوردگر آمیخته آن، یعنی  $\hat{\theta}_{mix}$  را قرار دهیم.

$$\frac{\alpha K}{\beta} \frac{\sum_{(i_1, i_2, \dots, i_k)} \alpha \beta^{\alpha-1} \hat{\theta}_{mix}^\alpha \sum_{j=1}^k (x_{i_j})^\alpha e^{-\hat{\theta}_{mix}(\beta^\alpha-1)\sum_{j=1}^k (x_{i_j})^\alpha}}{\sum_{(i_1, i_2, \dots, i_k)} e^{-\hat{\theta}_{mix}(\beta^\alpha-1)\sum_{j=1}^k (x_{i_j})^\alpha}} = 0, \quad (40)$$

که در آن با حل معادله (۴۰) برآوردگر آمیخته  $\beta$ ، یعنی  $\hat{\beta}_{mix}$  به دست می‌آید.

محاسبه حق بیمه خالص

در این بخش، برآوردگر حق بیمه خالص با استفاده از برآوردگرهای مختلف پارامترهای مجهول به دست می‌آید. می‌دانیم که حق بیمه خالص به صورت معادله (۴۱) تعریف شده است.

مقایسه برآوردگرهای حق بیمه

$$P = E(X), \quad (41)$$

در بخش‌های چهارم و پنجم، برآوردگرهای مختلف حق بیمه توزیع وایبول در حضور داده‌های دورافتاده به دست آمد. در این بخش، عملکرد همه برآوردگرها به طور عددی از لحاظ اریبی، میانگین توان دوم خطا و واریانس تعمیم یافته مقایسه شده‌اند. به منظور بررسی کارایی برآوردگرها، وقتی  $\beta$  و  $\theta$  مجهول و  $\alpha$  معلوم باشد، در نمونه‌هایی به اندازه  $n = 10, 20, 30, 40, 50, 60, 70, 80$  برای  $k = 1, 3, \alpha = 70, \theta = 3$  و  $\beta = 1.5$  این فرآیند ۵۰۰۰ بار تکرار شد. نتایج به شرح جداول ۱ تا ۶ و نمودارهای ۱ تا ۳ آمده است. برای شبیه‌سازی و محاسبات عددی از نرم‌افزار R استفاده شده است. خلاصه‌ای از نتایج حاصل از مطالعه شبیه‌سازی به شرح ذیل است:

که در آن  $X$  نمایانگر حق بیمه است (برای اطلاعات بیشتر درباره روش‌های تعیین حق بیمه به Young (2006) مراجعه کنید). با استفاده از معادله (۱۰) داریم.

$$p = E(X) = \frac{1}{\theta} \Gamma\left(1 + \frac{1}{\alpha}\right) \left(b \frac{1}{\beta} + \bar{b}\right). \quad (42)$$

در قسمت‌های قبل، برآوردهای  $\beta$  و  $\theta$  از روش‌های مختلف به دست آمد، ولی فرم بسته‌ای برای آن‌ها حاصل نشد، لذا با جایگزین کردن این برآوردگرها در معادله (۴۲)، برآوردگر حق بیمه به صورت برآورد جایگزین شده (Plug-in Estimator) حاصل می‌شود.

جدول ۱: واریانس تعمیم یافته  $\hat{\theta}$  و  $\hat{\beta}$ ، به ازای  $k = 1, \alpha = 70, \beta = 1.5, \theta = 3$

Table 1: Generalized variance of  $\hat{\beta}$  and  $\hat{\theta}$   $k = 1, \alpha = 70, \beta = 1.5, \theta = 3$

	80	70	60	50	40	30	20	10	
	3.87100E-05	1.26922E-06	6.06231E-07	3.21285E-07	2.14166E-07	1.69703E-07	1.55588E-07	1.92214E-07	$GV_{mm}$
	1.97538E-08	1.99333E-08	2.46108E-08	2.84662E-08	3.65628E-08	5.22965E-08	8.36677E-08	1.79596E-07	$GV_{ml}$
	3.11076E-08	3.15600E-08	3.85760E-08	4.25733E-08	5.63784E-08	7.93896E-08	1.21521E-07	2.44235E-07	$GV_{mix}$
	7.97425E-09	5.61779E-06	4.42858E-05	7.15781E-06	2.35275E-05	1.68057E-03	4.80476E-02	1.23522E-01	$GV_{ls}$
	2.45829E-08	2.46707E-08	2.56103E-06	9.78616E-05	9.19961E-03	4.59410E-02	1.32637E-01	8.15719E-02	$GV_{wls}$
	2.93741E-08	1.46580E-09	6.78105E-07	2.93950E-05	3.05403E-05	6.63587E-02	4.79328E-05	4.39627E-02	$GV_p$
	1.91250E-09	5.42587E-08	2.81324E-08	3.09258E-08	1.11028E-08	1.26374E-05	5.50953E-05	3.15833E-04	$GV_{cvm}$
	8.01101E-06	6.95291E-06	8.25041E-06	8.91714E-06	9.42261E-06	9.42261E-06	9.24930E-06	9.55168E-06	$GV_{ad}$

جدول ۲: اریبی حق بیمه،  $\hat{\beta}$  و  $\hat{\theta}$  به ازای  $k = 1$  و  $\alpha = 70, \beta = 1.5, \theta = 3$

Table 2: Bias of premium,  $\hat{\beta}$  and  $\hat{\theta}$  for  $\alpha = 70, \beta = 1.5, \theta = 3$  and  $k = 1$

	80	70	60	50	40	30	20	10	$n$ برآوردها
	6.613597	4.608409	3.402212	2.593288	2.013110	1.576279	1.235106	0.955721	$\hat{\beta}_{mm}$
	-0.020410	-0.021725	-0.023361	-0.025361	-0.028951	-0.034674	-0.045903	-0.080321	$\hat{\theta}_{mm}$
	0.009641	0.009457	0.009198	0.008803	0.008276	0.007367	0.005515	-0.000010	$\hat{p}_{mm}$
	0.013005	0.012189	0.012264	0.012404	0.012592	0.012583	0.012992	0.011409	$\hat{\beta}_{ml}$
	0.000300	0.000245	0.000301	0.000487	0.000497	0.000675	0.001127	0.002528	$\hat{\theta}_{ml}$
	0.009589	0.009397	0.009124	0.008731	0.008170	0.007218	0.005302	-0.000420	$\hat{p}_{ml}$
	0.023521	0.023343	0.024286	0.025545	0.027582	0.030612	0.037074	0.054289	$\hat{\beta}_{mix}$
	-0.020410	-0.021725	-0.023361	-0.025361	-0.028951	-0.034674	-0.045903	-0.080321	$\hat{\theta}_{mix}$
	0.011859	0.011802	0.011709	0.011547	0.011365	0.011029	0.010302	0.008020	$\hat{p}_{mix}$
	0.000038	0.000299	0.000659	0.015561	0.029723	0.039065	0.056561	0.152185	$\hat{\beta}_{ls}$
	-0.003550	-0.003666	-0.007294	-0.002184	-0.006010	-0.003766	-0.013328	-0.014508	$\hat{\theta}_{ls}$
	0.009245	0.009061	0.008408	0.008489	0.007509	0.006806	0.003836	-0.002298	$\hat{p}_{ls}$
	0.038261	0.039054	0.044906	0.061221	0.072476	0.106568	0.179438	0.164468	$\hat{\beta}_{wls}$
	0.000085	-0.000022	-0.000053	0.000231	-0.001159	-0.003675	-0.007245	-0.013306	$\hat{\theta}_{wls}$
	0.009568	0.009372	0.009097	0.008679	0.007986	0.006752	0.004361	-0.002463	$\hat{p}_{wls}$
	0.000078	0.000094	0.000322	0.024120	0.042782	0.077545	0.057565	0.162843	$\hat{\beta}_p$
	-0.003718	-0.000092	-0.005916	0.000690	-0.002708	-0.000921	-0.004132	0.002380	$\hat{\theta}_p$
	0.009000	0.009195	0.008256	0.008415	0.007366	0.006419	0.003743	-0.002876	$\hat{p}_p$
	0.000005	0.000015	0.000011	-0.000012	-0.004159	0.000115	0.003230	0.002263	$\hat{\beta}_{cvm}$
	-0.002246	-0.004731	-0.008341	-0.004387	-0.002177	-0.005818	-0.012754	-0.011819	$\hat{\theta}_{cvm}$
	0.009365	0.008914	0.008259	0.008249	0.007997	0.006678	0.004017	-0.001599	$\hat{p}_{cvm}$
	0.520921	0.508096	0.357149	0.430097	0.424165	0.428637	0.368719	0.403137	$\hat{\beta}_{ad}$
	-1.578636	-1.572513	-1.531119	-1.546045	-1.546783	-1.544725	-1.531580	-1.540284	$\hat{\theta}_{ad}$
	0.377298	0.373518	0.352403	0.358118	0.356877	0.353100	0.342251	0.329843	$\hat{p}_{ad}$



