Kharazmi University

Mathematical Research
Year 2026, Volume 12, Issue 1, pp. 78–91

Print ISSN: 2588-2546
Online ISSN: 2588-2554
DOI: xxxx

On Bayesian Prediction of k-Record Values via Record Ranked Set Sampling Scheme in a Rayleigh Distribution

Ehsan Golzade Gervi[(1)][1], Simin mansouri Boroujeni[(2)] and Kazem Fayyaz Heidari [(3)]

[(1),(2),(3)] Department of Statistics, Payame Noor University, Tehran, Iran.

Received: 8 February 2025 Accepted: 14 April 2026 Published online: 12 June 2026

Abstract:

Introduction

There are some tests where have been done sequentially, and only record values are observed. The record values have been used in applied examinations, such as quality control, hydrology, economics, climatology and etc. This is why many authors attracted to study on it. An observation X_j is an upper record value if its value exceeds all previous observations. Predicting future of k-record values as an influencing factor is one of the important aims in real life. In this article, two sample Bayesian prediction has been investigated via an upper record ranked set sampling scheme in a Rayleigh model. The record ranked set sampling scheme as a substitute method for producing record, has been formally suggested by Salehi and Ahmadi (2014). After a brief introduction to the record ranked set sampling scheme, the first aim of this article is to study the behavior of Bayesian point predictions of k-records from a future sequence based on an upper record ranked set sampling as an informative sample with respect to both squared error symmetric loss function and linear exponential (linex) asymmetric loss function for different values of α and β. We have compared the Bayesian point predictors in terms of the mean square prediction errors. Our second purpose of this article is to study the Bayesian prediction intervals for k-record in this scheme. Evaluation criteria such as the expected lengths and coverage probabilities are employed to consider the performance of this Bayesian prediction intervals.

Material and Methods

In this article, by considering a Rayleigh model, the record ranked set sampling scheme is used to comparing the Bayesian predictions for k-record values. In order to monitor the performance of the Bayesian k-record predictors, a Monte Carlo simulation is used. We consider $m = k = 1, 2, 3$. Also, the values of the a in linex loss function is considered to be -1, 1. The value of λ is considered to be 0.042 in Rayleigh model. We consider different values (1.5, 0.5 and 0) for α and β. We assumed that the case $(\alpha = \beta = 0)$ for non-informative prior. Next, an application example (related to wind speed) on a real data set for six years 2015 to 2020 of Brunswick new Province and Bas Caraquej Station for government Canada is presented to compare and show the applicability of the article's results. The given data are obtainable at IOC website. The p-value of the goodness of fit test Anderson-Darling of the Rayleigh distribution, is 0.971

[1]Corresponding author
E-mail addresses:(Ehsan Golzade Gervi)e_golzade_g@pnu.ac.ir, (Simin mansouri Boroujeni)s_mansouri@pnu.ac.ir, (Kazem Fayyaz Heidari)fayyaz@pnu.ac.ir

which supports the suitability of the fitting Rayleigh model on real data set. From this data set, it can be shown that $r = (3.05, 6.94, 6.39, 3.61, 6.39, 6.67)^T$, with $n = 6$. Data were analyzed by R software.

Results and discussion

In this article, k-record values predictors under the squared error loss function performs better than the linear exponential loss functions. The Bayesian k-record predictors based on $(\alpha = \beta = 1.5)$ are better than the others in terms of mean square prediction errors. The mean square prediction errors are increasing in m. Moreover, when m increases, the coverage probability become closer to coverage percentage (0.95). The expected lengths under the highest posterior density method are smaller than those of survival method and the coverage probabilities of the highest posterior density method is more than survival method. These Bayesian intervals improve which was expected as the α and β increase jointly. Also, the Bayesian prediction intervals for k-record based on $(\alpha = \beta = 1.5)$ and $(\alpha = \beta = 0.5)$ as compared to the Bayesian prediction intervals based on non-informative prior, performances very well, in terms of coverage probability and expected length. The results in the real data part confirm the results of the simulation part.

Conclusion

This paper contributes to the field of statistical analysis by emphasizing the appropriate sampling scheme to decrease the cost, increase efficiency and introducing the RRSS scheme. In this article, we have considered the behavior of predictors of future k-record via record ranked set sampling scheme in a Rayleigh model. Results showed that the Bayesian predictors of k-record under the squared error loss function performs better than the Bayes predictors under the linear exponential loss function. Also, the performance highest posterior density method are superior to survival method. Finally, using a record ranked set sampling scheme for Bayesian prediction of k-record is recommended if the conditions are met for it.

Keywords: Bayesian point prediction, Bayesian interval prediction, k-Record values, Bayesian predictive density, Record ranked set sampling.

©2026 Kharazmi University, Tehran, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution-NonComertial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

پژوهش‌های ریاضی
سال ۱۴۰۵، دوره ۱۲، شماره ۱، ص ۷۸–۹۱
شاپا چاپی: ۲۵۸۸-۲۵۴۶
شاپا الکترونیکی: ۲۲۸۸-۲۵۵۴
DOI: xxxx

درباره پیشگویی بیزی مقادیر k_رکورد بر اساس طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی در توزیع رایلی

احسان گل‌زاده گروی [(۱)][1]، سیمین منصوری بروجنی[(۲)] و کاظم فیاض حیدری[(۳)]

[(۱),(۲),(۳)] گروه آمار، دانشگاه پیام نور، تهران، ایران.

تاریخ دریافت: ۱۴۰۳/۱۱/۱۹ تاریخ پذیرش: ۱۴۰۵/۱/۲۵ تاریخ انتشار: ۱۴۰۵/۳/۲۲

چکیده: در این مقاله، با در نظر گرفتن توزیع رایلی، به پیشگویی بیزی مقادیر k_رکورد بالا، زمانی که داده‌ها به شیوه طرح نمونه‌گیری مجمو عه رتبه‌دار رکوردی جمع‌آوری شده باشد، پرداخته شده است. در ابتدا به پیشگویی های نقطه‌ای تحت تابع زیان متقارن توان دوم خطا و تابع زیان نامتقارن لاینکس و محاسبه میانگین توان دوم خطا پیشگویی آنها پرداخته می‌شود. سپس، برای ارزیابی بازه‌های پیشگویی به‌دست‌آمده به روش تابع بقا و روش HPD، متوسط طول بازه و احتمال پوشش ارائه می‌گردد. در پایان، با یک مطالعه شبیه‌سازی و ارائه یک مثال کاربردی از داده‌های واقعی مرتبط با میزان سرعت وزش باد، مقایسه نتایج و عملکرد نشان داده شده است.

واژه‌های کلیدی: پیشگویی بیزی نقطه‌ای، پیشگویی بیزی بازه‌ای، مقادیر k_رکورد، چگالی بیزی پیشگوی، نمونه‌گیری مجمو عه رتبه‌دار رکوردی.