جدول ۳: میانگین توان دوم خطای حقیقیه،  $\hat{\beta}$  و  $\hat{\theta}$  به ازای  $\alpha = 70, \beta = 1.5, \theta = 3$  و  $k = 1$

Table 3: Mean square error of insurance premium,  $\hat{\beta}$  and  $\hat{\theta}$  for  $\alpha = 70, \beta = 1.5, \theta = 3$  and  $k = 1$

80	70	60	50	40	30	20	10	n	برآوردها
43.852004	21.270536	11.588148	6.730986	4.055707	2.486428	1.526584	0.914074		$\hat{\beta}_{mm}$
0.000454	0.000513	0.000594	0.000700	0.000909	0.001299	0.002249	0.006739		$\hat{\theta}_{mm}$
0.000093	0.000090	0.000085	0.000078	0.000069	0.000055	0.000032	0.000003		$\hat{p}_{mm}$
0.000988	0.000928	0.000943	0.000900	0.000952	0.000980	0.001033	0.001006		$\hat{\beta}_{ml}$
0.000024	0.000026	0.000032	0.000039	0.000047	0.000066	0.000101	0.000226		$\hat{\theta}_{ml}$
0.000092	0.000089	0.000084	0.000077	0.000067	0.000053	0.000029	0.000002		$\hat{p}_{ml}$
0.001378	0.001330	0.001389	0.001402	0.001557	0.001763	0.002239	0.003824		$\hat{\beta}_{mix}$
0.000454	0.000513	0.000594	0.000700	0.000909	0.001299	0.002249	0.006739		$\hat{\theta}_{mix}$
0.000141	0.000140	0.000138	0.000134	0.000130	0.000123	0.000108	0.000068		$\hat{p}_{mix}$
0.000003	0.000408	0.001299	0.000739	0.001538	0.079502	0.612501	1.460405		$\hat{\beta}_{ls}$
0.021561	0.021594	0.043122	0.014423	0.035980	0.021606	0.079039	0.086155		$\hat{\theta}_{ls}$
0.000347	0.000342	0.000590	0.000246	0.000488	0.000304	0.000945	0.000994		$\hat{p}_{ls}$
0.002323	0.002343	0.070741	2.139438	1.274108	2.135238	3.111170	1.061570		$\hat{\beta}_{wls}$
0.000029	0.000030	0.000037	0.000046	0.007252	0.021644	0.043131	0.079042		$\hat{\theta}_{wls}$
0.000092	0.000088	0.000083	0.000076	0.000150	0.000303	0.000529	0.000971		$\hat{p}_{wls}$
0.000007	0.000043	0.000024	0.004645	0.002897	3.074393	0.004270	1.039910		$\hat{\beta}_p$
0.028742	0.007230	0.043148	0.007235	0.028801	0.021628	0.050318	0.043388		$\hat{\theta}_p$
0.000429	0.000171	0.000588	0.000158	0.000399	0.000299	0.000607	0.000503		$\hat{p}_p$
0.000000	0.000002	0.000001	0.000001	0.000094	0.000352	0.000710	0.004011		$\hat{\beta}_{cvm}$
0.014401	0.028755	0.050319	0.028763	0.014417	0.035954	0.079013	0.078996		$\hat{\theta}_{cvm}$
0.000262	0.000427	0.000675	0.000415	0.000237	0.000475	0.000949	0.000909		$\hat{p}_{cvm}$
0.300706	0.286001	0.139618	0.200382	0.194808	0.199080	0.149837	0.177763		$\hat{\beta}_{ad}$
2.501719	2.481907	2.350353	2.396420	2.398889	2.392302	2.352259	2.379324		$\hat{\theta}_{ad}$
0.144818	0.141826	0.125477	0.129700	0.128803	0.126068	0.118449	0.110076		$\hat{p}_{ad}$

جدول ۴: واریانس تعمیم یافته  $\hat{\beta}$  و  $\hat{\theta}$  به ازای  $\alpha = 70, \beta = 1.5, \theta = 3$  و  $k = 3$

Table 4: Generalized variance of  $\hat{\beta}$  and  $\hat{\theta}$  for  $\alpha = 70, \beta = 1.5, \theta = 3$  and  $k = 3$

80	70	60	50	40	30	20	10	n	برآوردها
-1.0912E08	-1.897E08	-1.9171E08	-1.5905E08	1.6600E-08	2.0900E-08	2.6200E-08	4.4200E-08		$GV_{mm}$
-5.2187E09	-7.4700E09	-7.5899E09	-7.3745E09	8.4200E-09	1.3900E-08	1.9300E-08	5.3100E-08		$GV_{ml}$
-1.0017E08	-1.0187E08	-1.1286E08	-1.0466E08	1.2600E-08	1.9500E-08	2.7200E-08	5.5000E-08		$GV_{mix}$
-0.0817E08	-1.1317E08	-1.4030E08	-2.6497E05	5.1500E-06	2.4500E-08	4.9900E-05	5.4200E-05		$GV_{ls}$
-1.0313E08	-1.1613E08	-1.2515E08	-1.3969E08	1.7200E-08	8.6000E-05	5.4400E-06	5.6300E-06		$GV_{wls}$
-1.1153E08	-1.1456E08	-1.1821E08	-1.0064E05	5.7900E-03	5.9181E-03	1.0552E-04	8.6500E-06		$GV_p$
-2.1246E09	-1.0146E08	-4.9640E05	-4.6588E06	1.3800E-05	7.5900E-06	9.7800E-06	9.3300E-08		$GV_{cvm}$
-1.0013E05	-1.0878E05	-1.0034E05	-7.1534E05	7.4300E-06	8.5700E-06	9.1200E-06	7.3300E-06		$GV_{ad}$

اریبی کمتری است و بعد از برآوردگر صدکی، برآوردگرهای کرامر\_فون\_میسز و کمترین توان دوم خطا دارای کمترین اریبی اند و برای حجم نمونه بزرگ به ترتیب برآوردگرهای صدکی، کرامر\_فون\_میسز و کمترین توان دوم خطای حقیقیه، دارای کمترین اریبی اند و عملکرد بهتری نسبت به سایر برآوردگرها دارند، از لحاظ اریبی (بتا) برآوردگر کرامر\_فون\_میسز دارای کمترین اریبی و بهترین برآوردگر است. برای حجم نمونه بزرگ از بتا برآوردگرهای کرامر\_فون\_میسز، صدکی و کمترین توان دوم خطا به ترتیب دارای کمترین اریبی اند، و از لحاظ اریبی (تتا) برآوردگر ماکسیمم درست‌نمایی و صدکی به ترتیب دارای کمترین

که برآوردگرها سازگارند و بررسی‌ها نشان‌دهنده آن است که برآوردگرهای ماکسیمم درست‌نمایی و آمیخته گشتاوری و ماکسیمم درست‌نمایی با افزایش حجم نمونه دارای عملکرد بهتری نسبت به دیگر برآوردگرها هستند، چون دارای واریانس تعمیم یافته کمتری در مقایسه با سایر برآوردگرها هستند. برای حجم نمونه بزرگ به ترتیب، برآوردگرهای کرامر\_فون\_میسز، ماکسیمم درست‌نمایی، صدکی و آمیخته ماکسیمم درست‌نمایی و گشتاوری دارای عملکرد بهتری در مقایسه با سایر برآوردگرها هستند. (ب) برای  $n = 10, 20, 30$  و  $k = 1$  حق‌بیمه برآوردگر صدکی در مقایسه با سایر برآوردگرها عملکرد بهتری دارد، زیرا دارای

جدول ۵: اربیبی حق بیمه،  $\hat{\theta}$  و  $\hat{\beta}$  به ازای  $\alpha = 70, \beta = 1.5, \theta = 3$  و  $k = 3$

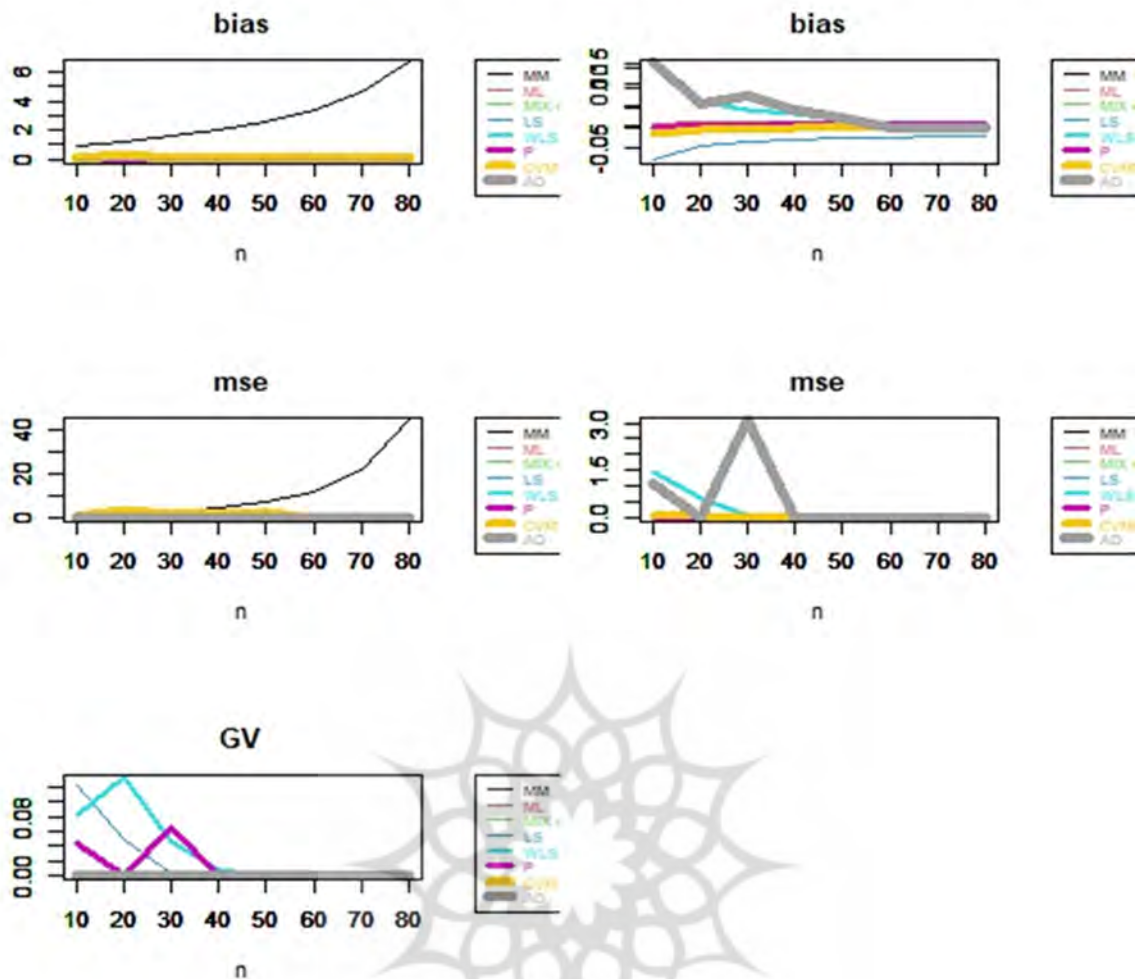
Table 5: Bias of premium  $\hat{\beta}$  and  $\hat{\theta}$  for  $\alpha = 70, \beta = 1.5, \theta = 3$  and  $k = 3$

								n
80	70	60	50	40	30	20	10	برآورد
1.470136	1.342069	1.234950	1.136308	1.043646	0.955707	0.870904	0.778491	$\hat{\beta}_{mm}$
-0.039010	-0.04101	-0.04544	-0.053053	-0.063128	-0.080208	-0.115863	-0.227768	$\hat{\theta}_{mm}$
0.027328	0.028338	0.027505	0.026485	0.024802	-0.000030	-0.005519	-0.022033	$\hat{p}_{mm}$
0.0056801	0.004580	0.003757	0.003776	0.003150	0.002876	0.002904	0.001890	$\hat{\beta}_{ml}$
0.000132	0.000503	0.000985	0.000281	0.000525	0.001172	0.001311	0.003346	$\hat{\theta}_{ml}$
0.029350	0.028260	0.027422	0.026391	0.024711	-0.000163	-0.005706	-0.022443	$\hat{p}_{ml}$
0.023550	0.025690	0.027387	0.030992	0.035729	0.044763	0.063962	0.127084	$\hat{\beta}_{mix}$
-0.039012	-0.04101	-0.04544	-0.053053	-0.063128	-0.080208	-0.115863	-0.227768	$\hat{\theta}_{mix}$
0.033010	0.032700	0.032358	0.032016	0.031339	0.008130	0.005710	-0.003167	$\hat{p}_{mix}$
0.015698	0.014598	0.013217	0.012882	0.012238	0.011512	0.011847	0.011612	$\hat{\beta}_{ls}$
-0.000105	-0.00020	0.000413	-0.015268	-0.003157	0.000107	-0.015360	-0.011244	$\hat{\theta}_{ls}$
0.031235	0.028275	0.027416	0.024753	0.024364	-0.000174	-0.007286	-0.023127	$\hat{p}_{ls}$
0.015267	0.014258	0.012579	0.012096	0.011509	0.009486	0.010852	0.010747	$\hat{\beta}_{wls}$
0.0000520	0.000022	0.000534	-0.000191	-0.000018	-0.002696	-0.003207	-0.004972	$\hat{\theta}_{wls}$
0.029147	0.028253	0.027408	0.026370	0.024679	-0.000525	-0.006025	-0.022543	$\hat{p}_{wls}$
0.015311	0.014612	0.013145	0.012780	0.028513	0.033665	0.014419	0.009138	$\hat{\beta}_p$
0.0009523	0.001034	0.001872	-0.004512	-0.000852	-0.002633	0.001883	0.009463	$\hat{\theta}_p$
0.029061	0.028141	0.027259	0.025535	0.024100	-0.001222	-0.006604	-0.023883	$\hat{p}_p$
0.003852	0.002779	0.001828	0.002216	0.001621	0.000954	0.001558	0.001181	$\hat{\beta}_{cvm}$
-0.000052	-0.00004	-0.02927	-0.003062	-0.008789	-0.002352	-0.005404	0.001550	$\hat{\theta}_{cvm}$
0.032144	0.028331	0.024214	0.026119	0.023796	-0.000395	-0.006226	-0.022229	$\hat{p}_{cvm}$
0.613400	0.512420	0.413913	0.344350	0.457913	0.440960	0.403016	0.349001	$\hat{\beta}_{ad}$
-1.593504	-1.57290	-1.54372	-1.532222	1.549700	-1.546344	-1.542829	-1.535349	$\hat{\theta}_{ad}$
0.386591	0.385821	0.369460	0.361285	0.363760	0.331751	0.314774	0.265815	$\hat{p}_{ad}$

جدول ۶: میانگین توان دوم خطای حق بیمه،  $\hat{\theta}$  و  $\hat{\beta}$  به ازای  $\alpha = 70, \beta = 1.5, \theta = 3$  و  $k = 3$

Table 6: Mean square error of insurance premium,  $\hat{\beta}$  and  $\hat{\theta}$  for  $\alpha = 70, \beta = 1.5, \theta = 3$  and  $k = 3$

								n
80	70	60	50	40	30	20	10	برآوردها
2.01253	1.801670	1.525489	1.291486	1.089438	0.913589	0.758654	0.606224	$\hat{\beta}_{mm}$
0.001360	0.001730	0.002116	0.002869	0.004055	0.006531	0.013572	0.052149	$\hat{\theta}_{mm}$
0.000856	0.000804	0.000757	0.000702	0.000616	0.000001	0.000032	0.000488	$\hat{p}_{mm}$
0.000275	0.000240	0.000244	0.000210	0.000197	0.000220	0.000210	0.000262	$\hat{\beta}_{ml}$
0.000027	0.000031	0.000035	0.000040	0.000048	0.000075	0.000112	0.000306	$\hat{\theta}_{ml}$
0.000851	0.000799	0.000752	0.000697	0.000611	0.000001	0.000034	0.000506	$\hat{p}_{ml}$
0.000761	0.000881	0.000977	0.001158	0.001466	0.002216	0.004294	0.016399	$\hat{\beta}_{mix}$
0.001456	0.001730	0.002116	0.002869	0.004055	0.006531	0.013572	0.052149	$\hat{\theta}_{mix}$
0.001092	0.001070	0.001048	0.001026	0.000983	0.000067	0.000034	0.000013	$\hat{p}_{mix}$
0.000538	0.000528	0.000512	0.000463	0.000435	0.000424	0.000703	0.001978	$\hat{\beta}_{ls}$
0.000032	0.000038	0.000042	0.089732	0.018042	0.000094	0.089638	0.053781	$\hat{\theta}_{ls}$
0.000852	0.000800	0.000752	0.001664	0.000803	0.000001	0.001042	0.001042	$\hat{p}_{ls}$
0.000652	0.000558	0.000480	0.000462	0.000446	0.004886	0.000486	0.001877	$\hat{\beta}_{wls}$
0.000029	0.000034	0.000039	0.000046	0.000058	0.018033	0.018027	0.018162	$\hat{\theta}_{wls}$
0.000852	0.000799	0.000752	0.000696	0.000610	0.000205	0.000232	0.000669	$\hat{p}_{wls}$
0.000588	0.000529	0.000456	0.000444	0.321345	0.166055	0.006054	0.002588	$\hat{\beta}_p$
0.000031	0.000039	0.000046	0.035922	0.018052	0.035891	0.018055	0.018468	$\hat{\theta}_p$
0.000840	0.000792	0.000744	0.001074	0.000790	0.000414	0.000244	0.000729	$\hat{p}_p$
0.000301	0.000291	0.000281	0.000265	0.000260	0.000421	0.000276	0.000346	$\hat{\beta}_{cvm}$
0.000052	0.000038	0.180193	0.017954	0.053900	0.018066	0.035930	0.000380	$\hat{\theta}_{cvm}$
0.000769	0.000803	0.002709	0.000894	0.001192	0.000206	0.000437	0.000497	$\hat{p}_{cvm}$
0.351410	0.291318	0.185775	0.128892	0.231765	0.211974	0.177168	0.132000	$\hat{\beta}_{ad}$
2.582651	2.483567	2.389717	2.352914	2.408907	2.397471	2.387199	2.362506	$\hat{\theta}_{ad}$
0.164085	0.151088	0.137909	0.131525	0.133899	0.111312	0.100208	0.071215	$\hat{p}_{ad}$