رده‌بندی ریاضی: 62G30 Primary; 62G25; 62C12

۱ مقدمه

مفهوم پیشگویی[2] یکی از مفاهیم ریشه‌دار است که تا به امروز دست‌مایه کار بسیاری از آماردانان بوده است. در مسائل پیشگویی، پژوهشگر علاقه‌مند است با استفاده از پیشامد E که به وقوع پیوسته است، پیشامد رخ نداده‌ای مثل F را پیشگویی کند. پیشگویی، در تصمیم‌گیری وضع آینده سیستم‌ها و تحلیل آماری آن نقش ویژه‌ای ایفا می‌کند که به‌صورت استنباط مقادیر متغیرهای تصادفی مجهول آینده یا توابعی از آن متغیرهای تصادفی تعریف می‌شود. به‌طور خاص، پیشگویی و توانایی حدس‌زدن جنبه‌های غیرتعیینی حوادث قبل از تصمیم‌گیری با استفاده از اطلاعات

[1] نویسنده مسئول مقاله
fayyaz@pnu.ac.ir, s_mansouri@pnu.ac.ir, e_golzade_g@pnu.ac.ir

[2] Prediction

موجود در نمونه از دیدگاه بیزی، به سوژهٔ مورد علاقه‌ای برای محققان تبدیل شده است. دیدگاه بیزی این امکان را به پژوهشگر می‌دهد تا اطلاعات قبلی را در تحلیل مسائل پیش رو وارد کند. از جمله پژوهش‌های ارزشمندی که در مبحث پیشگویی به ثبت رسیده است، می‌توان به تحقیقات در [۱۰، ۱۲، ۱۳، ۱۵، ۱۷، ۱۸، ۲۱، ۲۲، ۲۶، ۳۰، ۳۱] اشاره کرد. در این مطالعه، پیشگویی بیزی دونمونه‌ای در نظر گرفته شده است. در پیشگویی دو نمونه ای، فرض می‌کنیم داده‌های یک نمونه یا یک دنباله از متغیرهای تصادفی را در اختیار داریم و در صدد هستیم تا متغیرهای تصادفی آینده را از یک نمونه یا دنباله دیگر پیشگویی کنیم. در این روش، متغیرهایی که علاقه‌مند به پیشگویی آن‌ها در آینده هستیم مستقل از متغیرهای مشاهده شده است. این مقاله با در نظرگرفتن توزیع رایلی، به مساله پیشگویی بیزی از نوع نقطه‌ای و بازه‌ای برای مقادیر k‌ـ رکورد بر اساس داده‌های حاصل از طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی پرداخته است. ما کار خود را با اِرائه تعریف مقادیر رکوردی در زیر آغاز می‌کنیم:

تعریف ۱.۱ ([۱۱]). *فرض کنید $\{X_i, i \geq ۱\}$ یک دنباله نامتناهی از متغیرهای تصادفی مستقل و هم‌توزیع با تابع توزیع تجمعی مشترک $F(x;\theta)$ و تابع چگالی احتمال $f(x;\theta)$ باشد، در این صورت مشاهده‌ی X_j یک رکورد بالا است، هرگاه با احتمال یک، برای همه مقادیر i که $j > i$، $X_j > X_i$ برقرار باشد. به عبارت دیگر، برای اینکه یک رکورد جدید رخ دهد باید بتوان آن را با مشاهدات قبلی مقایسه نمود. گوئیم مشاهده X_j یک رکورد بالا است، هرگاه از تمام مشاهدات ما قبل خود بزرگتر باشد. به شیوه ای مشابه، رکوردهای پایین نیز قابل تعریف هستند. امروزه داده‌های رکوردی دارای کاربردهای وسیعی به ویژه در اقتصاد، زلزله نگاری، مهندسی، پزشکی، کنترل کیفیت، آب شناسی، استخراج نفت و معادن، قابلیت اطمینان، هوا و فضا، تحلیل داده‌های طول عمر و هواشناسی می‌باشد. برخی از این اطلاعات درباره رکوردها را در منابع کلیدی [۹، ۱۱، ۱۹، ۲۳] می‌توان دید. اکنون فرض کنیم $Y_۱, Y_۲, \ldots$ یک دنباله نامتناهی از متغیرهای تصادفی پیوسته و مستقل از $\{X_i, i \geq ۱\}$ باشد. در این صورت مقادیر kـ رکورد را می‌توان به‌عنوان تعمیمی از رکوردها در یک دنباله متغیرهای تصادفی هم‌توزیع و مستقل، مشاهدات متوالی از kامین بزرگ‌ترین مقدار نمونه تعریف کرد. در واقع مقادیر kـ رکورد بالا در دنباله‌ای از متغیرهای تصادفی، kامین بزرگ‌ترین مشاهده در یک نمونه‌گیری است، [۳۱]. برای اطلاع بیشتر درباره رکوردها بویژه kـ رکوردها، می‌توان به تحقیقات در 1,2,4,5,6,7,8,13,20 نیز مراجعه نمود.*

تعریف ۲.۱ ([۱۱]). *فرض کنیم $Y_{i:m}$، $\{T_{۱,k}, T_{۲,k}, \ldots\}$ و $\{R_{۱(k)}, R_{۲(k)}, \ldots\}$ به ترتیب نشان‌دهنده iامین آماره ترتیبی از یک نمونه تصادفی به حجم m، دنباله زمان‌های رخداد kـ رکورد بالا و مقادیر kـ رکورد بالا باشند. با قرار دادن $T_{۱,k} = k$ برای $m \geq ۲$ داریم:*

$$T_{m,k} = \min\{j : j > T_{m-۱,k}, Y_j > Y_{T_{m-۱,k}-k+۱:T_{m-۱,k}}\}$$

در این‌صورت دنباله kـ رکوردهای بالا عبارتند از: $R_{۱(k)} = Y_{۱:k}, R_{m(k)} = Y_{T_{m-۱,k}-k+۱:T_{m-۱,k}}$.

لم ۳.۱ ([۱۱]). *فرض کنیم $\{X_i, i \geq ۱\}$ دنباله‌ای نامتناهی از متغیرهای تصادفی مستقل با تابع توزیع تجمعی F و تابع چگالی f باشند. آنگاه تابع چگالی حاشیه‌ای mامین kـ رکورد عبارت است از*

$$f_{R_{m(k)}}(u) = \frac{k^m}{(m-۱)!}\{-\log(۱-F(u))\}^{m-۱}\{۱-F(u)\}^{k-۱}f(u) \qquad (۱.۱)$$

در تمام این مقاله، از kـ رکورد بالا استفاده می‌نماییم. برای تشریح بیشتر kـ رکورد بالا به مثال زیر توجه کنید:

مثال ۴.۱. *فرض کنیم یک نمونه تصادفی به حجم $n = ۱۶$ با مقادیر زیر داشته باشیم. این داده‌ها مربوط به زمان خرابی‌های پی در پی در سیستم تهویه مطبوع جت‌بوئینگ ۷۲۰ از ناوگان هواپیمایی ۸۰۴۵ است که توسط پورشان در مرجع [۲۵]، گزارش شده است:*

$$Y_i : ۱۰۲ - ۲۰۹ - ۱۴ - ۵۷ - ۵۴ - ۳۲ - ۶۷ - ۵۹ - ۱۳۴ - ۱۵۲ - ۲۷ - ۱۴ - ۲۳۰ - ۶۶ - ۶۱ - ۳۴$$