مودار ۱: اریبی، میانگین توان دوم خطا و واریانس تعمیم یافته برآوردگرهای مختلف براساس نتایج شبیه سازی به ازای  $k=1$  و  $\beta=1.5, \theta=3, \alpha=70$   
 Fig. 1: Bias, mean squared error and generalized variance of different estimators based on simulation results for  $\beta=1.5, \theta=3, \alpha=70$  and  $k=1$

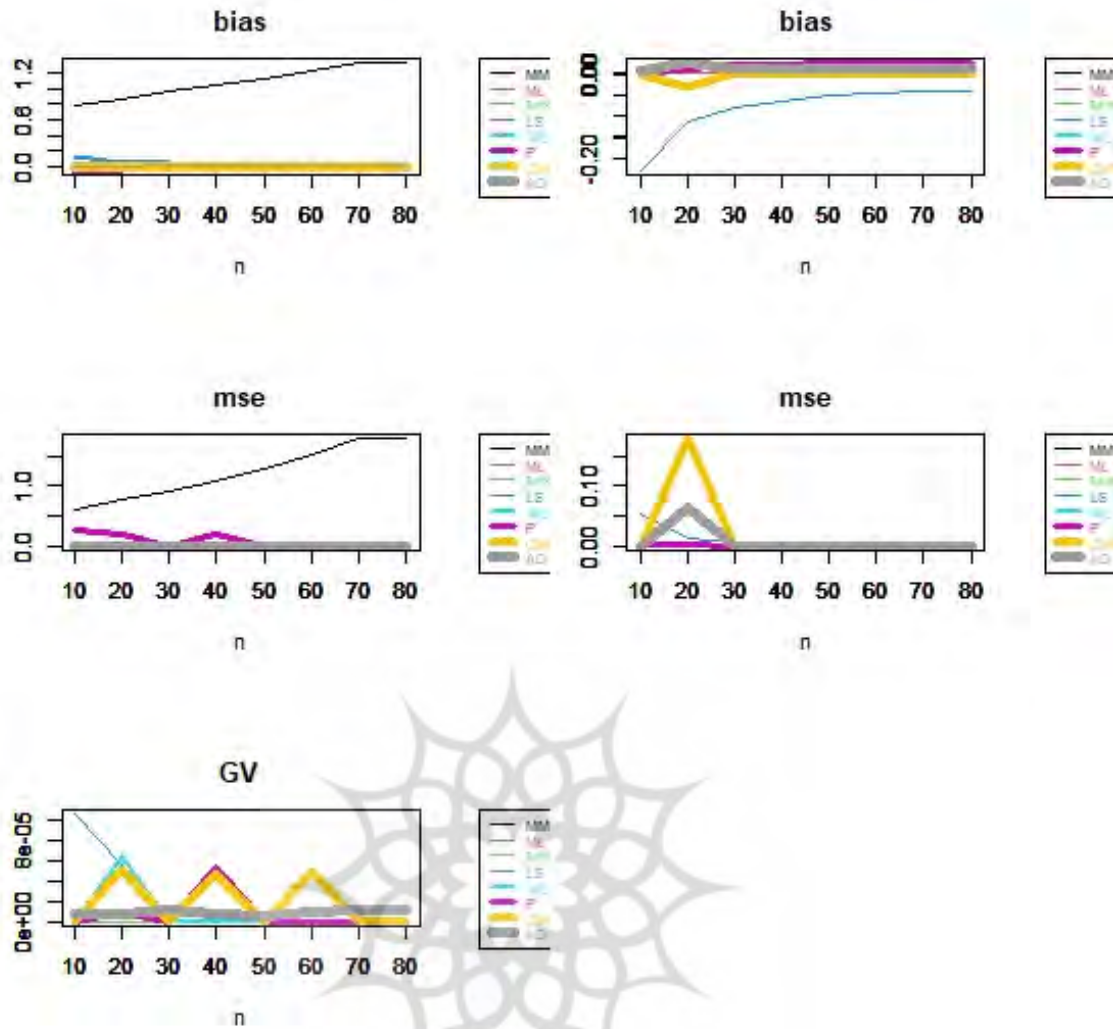
از داده‌های خسارت بیمه آتش سوزی دانمارک گرفته شده است. داده‌های دانمارکی توسط مت ریتگارد از کپنهاگ ارائه شده است (داده‌ها در بسته evir نرم افزار R موجود است). این داده‌ها شامل خسارت‌های بزرگ بیمه آتش سوزی در دانمارک است که از تاریخ پنجشنبه ۳ ژانویه ۱۹۸۰ تا دوشنبه ۳۱ دسامبر ۱۹۹۰ و شامل ۲۱۶۷ مشاهده است. نمونه تصادفی ۳۰ تایی انتخاب شده از این داده‌ها به شرح زیر است

۱.۶۶۴۶۸۵, ۱.۰۲۸۲۹۳۲, ۲۳.۲۸۳۸۵۹, ۱.۰۹۹۴۷۶, ۶.۵۱۱۰۹۱,  
 ۱۰.۱۷۸۰۲۴, ۴.۲۳۳۷۰۰, ۱.۳۰۰۰۰۰, ۱.۱۵۰۴۳۴  
 ۱.۰۶۰۰۰۰, ۱.۰۰۰۰۰۰, ۲.۳۶۷۹۸۷, ۱.۰۰۵۰۰۰, ۱.۲۶۲۶۴۴,  
 ۱.۸۴۱۷۵۳, ۱.۸۶۶۷۴۰, ۱.۳۲۲۲۵۴, ۱.۳۴۹۷۲۲,  
 ۱.۷۵۰۱۵۹, ۱.۶۳۹۳۴۴, ۱.۱۵۳۵۰۵, ۲.۷۸۱۸۴۵, ۲.۸۹۲۹۶۰,  
 ۲.۶۷۰۱۵۷, ۱۶.۴۱۵۲۶۲, ۱.۴۰۸۲۳۱, ۲.۳۴۲۴۴۹,  
 ۱.۷۷۴۶۲۳, ۱.۹۲۸۶۴۰, ۱.۰۰۴۶۳۸.

اریبی هستند و عملکرد بهتری نسبت به سایر برآوردگرها دارند و برای حجم نمونه بزرگ برآوردگر کمترین توان دوم خطا، ماکسیمم درست‌نمایی و صدکی به ترتیب دارای کمترین اریبی برای تتا هستند و دارای عملکرد بهتری نسبت به سایر برآوردگرها هستند. (پ) برای  $n=10, 20, 30$  و  $k=1$  به ترتیب حق بیمه از برآوردگرهای گشتاوری، ماکسیمم درست‌نمایی و آمیخته دارای کمترین میانگین توان دوم خطا هستند و عملکرد بهتری نسبت به سایر برآوردگرها دارند و برای حجم نمونه بزرگ به برآوردگر کمترین توان‌های دوم خطا عملکرد بهتری در مقایسه با سایر برآوردگرها دارد، زیرا میانگین توان دوم خطای کمتری نسبت به سایر برآوردگرها دارد.