الف: در این صورت آماره‌های رکوردی نمونه با استناد به تعاریف ۱.۱ و ۲.۱ به شرح زیر است:

j	۱	۲	۳
زمان رکورد بالا	۱	۲	۱۳
مقدار رکورد بالا	۱۰۲	۲۰۹	۲۳۰

ب: به‌ازای $k = ۲$، زمان‌های رخداد mامین kـ رکورد و مقادیر mامین kـ رکورد عبارت است از

m	۱	۲	۳	۴
$T_{m,k}$	۲	۹	۱۰	۱۳
$R_{m(k)}$	۱۰۲	۱۳۴	۱۵۲	۲۰۹

ج: به‌ازای $k = ۳$، زمان‌های رخداد mامین kـ رکورد و مقادیر mامین kـ رکورد عبارت است از

m	۱	۲	۳	۴	۵	۶
$T_{m,k}$	۳	۴	۷	۹	۱۰	۱۳
$R_{m(k)}$	۱۴	۵۷	۶۷	۱۰۲	۱۳۴	۱۵۲

مطالعات واقعی متعددی وجود دارد که به دلایل نامعلومی تعدادی از مشاهدات آن ناپدید شده اند و فقط تعداد محدودی از مشاهدات آن در دسترس است که مقادیر آن‌ها از تمام مشاهدات ماقبل خود بزرگتر یا کوچکتر می‌باشند. همین سبب می‌شود که پژوهشگر به همه مشاهدات دسترسی نداشته باشد. بنابراین ناچار است در چنین وضعیتی برای پیش‌گویی مقادیر آینده از همان تعداد محدود مشاهدات (مقادیر رکورد) به‌عنوان مقادیر نمونه‌ای مشاهده شده استفاده کند. به‌کارگیری یک روش آماری مناسب نمونه‌گیری بر اساس طرح‌های مختلف که موجب به حداکثر رساندن دقت خصیصه‌های جامعه، صرفه جویی در زمان و هزینه می‌شود، همواره مورد علاقه آماردانان بوده است. در این مطالعه، با طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی RRSS[1] که ایده‌ی آن بر مبنای ساختار نمونه مجموعه رتبه‌دار نخستین بار توسط صالحی و احمدی [۲۸]، پیشنهاد شد به اختصار آشنا می‌شویم:
فرض کنیم n دنباله‌ی مستقل از متغیرهای تصادفی در اختیار است. نمونه‌گیری در هر دنباله به نحوی است که تنها رکوردها ثبت می‌شوند. همچنین iامین نمونه‌گیری دنباله‌ای زمانی متوقف می‌شود که iامین رکورد مشاهده شود. برای تحلیل آماری در دنباله iام، تنها مقدار رکورد iام نیاز است. به این ساختار، طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی گفته می‌شود. با توجه به توضیحات فوق، مراحل استخراج نمونه‌ای با حجم n به روش نمونه‌گیری مجموعه رتبه‌دار رکوردی با استفاده از دیاگرام زیر دقیق‌تر نشان داده می‌شود:

$$\begin{array}{lccccc} ۱: & R_{(۱)۱} & & & \Rightarrow & R_{۱,۱} = R_{(۱)۱} \\ ۲: & R_{(۱)۲} & R_{(۲)۲} & & \Rightarrow & R_{۲,۲} = R_{(۲)۲} \\ & \dots & \dots & & \Rightarrow & \dots \quad \dots \\ n: & R_{(۱)n} & R_{(۲)n} & R_{(n)n} & \Rightarrow & R_{n,n} = R_{(n)n} \end{array} \qquad (۲.۱)$$

در واقع نمونه حاصل از طرح مذکور، بردار nـ بعدی، $R = (R_{۱,۱}, R_{۲,۲}, \dots, R_{n,n})^T$ است. در طرح فوق، متغیر $R_{i(j)}$، بر iامین رکورد در دنباله jام دلالت دارد. درایه‌های بردار R، مشاهدات حاصل از این طرح هستند که اولا مستقلند و ثانیا این امکان وجود دارد که مرتب نباشند. اکنون فرض کنیم $R_{i,i}$ها رکورد iام در دنباله iام باشند، آنگاه تابع چگالی احتمال حاشیه ای iامین رکورد و تابع چگالی احتمال توام n رکورد به ترتیب به فرم زیر می‌باشند، [۱۱]:

$$f_{i,i}(x) = \frac{\{-\log(۱ - F(x))\}^{i-۱}}{(i-۱)!} f(x) \qquad (۳.۱)$$

$$f_R(r;\theta) = \prod_{i=۱}^{n} \frac{\{-\log(۱ - F(r_{i,i};\theta))\}^{i-۱}}{(i-۱)!} f(r_{i,i};\theta) \quad , \quad \theta \in \Theta \qquad (۴.۱)$$

که در آن‌ها، $r = (r_{۱,۱}, r_{۲,۲}, \dots, r_{n,n})^T$ مقدار مشاهده‌شده $R = (R_{۱,۱}, R_{۲,۲}, \dots, R_{n,n})^T$ و Θ فضای پارامتر است. از زمان ایجاد ایدهٔ طرح RRSS، پژوهش‌های ارزشمندی در این حیطه به ثبت رسیده است. صالحی و احمدی برآورد مدل تنش‌ـ مقاومت را در طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی برای توزیع نمایی مطالعه کردند ([۲۹]). نظریه اطلاع نیز توسط اسکندرزاده و همکاران در این طرح بررسی شده است ([۱۴]). در مرجع

[1] Record Ranked Set Sampling

[۲۴]، پاوول و توماس طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی همراه را معرفی کردند. صفریان و همکاران برآوردگرهای بهبود یافته‌ای برای مدل تنش ـ مقاومت به‌دست آورده‌اند، ([۲۷]). در مرجع [۳]، گل‌زاده‌گروی و همکاران به مقایسه برآوردها و پیشگویی‌های بیزی تجربی بر اساس طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی با طرح نمونه‌گیری معکوس پرداختند. اخیرا استنباط آماری بر اساس رکوردهای بالا در این طرح با استفاده از توزیع پارتو نوع اول، توسط گل‌زاده‌گروی مطالعه شده است، ([۱۶]).

روند نگارش این مقاله بدین شرح است که، پس از ارائه مقدمه و پیشنیازها، در بخش ۲، پیشگویی بیزی نقطه‌ای مقادیر k ـ رکورد، بر اساس داده‌های حاصل از طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی مورد بحث قرار گرفته است و به دنبال آن، در بخش ۳، به پیشگویی بازه‌ای مقادیر k ـ رکورد به روش تابع بقا[1] و روش HPD[2] پرداخته شده است. در بخش ۴، نتایج به‌دست آمده به کمک شبیه‌سازی را مورد مطالعه قرار داده‌ایم. بخش ۵ نیز به نحوه کاربست پیشگویی‌های به‌دست‌آمده بر اساس داده‌های واقعی اختصاص داده شده است. در بخش پایانی، نتیجه‌گیری و بحث آورده شده است.