#### مثال‌های کاربردی

مثال ۱. در این مقاله به منظور برآورد پارامترهای توزیع وایبول و حق بیمه، یک مثال واقعی ارائه شده است. یک نمونه به اندازه ۳۰



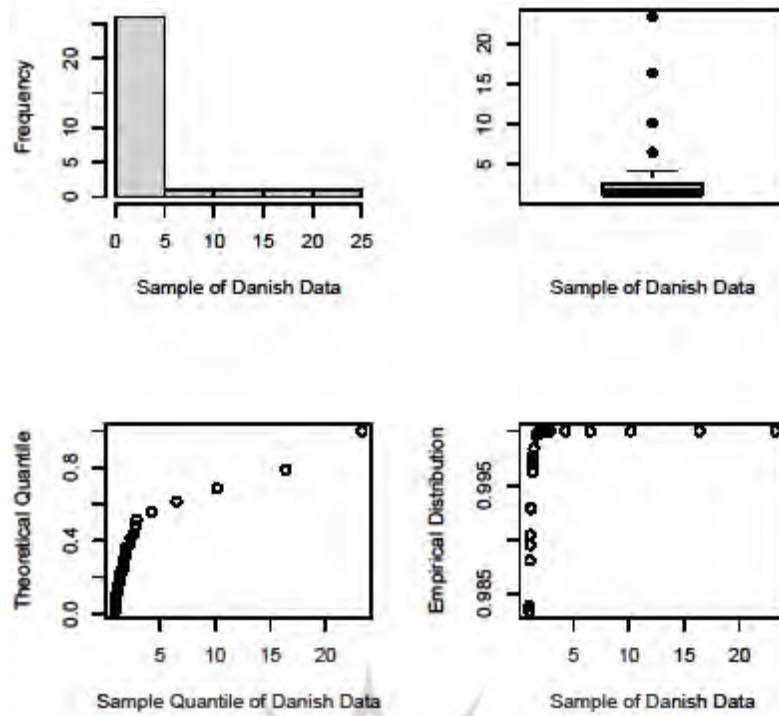
نمودار ۲: اربیبی، میانگین توان دوم خطا و واریانس تعمیم‌یافته برآوردگرهای مختلف براساس نتایج شبیه‌سازی به‌ازای  $\beta = 1.5, \theta = 3, \alpha = 70$  و  $k = 3$   
 Fig. 2: Bias, mean squared error and generalized variance of different estimators based on simulation results for  $\beta = 1.5, \theta = 3, \alpha = 70$  and  $k = 3$

$$\beta = 1.5, \theta = 3, \alpha = 70 \text{ and } k = 3$$

است که توزیع وایبول را در مواجهه با داده‌های دورافتاده برازش دهیم. از طرفی دیگر، تعداد داده‌های دورافتاده ( $k$ )، علاوه بر پارامترهای  $\theta$  و نیز مجهول است. لذا به‌ازای  $k = 1, 2, \dots, 5$  برآورد پارامترهای  $\beta$  و  $\theta$  به روش‌های گوناگون به‌دست آمده و با محاسبه تابع درست‌نمایی به‌ازای  $k$ ‌های مختلف و برآوردگرهای پارامتر، آن  $k$ ‌یی را به‌دست می‌آوریم که تابع درست‌نمایی تجربی را ماکسیمم کند. برآوردگر پارامترها به روش‌های مختلف در **جدول ۷** و **۸** به‌ازای  $k$ ‌های مختلف آمده است. تابع درست‌نمایی تجربی آن نیز به‌ازای  $k$ ‌های مختلف محاسبه شده است که در **جدول ۹** نشان داده شده است. بنابراین، با توجه به مقادیر تابع درست‌نمایی نسبت به  $k$ ، مشخص است که به‌طور مثال در روش گشتاوری، تابع درست‌نمایی وقتی ماکسیمم است که  $k$  برابر ۵ باشد. با قرار دادن مقادیر برآورد پارامترهای  $\beta$  و  $\theta$  در فرمول محاسبه حقیق‌بیمه،

برای اینکه بررسی کنیم داده‌های خسارت از توزیع وایبول تبعیت می‌کند یا خیر، از روش‌های توصیفی شامل رسم نمودارهای جعبه‌ای، بافت‌نگار، چندک-چندک و تابع توزیع تجربی استفاده می‌کنیم. همچنین آزمون نیکویی برازش کلموگروف-اسمیرنوف یک نمونه‌ای برای برازش توزیع وایبول به کار برده می‌شود. نتایج مربوط به روش توصیفی در **نمودار ۳** آمده است. با توجه به نتیجه آزمون کلموگروف-اسمیرنوف و نیز **نمودار ۳**، توزیع متغیر خسارت وایبول است ( $p = 0.081$ ). شایان ذکر است که برآورد پارامترهای این توزیع به روش ماکسیمم درست‌نمایی به‌ترتیب  $\hat{\alpha} = 1.3248$  و  $\hat{\theta} = 2.9001$  است.

چون می‌دانیم که ممکن است داده دورافتاده در داده‌ها وجود داشته باشد، به این دلیل که خسارت‌هایی که به واحدهای مسکونی وارد می‌شود، بعضی اوقات بیش از حد معمول است، لازم



نمودار ۳: بافت‌نگار، نمودار جعبه‌ای، چندک-چندک و تابع توزیع تجربی داده‌های خسارت آتش‌سوزی دانمارک  
 Fig. 3: Histogram, box plot, quantile-quantile and empirical distribution function of Danish fire damage data

جدول ۷: برآوردهای مختلف  $\beta$  برای  $\alpha = 1.3248$  نسبت به  $k$

Table 7: Various estimators of  $\beta$  for  $\alpha = 1.3248$  in respect to  $k$

$\hat{\beta}_{ad}$	$\hat{\beta}_{cvm}$	$\hat{\beta}_p$	$\hat{\beta}_{wls}$	$\hat{\beta}_{ls}$	$\hat{\beta}_{mix}$	$\hat{\beta}_{ml}$	$\hat{\beta}_{mm}$	$k$
2.793501	0.9999996	0.9999997	1.00002044	0.9999995	3.779443	0.1336535	0.1254932	1
2.793501	1.0000236	1.0000209	0.06839348	1.0000285	3.761079	2.8841202	0.2699722	2
2.793501	1.0000243	1.0000214	1.00002152	1.0000293	3.742452	0.1195378	0.4379768	3
2.793501	1.0000247	1.0000183	0.99999969	1.0000297	3.723491	2.9819105	0.6356001	4
2.793501	1.0000255	1.0000225	1.00002267	1.0000305	3.704073	3.0357152	0.8712122	5

جدول ۸: برآوردهای مختلف  $\theta$  برای  $\alpha = 1.3248$  نسبت به  $k$

Table 8: Different estimators of  $\theta$  for  $\alpha = 1.3248$  respect to  $k$

$\hat{\theta}_{ad}$	$\hat{\theta}_{cvm}$	$\hat{\theta}_p$	$\hat{\theta}_{wls}$	$\hat{\theta}_{ls}$	$\hat{\theta}_{mix}$	$\hat{\theta}_{ml}$	$\hat{\theta}_{mm}$	$k$
6.32406	0.4112209	0.4274100	0.3614717	0.4120285	0.1910469	0.3219974	0.1910469	1
6.32406	0.4112198	0.4274089	0.3873936	0.4120272	0.1870536	0.2344257	0.1870536	2
6.32406	0.4112194	0.4274086	0.3614712	0.4120268	0.1830430	0.4966554	0.1830430	3
6.32406	0.4112191	0.4274084	0.3614725	0.4120264	0.1790135	0.2247280	0.1790135	4
6.32406	0.4112187	0.4274079	0.3614706	0.4120259	0.1749633	0.2197206	0.1749633	5