۲ پیشگویی بیزی نقطه‌ای مقادیر k ـ رکورد

فرض کنید $\{X_i, i \geq 1\}$ دنباله‌ای نامتناهی از متغیرهای تصادفی مستقل و هم‌توزیع دارای توزیع رایلی با تابع چگالی

$$f(x;\lambda) = 2\lambda x e^{-\lambda x^2}, \qquad x > 0, \quad \lambda > 0, \tag{۱.۲}$$

و تابع توزیع تجمعی

$$F(x;\lambda) = 1 - e^{-\lambda x^2}, \tag{۲.۲}$$

باشد. اگر $L(\lambda|r)$ تابع درستنمایی نمونه‌های استخراج شده بر اساس طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی باشد، با استفاده از روابط (۱.۲) و (۲.۲) و قرار دادن آنها در رابطه (۴.۱) می‌توان نوشت

$$L(\lambda|r) \propto \lambda^N e^{-\lambda t}, \tag{۳.۲}$$

که در آن $N = \frac{n(n+1)}{2}$ و $t = \sum r_{i,i}^2$ مقدار مشاهده شده $T = \sum R_{i,i}^2$ است. می‌توان نشان داد $R_{i,i}^2$ دارای توزیع گاما با پارامترهای i و λ است. لذا T دارای توزیع گاما با پارامترهای N و λ است. در ادامه می‌خواهیم براساس طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی، بر روی مساله پیشگویی بیزی نقطه‌ای mامین k ـ رکورد، $R_{m(k)}$، تحت تابع زیان متقارن توان دوم خطا[3] که به‌صورت $L_1(\lambda, \hat{\lambda}) = (\hat{\lambda} - \lambda)^2$ است، متمرکز شویم.

۱.۲ پیشگویی بیزی نقطه‌ای $R_{m(k)}$ تحت تابع زیان توان دوم خطا

فرض کنیم $\pi(\lambda)$ توزیع پیشین مزدوج گاما، به فرم زیر باشد

$$\Pi(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}, \tag{۴.۲}$$

که در آن ابرپارامترهای α و β مقادیری مثبت و معلوم‌اند و $\Gamma(\alpha) = \int_0^\infty \lambda^{\alpha-1} e^{-\beta\lambda}\, \mathrm{d}\lambda$ است. با استفاده از روابط (۳.۲) و (۴.۲)، تابع چگالی پسین λ به فرم زیر به‌دست می‌آید

$$\Pi(\lambda|r) = \frac{(\beta+t)^{N+\alpha}}{\Gamma(N+\alpha)} \lambda^{N+\alpha-1} e^{-\lambda(\beta+t)}. \tag{۵.۲}$$

[1] Survival Function Method
[2] Highest Posterior Density Method
[3] Squared Error Loss Function

اکنون فرض کنیم $R_{m(k)}$، mامین k ـ رکورد حاصل از یک دنباله آینده باشد. آنگاه تابع چگالی احتمال پیشگوی بیزی[1] $R_{m(k)}$، به شرط r عبارت خواهد بود از

$$f^*_{R_{m(k)}}(u|r)=\int_{\circ}^{\infty} f_{R_{m(k)}}(u|\lambda)\Pi(\lambda|r)\,\mathrm{d}\,\lambda,$$

که در آن u مقدار مشاهده شده $R_{m(k)}$ می‌باشد. با بکارگیری توزیع رایلی و رابطه (۱.۱) خواهیم داشت

$$f_{R_{m(k)}}(u|\lambda)=\frac{۲(k\lambda)^m}{\Gamma(m)}u^{۲m-۱}e^{-k\lambda u^۲},\qquad u>\circ, \tag{۶.۲}$$

در نتیجه طبق تعریف امید ریاضی داریم

$$E(R^c_{m(k)})=\int_{\circ}^{\infty} u^c f_{R_{m(k)}}(u|\lambda)\,\mathrm{d}\,u=\frac{\Gamma(m+\frac{c}{۲})}{\Gamma(m)(k\lambda)^{\frac{c}{۲}}}. \tag{۷.۲}$$

در ادامه با استفاده از روابط (۵.۲) و (۶.۲) و اندکی می‌توان نوشت

$$f^*_{R_{m(k)}}(u|r)=\frac{۲}{B(m,N+\alpha)u}p(u)^m(۱-p(u))^{N+\alpha}, \tag{۸.۲}$$

که در آن $p(u)=\frac{ku^۲}{\beta+t+ku^۲}$ و $B(m,N+\alpha)=\frac{\Gamma(m)\Gamma(N+\alpha)}{\Gamma(m+N+\alpha)}$ است.

از ویژگی‌های تابع چگالی احتمال پیشگوی بیزی، $f^*_{R_{m(k)}}(u|r)$، مستقل بودن آن از پارامتر λ است. این ویژگی مثبت، تابع چگالی احتمال پیشگوی بیزی را به یک معیار مناسب برای استنباط آماری در خصوص مقادیر k ـ رکورد آینده تبدیل کرده است. پیشگویی بیزی $R_{m(k)}$ تحت تابع زیان متقارن توان دوم خطا، $\hat{R}^S_{m(k)}$، میانگین تابع چگالی احتمال پیشگوی بیزی است که به فرم

$$\hat{R}^S_{m(k)}=E(R_{m(k)}|r)=\int_{\circ}^{\infty} uf^*_{R_{m(k)}}(u|r)\,\mathrm{d}\,u=\sqrt{\frac{\beta+t}{k}}\frac{\Gamma(m+\frac{۱}{۲})\Gamma(N+\alpha-\frac{۱}{۲})}{\Gamma(m)\Gamma(N+\alpha)}, \tag{۹.۲}$$

محاسبه می‌شود. میانگین توان دوم خطا پیشگویی[2] (MSPE) برای $\hat{R}^S_{m(k)}$، به کمک روابط (۷.۲) و (۹.۲) با اندکی محاسبات برابر است با

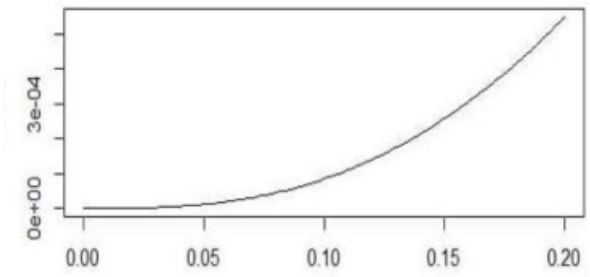
$$\begin{aligned}E(\hat{R}^S_{m(k)}-R_{m(k)})^۲=&E(\hat{R}^S_{m(k)})^۲+E(R_{m(k)})^۲-۲E(\hat{R}^S_{m(k)})E(R_{m(k)})\\ =&\frac{\Gamma^۲(m+\frac{۱}{۲})\Gamma^۲(N+\alpha-\frac{۱}{۲})}{k\Gamma^۲(m)\Gamma^۲(N+\alpha)}(\beta+E(T))\\ &+\frac{m}{k\lambda}-\frac{۲\Gamma^۲(m+\frac{۱}{۲})\Gamma(N+\alpha-\frac{۱}{۲})}{k\lambda^{\frac{۱}{۲}}\Gamma^۲(m)\Gamma(N+\alpha)}E\sqrt{\beta+t}\end{aligned}$$

با بهره‌گیری از بسط دوجمله‌ای، می‌توان آن را به‌صورت زیر دوباره نویسی کرد

$$\begin{aligned}E(\hat{R}^S_{m(k)}-R_{m(k)})^۲=&\frac{۱}{k\lambda}\Big(\frac{\Gamma^۲(m+\frac{۱}{۲})\Gamma^۲(N+\alpha-\frac{۱}{۲})}{k\Gamma^۲(m)\Gamma^۲(N+\alpha)}\Big)(\beta\lambda+N)\\ &+m-\frac{۲(\beta\lambda)^{\frac{۱}{۲}}\Gamma^۲(m+\frac{۱}{۲})\Gamma(N+\alpha-\frac{۱}{۲})}{k\lambda^{\frac{۱}{۲}}\Gamma^۲(m)\Gamma(\alpha)\Gamma(N+\alpha)}\sum_{i=\circ}^{\infty}\binom{\frac{۱}{۲}}{l}\frac{\Gamma(\alpha+l)}{\beta^{۲l}}.\end{aligned} \tag{۱۰.۲}$$

این تابع با فرض ثابت نگه داشتن سایر پارامترها، تابعی یکنوا صعودی از m است. شکل ۱ را ببینید.

[1] Bayesian Predictive Density Function
[2] Mean square prediction error



شکل ۱: نمودار میانگین توان دوم خطا پیشگویی $\hat{R}^{S}_{m(k)}$ نسبت به m برای $k = \lambda = \alpha = \beta = ۲$.

۲.۲ پیشگویی بیزی نقطه‌ای $R_{m(k)}$ تحت تابع زیان لاینکس

تابع زیان نامتقارن لاینکس[۱] به فرم $L_{۲}(\lambda, \hat{\lambda}) = b(e^{a(\hat{\lambda}-\lambda)} - a(\hat{\lambda} - \lambda) - ۱)$[۲] تعریف می‌شود که در آن ثابت $a \neq ۰$ شکل تابع زیان را مشخص می‌کند و $b > ۰$ نیز برای مقیاس تابع زیان به کار می‌رود. در این بخش، پیشگویی بیزی نقطه‌ای $R_{m(k)}$ تحت تابع زیان لاینکس، $\hat{R}^{L}_{m(k)}$، با استفاده از داده‌های حاصل از طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی ارائه می‌شود. پیش گویی بیزی نقطه‌ای برای $\hat{R}^{L}_{m(k)}$، به فرم $-\frac{۱}{a} \ln E(e^{-aR_{m(k)}}|r)$ است. با انجام اندکی محاسبات می‌توان دید

$$\hat{R}^{L}_{m(k)} = -\frac{۱}{a} \ln \int_{۰}^{\infty} e^{-au} f^{*}_{R_{m(k)}}(u|r)\, \mathrm{d}\, u = -\frac{۱}{a} \ln \sum_{i=۰}^{\infty} \frac{\left(-a\sqrt{\frac{\beta+t}{k}}\right)^{l}}{l!} \frac{\Gamma(m+\frac{۱}{۲})\Gamma(N+\alpha-\frac{۱}{۲})}{\Gamma(m)\Gamma(N+\alpha)}.$$

میانگین توان دوم خطا پیشگویی برای $\hat{R}^{L}_{m(k)}$، فرم بسته‌ای ندارد. برای حل آن باید به سراغ روش‌های عددی رفت.

۳ پیشگویی بیزی بازه‌ای مقادیر k ـ رکورد

پیشگویی‌های نقطه‌ای، اطلاعات زیادی درباره دقت پیشگویی ارائه نمی دهند. این نوع از پیشگویی‌ها یا صد درصد درست هستند و یا صد درصد نادرست. لذا در بسیاری از موارد، به‌دست آوردن یک بازه پیشگویی که احتمال قرار گرفتن متغیر تصادفی قابل پیشگویی در آن زیاد است کفایت می‌کند. در این بخش، قصد ایجاد یک بازه پیشگویی $(۱ - \gamma)$ درصدی برای $R_{m(k)}$ به روش تابع بقا و روش HPD با استفاده از داده‌های حاصل از طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی را داریم.

۱.۳ پیشگویی بازه‌ای به روش تابع بقا

برای ایجاد یک بازه پیشگویی بیزی برای $R_{m(k)}$ به روش تابع بقا، به کمک تابع بقا پیشگوی، $F^{*}_{R_{m(k)}}(u|r)$، داریم

$$F^{*}_{R_{m(k)}}(u|r) = P(R_{m(k)} > u) = \int_{۰}^{\infty} \frac{۲}{B(m, N+\alpha)y} p(y)^{m}(۱ - p(y))^{N+\alpha}\, \mathrm{d}\, y$$

حدود بازه پیشگویی بیزی $(۱ - \gamma)$ درصدی، از حل دو معادله زیر حاصل می‌شود

$$F^{*}_{R_{m(k)}}(L(r)) = ۱ - \frac{\gamma}{۲} \quad , \quad F^{*}_{R_{m(k)}}(U(r)) = \frac{\gamma}{۲} \qquad (۱.۳)$$

[1] Linex Loss Function

که در آن‌ها $L(r)$ و $U(r)$ به ترتیب حدود پائین و بالای بازه هستند. از رابطه (۱.۳) داریم

$$\int_{۰}^{L(r)} \frac{۲}{B(m,N+\alpha)y} p(y)^m (۱-p(y))^{N+\alpha}\,\mathrm{d}\,y = \frac{\gamma}{۲}$$

با استفاده از تغییر متغیر $p(y)=v$ و محاسباتی نه چندان دشوار، رابطه اخیر را می‌توان به فرم

$$\int_{۰}^{(۱+\frac{\beta+t}{kL(r)^۲})^{-۱}} \frac{۱}{B(m,N+\alpha)} v^{m-۱}(۱-v)^{N+\alpha-۱}\,\mathrm{d}\,v = \frac{\gamma}{۲} \qquad (۲.۳)$$

دوباره نویسی کرد. با استفاده از رابطه (۲.۳) خواهیم داشت، $L(r) = \sqrt{\frac{\beta+t}{k}\frac{B_{۱-\frac{\gamma}{۲}}(m,N+\alpha)}{۱-B_{۱-\frac{\gamma}{۲}}(m,N+\alpha)}}$ به‌طور مشابه نیز داریم: $U(r) = \sqrt{\frac{\beta+t}{k}\frac{B_{\frac{\gamma}{۲}}(m,N+\alpha)}{۱-B_{\frac{\gamma}{۲}}(m,N+\alpha)}}$. مقصود از نماد $B_\gamma(m,N+\alpha)$، چندک مرتبه γام توزیع بتا با پارامترهای m و $N+\alpha$ است. در نتیجه بازه پیشگویی بیزی برای $R_{m(k)}$ عبارت است از

$$\sqrt{\frac{۱}{k}\frac{B_{۱-\frac{\gamma}{۲}}(m,N+\alpha)}{۱-B_{۱-\frac{\gamma}{۲}}(m,N+\alpha)}}, \sqrt{\frac{۱}{k}\frac{B_{\frac{\gamma}{۲}}(m,N+\alpha)}{۱-B_{\frac{\gamma}{۲}}(m,N+\alpha)}}. \qquad (۳.۳)$$