جدول ۹: تابع درست‌نمایی تجربی نسبت به  $k$

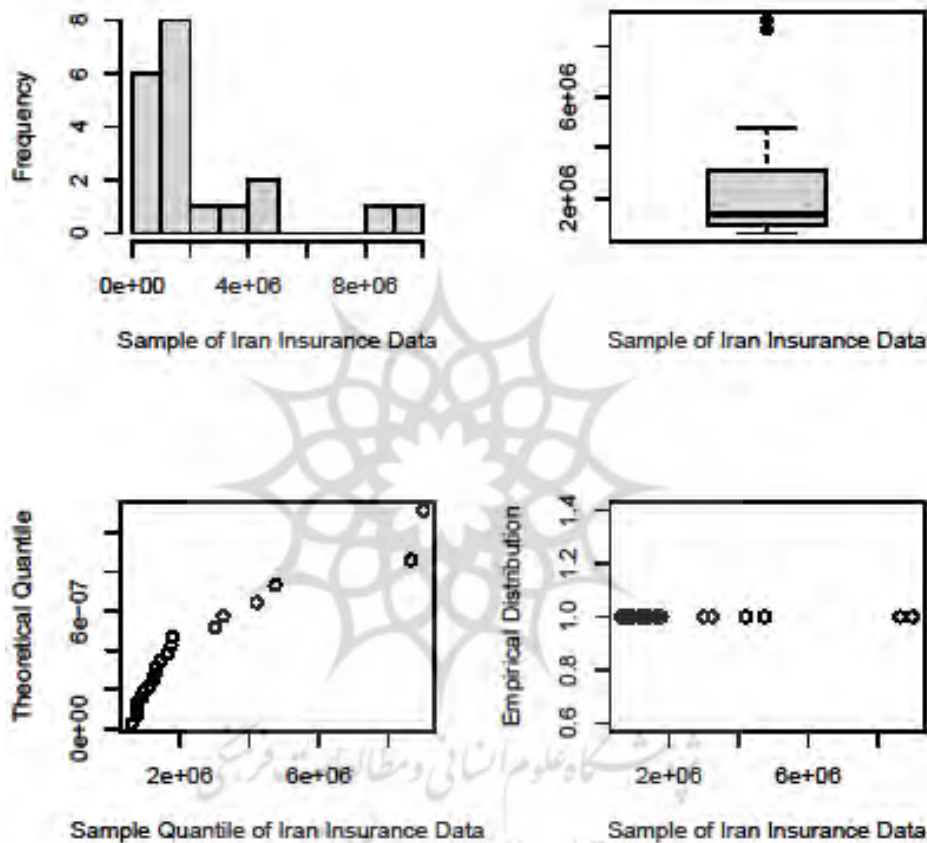
Table 9: Empirical likelihood function with respect to  $k$

$l(\hat{\theta}_{ad}, \hat{\beta}_{ad})$	$l(\hat{\theta}_{cvm}, \hat{\beta}_{cvm})$	$l(\hat{\theta}_p, \hat{\beta}_p)$	$l(\hat{\theta}_{wls}, \hat{\beta}_{wls})$	$l(\hat{\theta}_{ls}, \hat{\beta}_{ls})$	$l(\hat{\theta}_{mix}, \hat{\beta}_{mix})$	$l(\hat{\theta}_{ml}, \hat{\beta}_{ml})$	$l(\hat{\theta}_{mm}, \hat{\beta}_{mm})$	$k$
-2183.302	-76.22000	-77.83384	-71.92480	-76.29815	-67.79172	-61.97735	-67.69032	1
-2215.391	-73.54580	-75.15964	-55.36909	-73.62395	-64.68653	-63.55078	-65.28727	2
-2248.454	-71.31221	-72.92605	-67.01706	-71.39036	-62.01014	-50.28831	-63.75286	3
-2282.949	-69.40267	-71.01651	-65.10756	-69.48082	-59.64774	-58.57895	-62.58078	4
-2318.622	-67.75401	-69.36785	-63.45886	-67.83216	-57.53808	-56.50286	-61.31777	5

جدول ۱۰: مقادیر  $\hat{\beta}_{max}$ ,  $\hat{\theta}_{max}$ ,  $k_{max}$  و حق بیمه خالص برآودگرهای مختلف

Table 10: Values of  $\hat{\beta}_{max}$ ,  $\hat{\theta}_{max}$ ,  $k_{max}$  and net premium of different estimators

$k_{max}$	$\hat{\beta}_{max}$	$\hat{\theta}_{max}$	$\hat{p}$
$k_{mm} = 5$	$\hat{\beta}_{mm} = 0.8712122$	$\hat{\theta}_{mm} = 0.1749633$	$\hat{p}_{mm} = 3.37657$
$k_{ml} = 3$	$\hat{\beta}_{ml} = 0.1195378$	$\hat{\theta}_{ml} = 0.4966554$	$\hat{p}_{ml} = 3.217375$
$k_{mix} = 5$	$\hat{\beta}_{mix} = 3.704073$	$\hat{\theta}_{mix} = 0.1749633$	$\hat{p}_{mix} = 4.619323$
$k_{ls} = 5$	$\hat{\beta}_{ls} = 1.000031$	$\hat{\theta}_{ls} = 0.4120259$	$\hat{p}_{ls} = 2.233271$
$k_{wls} = 2$	$\hat{\beta}_{wls} = 0.06839348$	$\hat{\theta}_{wls} = 0.3873936$	$\hat{p}_{wls} = 4.532246$
$k_p = 5$	$\hat{\beta}_p = 1.000023$	$\hat{\theta}_p = 0.4274079$	$\hat{p}_p = 2.152901$
$k_{cvm} = 5$	$\hat{\beta}_{cvm} = 1.0000255$	$\hat{\theta}_{cvm} = 0.4112187$	$\hat{p}_{cvm} = 2.237657$
$k_{ad} = 1$	$\hat{\beta}_{ad} = 2.793501$	$\hat{\theta}_{ad} = 6.32406$	$\hat{p}_{ad} = 0.1423892$



نمودار ۴: بافت‌نگار، نمودارهای جمعی‌ای، چندک-چندک و تابع توزیع تجربی داده‌های خسارت بیمه ماشین بیمه ایران

Fig. 4: Histogram, box plot, quantile-quantile and empirical distribution function of Iran Insurance car insurance claim data

مثال ۲. در یک شرکت بیمه یکی از خدمات، بیمه وسایل نقلیه است. موضوع خسارت وارده به خودروهای در حال تردد در داخل شهرها را بر اثر تصادف در نظر بگیرید. اگر خودروها دارای بیمه باشند، وقتی تصادفی صورت می‌گیرد شرکت بیمه متعهد است که از طرف راننده مقصر مبلغی را به‌عنوان جبران خسارت پرداخت کند. می‌دانیم در صورتی که خسارت‌های وارده کمتر از ۵۰۰,۰۰۰ ریال مقرون‌به‌صرفه نیست که برای دریافت خسارت از شرکت بیمه اقدام صورت پذیرد. اما اگر در نظر بگیریم که تعدادی از خودروهایی که تصادف کرده‌اند گران‌قیمت‌اند و یا خسارت وارده به آن‌ها شدید

می‌توان برآوردگر حق بیمه را به روش‌های مختلف به‌دست آورد که نتایج آن در جدول ۱۰ آمده است.

در جدول ۱۰ مقادیر ماکسیمم  $k$  برای هر یک از روش‌های برآورد، به همراه برآورد پارامترهای  $\theta$  و  $\beta$  متناظر با آن (تعداد داده‌های دورافتاده) و همچنین حق بیمه خالص ارائه شده است. به‌طور مثال، در روش گشتاوری تعداد داده‌های دورافتاده  $k=5$  برآورد شده است، بنابراین به‌ازای  $\hat{\beta}_{mm} = 0.8712122$ ،  $k=5$  و  $\hat{\theta}_{mm} = 0.1749633$  است و حق بیمه خالص متناظر با آن  $\hat{p}_{mm} = 3.37657$  است.

جدول ۱۱: برآوردگرهای مختلف  $\beta$  برای  $\alpha = 1.372$  نسبت به  $k$

Table 11: Various estimators of  $\beta$  for  $\alpha = 1.372$  respect to  $k$

$\hat{\beta}_{ad}$	$\hat{\beta}_{cvm}$	$\hat{\beta}_p$	$\hat{\beta}_{wls}$	$\hat{\beta}_{ls}$	$\hat{\beta}_{mix}$	$\hat{\beta}_{ml}$	$\hat{\beta}_{mm}$	$k$
5.000085	4.9999666	4.9999742	4.999959	4.999990	3.142297	18.353170	0.4316495	1
5.000085	0.1239608	1.5993283	1.780974	2.124528	3.099228	9.098362	1.1095494	2
5.000085	4.9984942	0.2618328	1.819728	2.150974	3.055664	8.250227	2.3236309	3
5.499902	4.9984942	1.6521680	5.000005	2.187310	3.008829	7.195901	5.1099639	4
5.499902	5.0000020	1.6891253	4.999372	5.000003	2.960044	7.813507	17.9738177	5

جدول ۱۲: برآوردگرهای مختلف  $\theta$  برای  $\alpha = 1.372$  نسبت به  $k$

Table 12: Different estimators of  $\theta$  for  $\alpha = 1.372$  with respect to  $k$

$\hat{\theta}_{ad}$	$\hat{\theta}_{cvm}$	$\hat{\theta}_p$	$\hat{\theta}_{wls}$	$\hat{\theta}_{ls}$	$\hat{\theta}_{mix}$	$\hat{\theta}_{ml}$	$\hat{\theta}_{mm}$	$k$
	4.355500E-07	4.676806E-07	3.957359E-07	4.363649E-07	3.098155E-07	2.796704E-07	3.098155E-07	1
1.376386E-05	5.291053E-07	4.771573E-07	4.052705E-07	4.327867E-07	3.007932E-07	5.062374E-07	3.007932E-07	2
1.376386E-05	3.680291E-07	5.825220E-07	3.910359E-07	4.143691E-07	2.917207E-07	4.106268E-07	2.917207E-07	3
1.376386E-05	3.345015E-07	4.507850E-07	2.941704E-07	3.958547E-07	2.825896E-07	4.777218E-07	2.825896E-07	4
1.376386E-05	3.026454E-07	4.362148E-07	2.655536E-07	3.025644E-07	2.733899E-07	1.541216E-07	07	5