طول این بازه تصادفی تابعی از T است. متوسط طول این بازه برابر است با

$$\left(\sqrt{\frac{۱}{k}\frac{B_{\frac{\gamma}{۲}}(m,N+\alpha)}{۱-B_{\frac{\gamma}{۲}}(m,N+\alpha)}} - \sqrt{\frac{۱}{k}\frac{B_{۱-\frac{\gamma}{۲}}(m,N+\alpha)}{۱-B_{۱-\frac{\gamma}{۲}}(m,N+\alpha)}}\right)\sum_{i=۰}^{\infty}\binom{\frac{۱}{۲}}{l}\frac{\Gamma(\alpha+l)}{\Gamma(\alpha)\beta^{۲l-\frac{۱}{۲}}}.$$

۲.۳ پیشگویی بازه‌ای به روش HPD

شکل ۲، نمودار تابع چگالی احتمال پیشگوی بیزی $f^*_{R_{m(k)}}(u|r)$ و هیستوگرام مقادیر تولید شده از آن بر اساس نمونه‌های شبیه‌سازی شده را نشان می‌دهد. همان‌طور که ملاحظه می‌شود، تابع چگالی احتمال پیشگوی بیزی تک مدی است. لذا به منظور ایجاد یک بازه پیشگویی بیزی به روش HPD (بازه دارای کوتاهترین متوسط طول) برای $R_{m(k)}$ تحت داده‌های حاصل از طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی، نیاز به حل همزمان دو معادله زیر داریم

$$\int_{\xi_۲}^{\xi_۱} f^*_{R_{m(k)}}(u|r)\,\mathrm{d}\,u = ۱-\gamma \quad , \quad f^*_{R_{m(k)}}(\xi_۱|r) = f^*_{R_{m(k)}}(\xi_۲|r), \qquad ۰<\gamma<۱$$

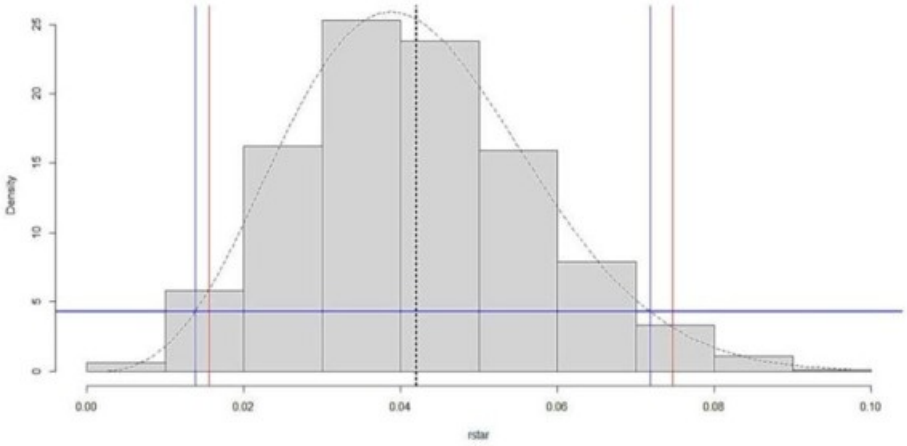
یا به‌طور معادل

$$\begin{aligned} &\int_{\xi_۲}^{\xi_۱} \frac{۲}{B(m,N+\alpha)u} p(u)^m (۱-p(u))^{N+\alpha}\,\mathrm{d}\,u = ۱-\gamma, \\ &\left(\frac{\xi_۱}{\xi_۲}\right)^{۲m-۱} = \left(\frac{\beta+t+k\xi_۱^۲}{\beta+t+k\xi_۲^۲}\right)^{m+N+\alpha} \end{aligned} \qquad (۴.۳)$$

به $(\xi_۱,\xi_۲)$ یک بازه پیشگویی بیزی $(۱-\gamma)$ درصدی برای $R_{m(k)}$ به روش HPD می‌گویند.

۴ مطالعه شبیه‌سازی

در این بخش، ارزیابی عملکرد طرح معرفی شده در پیشگویی بیزی مقادیر k‌ـرکورد آینده با استفاده از شبیه‌سازی مورد مطالعه قرار می‌گیرد. با این هدف، برای ابرپارامترهای (α,β) در رابطه (۴.۲)، از مقادیر (۱٫۵, ۱٫۵)، (۰٫۵, ۰٫۵) و



شکل ۲: منحنی تابع چگالی احتمال پیشگوی بیزی $f^*_{R_{m(k)}}(u|r)$ و هیستوگرام مقادیر تولید شده از آن بر اساس نمونه‌های شبیه سازی شده با ۱۰۰۰ بار تکرار

$(\circ,\circ)$ برای حالتهای آگاهی‌بخش و ناآگاهی‌بخش استفاده شده است. با استفاده از نرم‌افزار R، بر اساس یک نمونه تصادفی شبیه سازی شده با ۱۰۰۰ بار تکرار، پارامتر مجهول λ در مدل رایلی، تحت تابع زیان توان دوم خطا، برابر با ۰٫۴۲ تخمین زده شده است. در شکل ۲، خطچین عمودی مشکی، برآورد بیز پارامتر λ تحت توابع زیان توان دوم خطا و لاینکس است. این برآوردها چون نزدیک هم هستند، بر روی یک خط افتاده اند. معیارهای زیر برآورد میانگین توان دوم خطا پیشگویی تحت توابع زیان توان دوم خطا و لاینکس در ۱۰۰۰ بار تکرار را نمایش می‌دهد:

$$MSPE_1(\hat{Y}_i,Y_i)=\frac{1}{1000}\sum_{i=1}^{1000}L_1(\lambda,\hat{\lambda}) \quad , \quad MSPE_2(\hat{Y}_i,Y_i)=\frac{1}{1000}\sum_{i=1}^{1000}L_2(\lambda,\hat{\lambda})$$

کوچک بودن معیار MSPE بیانگر دقت بیشتر و عملکرد بهتر است. به‌ازای مقادیر $k=1,2,3$، $m=1,2,3$ و $a=-1,1$ نتایج شبیه سازی مربوط به مقادیر پیش‌گویی $R_{m(k)}$ و میانگین توان دوم خطا پیشگویی متناظر آن تحت توابع زیان توان دوم خطا و لاینکس در جدول ۱ گزارش شده است. دیده می‌شود، با فرض ثابت نگه داشتن سایر پارامترها و افزایش پارامتر m، برآورد میانگین توان دوم خطا پیشگویی نیز افزایش می‌یابد. این نتیجه در شکل ۱ برای $k=\lambda=\alpha=\beta=2$ در قالب نمودار نیز آورده شده است. همچنین در جدول ۱، با استفاده از معیار MSPE می‌توان دید که، برآورد mامین kـ رکورد آینده $R_{m(k)}$ تحت تابع زیان توان دوم خطا، کمتر از برآورد آن تحت تابع زیان لاینکس و در نتیجه بهتر از آن است.