جدول ۱۳: تابع درست‌نمایی تجربی نسبت به

Table 13: Different estimators of  $\theta$  for  $\alpha = 1.372$  with respect to  $k$

$l(\hat{\theta}_{ad}, \hat{\beta}_{ad})$	$l(\hat{\theta}_{cvm}, \hat{\beta}_{cvm})$	$l(\hat{\theta}_p, \hat{\beta}_p)$	$l(\hat{\theta}_{wls}, \hat{\beta}_{wls})$	$l(\hat{\theta}_{ls}, \hat{\beta}_{ls})$	$l(\hat{\theta}_{mix}, \hat{\beta}_{mix})$	$l(\hat{\theta}_{ml}, \hat{\beta}_{ml})$	$l(\hat{\theta}_{mm}, \hat{\beta}_{mm})$	$k$
-3369.084	-312.8188	-313.7861	-311.9326	-312.8407	-311.2676	-315.5855	-311.7098	1
-3566.087	-307.9107	-311.0299	-309.2609	-309.8204	-308.7950	-318.8814	-309.4394	2
-3766.795	-308.2520	-305.2523	-307.1960	-307.6209	-306.7795	-313.5469	-306.8922	3
-Inf	-306.4316	-307.1504	-305.8656	-305.7868	-305.1112	-316.8777	-305.8934	4
-Inf	-304.8807	-305.6489	-304.5234	-304.8793	-303.7345	-308.8130	-324.9688	5

جدول ۱۴: مقادیر  $\hat{\beta}_{max}, \hat{\theta}_{max}, k_{max}$  و حق بیمه خالص برآوردگرهای مختلف

Table 14: Values of  $\hat{\beta}_{max}, \hat{\theta}_{max}, k_{max}$  and net premium of different estimators

$k_{max}$	$\hat{\beta}_{max}$	$\hat{\theta}_{max}$	$\hat{p}$
$k_{mm} = 4$	$\hat{\beta}_{mm} = 5.109964$	$\hat{\theta}_{mm} = 2.825896E - 07$	$\hat{p}_{mm} = 2462050$
$k_{ml} = 5$	$\hat{\beta}_{ml} = 7.813507$	$\hat{\theta}_{ml} = 1.541216E - 07$	$\hat{p}_{ml} = 4639688$
$k_{mix} = 5$	$\hat{\beta}_{mix} = 2.960044$	$\hat{\theta}_{mix} = 2.733899E - 07$	$\hat{p}_{mix} = 2791065$
$k_{ls} = 5$	$\hat{\beta}_{ls} = 5.000003$	$\hat{\theta}_{ls} = 3.025644E - 07$	$\hat{p}_{ls} = 2417798$
$k_{wls} = 5$	$\hat{\beta}_{wls} = 4.999372$	$\hat{\theta}_{wls} = 2.655536E - 07$	$\hat{p}_{wls} = 2754793$
$k_p = 3$	$\hat{\beta}_p = 0.2618328$	$\hat{\theta}_p = 5.82522E - 07$	$\hat{p}_p = 2233599$
$k_{cvm} = 5$	$\hat{\beta}_{cvm} = 5.000002$	$\hat{\theta}_{cvm} = 3.026454E - 07$	$\hat{p}_{cvm} = 2417151$
$k_{ad} = 1$	$\hat{\beta}_{ad} = 5.000085$	$\hat{\theta}_{ad} = 1.376386E - 05$	$\hat{p}_{ad} = 63779.17$

برازش توزیع وایبول به داده‌های خسارت بیمه ایران می‌پردازیم. نتایج مربوط به روش توصیفی در نمودار ۴ آمده است. توزیع وایبول با پارامترهای  $\hat{\alpha} = 1.372$  و  $\hat{\theta} = 2.3327e+6$  که به روش ماکسیمم درست‌نمایی به‌دست آمده است، داده‌های خسارت به‌خوبی برازش می‌یابد ( $p = 0.375$ ).

علاوه بر پارامترهای  $\theta$  و  $\beta$ ، تعداد داده‌های دورافتاده ( $k$ )، نیز مجهول است. لذا به‌ازای  $k = 1, 2, \dots, 5$  برآورد پارامترهای  $\theta$  و  $\beta$  به روش‌های گوناگون به‌دست می‌آید و همانند مثال قبل، ابتدا تعداد داده‌های دورافتاده تعیین می‌شود. برآوردگر پارامترها به روش‌های مختلف در جداول ۱۱ و ۱۲ به‌ازای  $k$ ‌های مختلف گزارش شده است. جدول ۱۳ مقدار تابع درست‌نمایی تجربی به‌ازای  $k$ ‌های

است (طبیعی است که تعداد این خودروها نسبت به عمده خودروها کمتر است) باید حق بیمه را طوری در نظر گرفت که بیمه‌گر دچار خسارت و در نهایت ورشکستگی نشود. در اینجا، تعداد داده‌های دورافتاده  $k$  نیز مجهول است. در این مقاله یک نمونه به اندازه ۲۰ از داده‌های خسارت ماشین بیمه ایران در سال ۲۰۰۹ میلادی را در نظر گرفته‌ایم که به شرح ذیل‌اند.

۷۵۰۰۰۰، ۷۸۰۰۰۰، ۶۳۰۰۰۰، ۱۷۵۰۰۰۰، ۱۴۵۰۰۰۰، ۳۰۰۰۰۰۰، ۸۶۵۰۰۰۰، ۴۲۱۰۰۰۰، ۸۹۰۰۰۰۰، ۹۵۰۰۰۰۰، ۱۲۴۰۰۰۰۰، ۱۸۰۰۰۰۰۰، ۱۶۳۰۰۰۰۰، ۹۰۱۰۰۰۰۰، ۴۷۵۰۰۰۰۰، ۳۲۵۰۰۰۰۰، ۱۱۳۵۰۰۰۰، ۱۳۲۶۰۰۰۰، ۱۲۸۰۰۰۰۰، ۷۶۰۰۰۰۰

همانند مثال قبل به روش توصیفی و استنباطی به بررسی

### تشکر و قدردانی

نویسندگان مقاله از زحمات داوران و سردبیر محترم، که با ارائه نکات موثر و مفید موجب ارتقا علمی و کیفی مقاله شده اند کمال تشکر و قدردانی را دارند و همچنین از دانشگاه فردوسی مشهد که با قرار دادن امکانات مادی و معنوی نقش اساسی در نگارش این مقاله داشته اند نیز سپاسگزاری می کنند.

### تعارض منافع

نویسندگان اعلام می دارند که در مورد انتشار این مقاله تضاد منافع وجود ندارد. علاوه بر این، موضوعات اخلاقی شامل سرعت ادبی، رضایت آگاهانه، سوءرفتار، جعل داده‌ها، انتشار و ارسال مجدد و مکرر توسط نویسندگان رعایت شده است.

### دسترسی آزاد

کپی‌رایت نویسنده (ها) ©2024: این مقاله تحت مجوز بین‌المللی Creative Commons Attribution 4.0 اجازه استفاده، اشتراک‌گذاری، اقتباس، توزیع و تکثیر را در هر رسانه یا قالبی مشروط بر درج نحوه دقیق دسترسی به مجوز CC، منوط به ذکر تغییرات احتمالی بر روی مقاله می‌داند. لذا به استناد مجوز مذکور، درج هرگونه تغییرات در تصاویر، منابع و ارجاعات یا سایر مطالب از اشخاص ثالث در این مقاله باید در این مجوز گنجانده شود، مگر اینکه در راستای اعتبار مقاله به اشکال دیگری مشخص شده باشد. در صورت عدم درج مطالب مذکور و یا استفاده فراتر از مجوز فوق، نویسنده ملزم به دریافت مجوز حق نسخه‌برداری از شخص ثالث است.

به منظور مشاهده مجوز بین‌المللی Creative Commons Attribution 4.0 به نشانی زیر مراجعه شود:

<http://creativecommons.org/licenses/by/4.0>

### یادداشت ناشر

ناشر نشریه پژوهشنامه بیمه با توجه به مرزهای حقوقی در نقشه‌های منتشر شده بی طرف باقی می‌ماند.