جدول ۱: پیشگویی $R_{m(k)}$ و MSPE متناظرش (در داخل پرانتز) تحت توابع زیان توان دوم خطا و لاینکس

$\alpha=\beta=\circ$			$\alpha=\beta=\circ/\Delta$			$\alpha=\beta=1/\Delta$			k	m
Sel	Linex$(a=-1)$	Linex$(a=1)$	Sel	Linex$(a=-1)$	Linex$(a=1)$	Sel	Linex$(a=-1)$	Linex$(a=1)$		
۰٫۰۳۹۵(۰٫۷۱۶۹)	۰٫۰۳۹۳(۰٫۷۱۷۲)	۰٫۰۳۹۱(۰٫۷۱۷۶)	۰٫۰۳۹۵(۰٫۷۱۶۸)	۰٫۰۳۹۱(۰٫۷۱۷۱)	۰٫۰۳۹۳(۰٫۷۱۷۵)	۰٫۰۳۹۶(۰٫۷۱۶۷)	۰٫۰۳۹۴(۰٫۷۱۷۰)	۰٫۰۳۹۲(۰٫۷۱۷۴)	۱	۱
۰٫۰۱۹۷(۰٫۱۷۹۳)	۰٫۰۹۸۱(۱٫۶۶۶۱)	۰٫۵۵۳۶(۱٫۶۶۸۳)	۰٫۰۱۹۸(۰٫۱۷۹۱)	۰٫۰۹۰۷(۱٫۵۳۶۸)	۰٫۵۰۷۸(۱٫۶۳۱۴)	۰٫۰۱۹۸(۰٫۱۷۹۱)	۰٫۰۷۲۹۶(۱٫۵۱۸۵)	۰٫۰۴۱۷۷(۱٫۵۶۰۰)	۲	
۰٫۰۱۳۱(۰٫۰۷۹۷)	۰٫۰۹۸۲۱(۰٫۶۶۲۳)	۰٫۸۵۵۲(۱٫۶۸۱۴)	۰٫۰۱۳۲(۰٫۰۷۹۶)	۰٫۰۵۴۰۷(۰٫۶۶۱۴)	۰٫۸۲۴۶(۱٫۵۸۰۴)	۰٫۰۱۳۲(۰٫۰۷۹۵)	۰٫۰۹۲۳۱(۰٫۰۴۱۸۲)	۰٫۰۵۱۸۸(۱٫۴۶۰۱)	۳	
۰٫۰۸۳۸(۱٫۳۵۱۲)	۰٫۰۸۳۴(۱٫۴۵۲۲)	۰٫۰۸۲۹(۱٫۶۶۳۴)	۰٫۰۸۳۹(۱٫۲۵۱۰)	۰٫۰۸۳۴(۱٫۳۵۲۱)	۰٫۰۸۳۰(۱٫۵۵۳۳)	۰٫۰۸۴۰(۱٫۲۵۰۸)	۰٫۰۸۳۵(۱٫۲۵۱۹)	۰٫۰۸۳۰(۱٫۵۵۳۱)	۱	۲
۰٫۰۴۱۹(۰٫۳۸۷۹)	۰٫۰۴۱۸(۱٫۷۶۷۹)	۰٫۰۴۱۷(۱٫۷۷۸۳)	۰٫۰۴۱۹(۰٫۳۸۷۷)	۰٫۰۴۱۸(۱٫۶۸۷۸)	۰٫۰۴۱۷(۱٫۶۸۸۰)	۰٫۰۴۲۰(۰٫۳۸۷۳)	۰٫۰۵۴۰(۱٫۶۸۷۶)	۰٫۰۴۱۸(۱٫۶۸۷۹)	۲	
۰٫۰۲۷۹(۰٫۱۷۶۶)	۰٫۰۳۱۱۳(۰٫۹۵۸۳)	۰٫۵۱۵۲(۱٫۶۶۹۴)	۰٫۰۲۷۹(۰٫۱۷۲۴)	۰٫۰۴۴۱۴(۰٫۹۵۲۰)	۰٫۰۲۷۵۰(۱٫۶۵۵۲)	۰٫۰۲۸۰(۰٫۱۷۲۲)	۰٫۰۲۹۴۹(۰٫۸۶۳۶)	۰٫۰۲۰۱۷(۱٫۴۸۴۵)	۳	
۰٫۰۱۲۸۴(۱٫۳۵۱۷)	۰٫۰۱۲۷۷(۱٫۵۴۲۹)	۰٫۰۱۲۷۰(۱٫۷۶۵۰)	۰٫۰۱۲۸۵(۱٫۳۵۰۶)	۰٫۰۱۲۷۷(۱٫۳۵۲۷)	۰٫۰۱۲۷۰(۱٫۶۵۴۹)	۰٫۰۱۲۸۵(۱٫۳۵۰۳)	۰٫۰۱۲۷۸(۱٫۳۵۲۵)	۰٫۰۱۲۷۱(۱٫۶۵۴۷)	۱	۳
۰٫۰۶۴۲(۰٫۵۸۷۹)	۰٫۰۶۴۰(۱٫۷۸۸۵)	۰٫۰۶۳۸(۱٫۸۸۸۲)	۰٫۰۶۴۲(۰٫۵۸۷۶)	۰٫۰۶۴۰(۱٫۷۸۷۸)	۰٫۰۶۳۹(۱٫۸۵۸۱)	۰٫۰۶۴۳(۰٫۵۸۷۴)	۰٫۰۶۴۱(۱٫۷۷۷۳)	۰٫۰۶۴۰۳(۱٫۷۸۸۰)	۲	
۰٫۰۴۲۸(۰٫۲۶۱۹)	۰٫۰۴۲۷(۱٫۰۶۱۲)	۰٫۰۴۲۶(۱٫۶۶۷۳)	۰٫۰۴۲۸(۰٫۲۶۱۳)	۰٫۰۴۲۷(۰٫۹۶۱۲)	۰٫۰۴۲۷(۱٫۶۶۱۳)	۰٫۰۴۲۹(۰٫۲۶۱۰)	۰٫۰۴۲۸(۰٫۹۶۱۱)	۰٫۰۴۲۷(۱٫۵۶۱۲)	۳	

جدول ۲ برآورد مقدار احتمال پوشش به همراه برآورد میانگین طول بازه پیشگویی بیزی با ضریب پیشگویی ۰٫۹۵ برای $R_{m(k)}$ را نشان می‌دهد. برآورد مقدار احتمال پوشش، نسبت تعداد دفعاتی است که در ۱۰۰۰ بار تکرار، مقادیر واقعی $R_{m(k)}$ آینده در بازه‌های پیشگویی مربوطه قرار گیرند و به‌صورت $CP=\frac{1}{1000}\sum_{r=1}^{1000}I(L_r<Y_i<U_r)$ تعریف می‌شود که در آن مقصود از $I(\cdot)$، تابع نشانگر است. هم چنین معیار $EL=\frac{1}{1000}\sum_{r=1}^{1000}(L_r-U_r)$،

میانگین طول بازه‌های پیشگویی است. نتایج جدول ۲ بیانگر آن است که، طول بازه‌های پیشگویی به‌دست آمده به روش HPD به‌طور متوسط کوتاه‌تر از طول بازه‌های پیشگویی به‌دست آمده به روش تابع بقا است. احتمال پوشش نیز در بسیاری از حالات تفاوت چندانی با ضریب پیش‌گویی ۰٫۹۵ ندارد و از این رویکرد نیز خوب عمل کرده است. در شکل ۲، خطوط عمودی قرمز بازه پیشگویی تحت روش تابع بقا با دمهای برابر است. خطوط عمودی آبی نیز بازه اطمینان تحت روش HPD و خط افقی آبی، چگالی در دو نقطه است که چون نزدیک به هم هستند روی یک خط قرارگرفته است. شکل ۲ نتایج فوق را تایید می‌کند.