مختلف را نشان می‌دهد. در انتها، همانند قبل می‌توان حق بیمه را به روش‌های مختلف برآورد کرد. نتایج آن در جدول ۱۴ آمده است. در جدول ۱۴ مقادیر ماکسیم  $k$  برای هر یک از روش‌های برآورد، به همراه برآورد پارامترهای  $\beta$  و  $\theta$  متناظر با آن و همچنین حق بیمه خالص ارائه شده است.

### نتایج و بحث

برآورد حق بیمه خالص یکی از مهم‌ترین مسائل در بیمه است با توجه به اینکه می‌توان برای توزیع‌های مختلفی حق بیمه را محاسبه کرد، هدف اصلی این پژوهش محاسبه حق بیمه و برآورد پارامترهای توزیع وایبول با استفاده از روش‌های مختلف است. با توجه به وجود داده دورافتاده، لازم است که توزیع وایبول را در مواجهه با داده‌های دورافتاده برازش دهیم، در اینجا، ابتدا پارامترهای توزیع وایبول به روش‌های گوناگون برآورد و کارایی آن‌ها با یکدیگر مقایسه شد که به‌طور کلی برآوردگرهای ماکسیمم درست‌نمایی و آمیخته گشتاوری و ماکسیمم درست‌نمایی برای حجم نمونه کم به‌ازای پارامترهای مختلف کارا تر بودند و برای حجم نمونه بزرگ برآوردگر کمترین توان‌های دوم خطا عملکرد بهتری در مقایسه با سایر برآوردگرها دارد، زیرا میانگین توان دوم خطای کمتری نسبت به سایر برآوردگرها دارد و به‌ازای پارامترهای مختلف کارا تر بود. لذا بر این اساس، حق بیمه خالص برآورد و گزارش شد. برای بررسی وضعیت این برآوردگرها در مسائل واقعی دو مثال از داده‌های بیمه خسارت آتش‌سوزی دانمارک و بیمه خودروی شرکت بیمه ایران انتخاب و ضمن برآورد پارامترهای توزیع وایبول در مواجهه با داده‌های دورافتاده، حق بیمه خالص برآورد شد.

### مشارکت نویسندگان

ابراهیم سهرابی: جمع‌آوری مطالعات مرتبط، جمع‌آوری و اخذ داده‌ها، تجزیه و تحلیل و تفسیر داده‌ها، تحلیل آماری. مهدی جباری نوقایی: ایده و مطالب نظری، تجزیه و تحلیل و تفسیر داده‌ها، تجزیه و تحلیل و تفسیر داده‌ها، تحلیل آماری و نظارت و سرپرستی.



## منابع

- Abubakar, H.; Muhammad Sabri, S.R., (2021). A simulation study on modified Weibull distribution for modelling of investment return. *J. Sci. Technol.*, 29(4): 2767-2790 **(24 Pages)**.
- Alyami, S.A.; Elbatal, I.; Alotaibi, N.; Almetwally, E.M.; Okasha, H.M.; Elgarhy, M., (2022). Topp–Leone modified Weibull model: Theory and applications to medical and engineering data. *J. Appl. Sci.*, 12(20).
- Anderson, T.W.; Darling, D.A., (1954). A test of goodness-of-fit. *J. Am. Stat. Assoc.*, 49(268): 765-769 **(5 Pages)**.
- Ancombe, F.J., (1960). Rejection of outliers. *J. Technometrics.*, 2(2): 123-147 **(25 Pages)**.
- Barnett, V.; Lewis, T., (1984). Outliers in statistical data. *J. Biom.*, 30(7): 866-867 **(2 Pages)**.
- Chaturvedi, A.; Pati, M.; Tomer, S.K., (2014). Robust bayesian analysis of Weibull failure model. *J. Metron.*, 72: 77-95 **(19 Pages)**.
- Dixit, U.J., (1987). Characterization of the Gamma distribution in the presence of K outliers., 4: 54-59 **(6 Pages)**.
- Dixit, U.J., (1989). Estimation of parameters of the Gamma distribution in the presence of outliers., 18(8): 3071-3085 **(15 Pages)**.
- Dixit, U.J.; Jabbari Nooghabi, M., (2011). Efficient estimation of the parameters of the Pareto distribution in the presence of outliers. *J. Commun. Stat. Appl. Methods.*, 18(6): 817-835 **(19 Pages)**.
- Fang, L.; Barmalzan, G.; Ling, J., (2016). Dispersive order of lifetimes of series systems in multiple-outlier Weibull models. *J. Syst. Sci. Complexity.*, 29(6): 1693-1702 **(10 Pages)**.
- Gupta, P.K.; Singh, A.K., (2017). Classical and bayesian estimation of Weibull distribution in the presence of outliers. , 4(1): 1-9 **(9 Pages)**.
- Hawkins, D.M., (1980). Identification of outliers.
- Johnson, N.L.; Kotz, S.; Balakrishnan, N., (1995). Continuous univariate distribution., 2.
- Kale, B.K.; Sinha, S.K., (1971). Estimation of expected life in the presence of an outlier observation. *J. Technometrics.*, 13(4): 755-759 **(5 Pages)**.
- Kao, J.H.K., (1958). Computer methods for estimating Weibull parameters in reliability studies., 13: 15-22 **(8 Pages)**.
- Kao, J.H.K., (1959). A graphical estimation of mixed Weibull parameters in life testing electron tube. *J. Technometrics.*, 1(4): 389-407 **(19 Pages)**.
- Kendal, M.G.; Buckland, W.R., (1957). A dictionary of statistical terms.
- Lieblein, J.; Zelen, M., (1956). Statistical investigation of the fatigue life of deep groove ball bearings. *J. Res. Natl. Bur. Stand.*, 57(5): 273-316 **(44 Pages)**.
- Liu, X.; Ahmad, Z.; Gemeay, A.M.; Abdulrahman, A.T.; Hafez, E.H.; Khalil, N., (2021). Modeling the survival times of the COVID-19 patients with a new statistical model: A case study from China., 16(7).
- Pettit, L.I., (1988). Bayes methods for outliers in exponential samples. *J. R. Stat. Soc.*, 50(3): 371-380 **(10 Pages)**.
- Pettitt, A.N., (1976). A two-sample Anderson-Darling rank statistic. *J. Biom.*, 63(1): 161-168 **(8 Pages)**.
- Read, R.R., (1981). Representation of certain covariance matrices with application to asymptotic efficiency. *J. Am. Stat. Assoc.*, 76(373): 148-154 **(7 Pages)**.
- Sindhu, T.N.; Atangana, A., (2021). Reliability analysis incorporating exponentiated inverse Weibull distribution and inverse power law. *J. Qual. Reliab. Eng. Int.*, 37(6): 2399-2422 **(24 Pages)**.
- Singh, B.; Rathi, S.; Kumar, S., (2013). Inferential statistics on the dynamic system model with time-dependent failure-rate. *J. Stat. Comput. Simul.*, 83(1): 1-24 **(24 Pages)**.
- Stephens, M.A., (1974). EDF statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc.*, 69(347): 730-737 **(8 Pages)**.
- Strzelecki, P., (2021). Determination of fatigue life for low probability of failure for different stress levels using 3-parameter Weibull distribution. *J. Fatigue.*, 145.
- Weibull, W., (1951). A statistical distribution function of wide applicability. *J. Appl. Mech.*
- Young, V.R., (2006). Premium principles in encyclopedia of actuarial science.

AUTHOR(S) BIOSKETCHES	معرفی نویسندگان
<p>ابراهیم سهرابی، دانشجوی دکتری آمار، گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، مشهد، ایران</p> <ul style="list-style-type: none"><li>Email: <a href="mailto:sohrabi.ebrahim@mail.um.ac.ir">sohrabi.ebrahim@mail.um.ac.ir</a></li><li>ORCID: 0009-0000-9737-0021</li><li>Homepage: <a href="https://statistics.um.ac.ir/index.php/fa/">https://statistics.um.ac.ir/index.php/fa/</a></li></ul>	<p>مهدی جباری نوقابی، دانشیار گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، مشهد، ایران</p> <ul style="list-style-type: none"><li>Email: <a href="mailto:jabbarinm@um.ac.ir">jabbarinm@um.ac.ir</a></li><li>ORCID: 0000-0002-5636-2209</li><li>Homepage: <a href="https://mathstat.um.ac.ir/index.php/fa/">https://mathstat.um.ac.ir/index.php/fa/</a></li></ul>

HOW TO CITE THIS ARTICLE	
<p>Sohrabi, E.; Jabbari Nooghabi, M., (2024). Estimation of the parameters of Weibull distribution and net premium in the presence of outliers. <i>Iran. J. Insur. Res.</i>, 13(1): 43-60.</p> <p>DOI: 10.22056/ijir.2024.01.04</p> <p>URL: <a href="https://ijir.irc.ac.ir/article_160310.html?lang=en">https://ijir.irc.ac.ir/article_160310.html?lang=en</a></p>	