جدول ۲: برآورد میانگین طول بازه پیشگویی بیزی و احتمال پوشش (در داخل پرانتز) $R_{m(k)}$

m	k	$\alpha=\beta=1/5$		$\alpha=\beta=0/5$		$\alpha=\beta=0$	
		روش تابع بقا	روش HPD	روش تابع بقا	روش HPD	روش تابع بقا	روش HPD
۱	۱	۰٫۷۸۷(۰٫۹۶۱)	۰٫۷۴۷(۰٫۹۸۵)	۰٫۷۸۵(۰٫۹۸۹)	۰٫۷۴۵(۰٫۹۸۳)	۰٫۷۸۵(۰٫۹۷۹)	۰٫۷۴۴(۰٫۹۷۶)
	۲	۰٫۳۹۵(۰٫۹۰۶)	۰٫۳۷۵(۰٫۹۶۸)	۰٫۳۹۲(۰٫۹۹۹)	۰٫۳۷۲(۰٫۹۹۴)	۰٫۳۹۳(۰٫۹۸۴)	۰٫۳۷۲(۰٫۹۹۵)
	۳	۰٫۲۶۴(۰٫۹۵۵)	۰٫۲۵۱(۰٫۹۲۹)	۰٫۲۶۱(۰٫۹۹۳)	۰٫۲۴۸(۰٫۹۹۵)	۰٫۲۶۱(۰٫۹۶۲)	۰٫۲۴۷(۰٫۹۷۷)
۲	۱	۰٫۱۱۸۰(۰٫۹۷۹)	۰٫۱۱۵۱(۰٫۹۹۵)	۰٫۱۱۷۹(۰٫۹۹۵)	۰٫۱۱۴۹(۰٫۹۷۵)	۰٫۱۱۷۹(۰٫۹۷۱)	۰٫۱۱۴۹(۰٫۹۲۳)
	۲	۰٫۵۹۱(۰٫۹۱۲)	۰٫۵۷۶(۰٫۹۹۲)	۰٫۵۸۹(۰٫۹۹۹)	۰٫۵۷۴(۰٫۹۹۵)	۰٫۵۸۹(۰٫۹۰۴)	۰٫۵۷۳(۰٫۹۵۴)
	۳	۰٫۳۹۴(۰٫۹۴۳)	۰٫۳۸۴(۰٫۹۸۱)	۰٫۳۹۳(۰٫۹۷۶)	۰٫۳۸۳(۰٫۹۹۳)	۰٫۳۹۴(۰٫۹۰۴)	۰٫۳۸۳(۰٫۹۳۲)
۳	۱	۰٫۱۴۷۱(۰٫۹۸۷)	۰٫۱۴۴۲(۰٫۹۷۹)	۰٫۱۴۶۹(۰٫۹۶۹)	۰٫۱۴۴۰(۰٫۹۸۵)	۰٫۱۴۶۹(۰٫۹۶۲)	۰٫۱۴۴۱(۰٫۹۴۲)
	۲	۰٫۷۳۶(۰٫۹۱۱)	۰٫۷۲۱(۰٫۹۹۲)	۰٫۷۳۴(۰٫۹۵۷)	۰٫۷۲۰(۰٫۹۹۹)	۰٫۷۳۴(۰٫۹۰۶)	۰٫۷۲۱(۰٫۹۵۴)
	۳	۰٫۴۹۱(۰٫۹۹۱)	۰٫۴۸۱(۰٫۹۷۹)	۰٫۴۸۹(۰٫۹۱۳)	۰٫۴۸۰(۰٫۰۰۰)	۰٫۴۸۹(۰٫۹۹۴)	۰٫۴۸۰(۰٫۹۳۵)

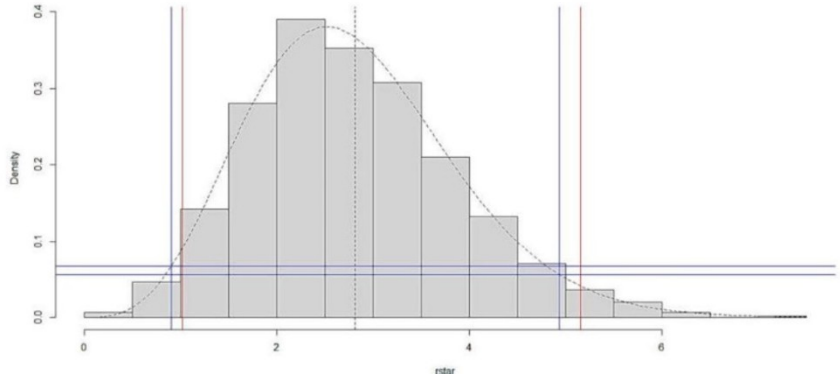
۵ مثال واقعی مرتبط با میزان سرعت وزش باد

درایه‌های بردار مشاهدات $r=(۳٫۰۵, ۶٫۹۴, ۶٫۳۹, ۳٫۶۱, ۶٫۳۹, ۶٫۶۷)^T$، رکوردهای به‌دست آمده بر اساس دیاگرام RRSS در رابطه (۲.۱) و داده‌های واقعی مرتبط با میزان سرعت وزش باد کشور کانادا، استان Brunswick new و ایستگاه Bas caraquej بین سال‌های ۲۰۱۵ تا ۲۰۲۰ میلادی می‌باشد که از سایت کمیسیون بین المللی اقیانوس شناسی IOC استخراج و برای ارزیابی مورد استفاده قرار گرفته است. در بررسی آزمون نیکویی برازش توزیع رایلی بر این داده‌ها، p_مقدار آزمون اندرسون_دارلینگ برابر ۰٫۹۷۱ محاسبه شده است. بیشتر بودن این مقدار احتمال از مقدار خطای نوع اول ($\gamma=0/05$) بیانگر مناسب بودن برازش توزیع رایلی بر روی این داده‌ها است. شکل ۳ نمودارهای احتمال P-P و احتمال تجمعی Q-Q را نشان می‌دهد. در بررسی شکل ۳، روند خطی دیده می‌شود و این یعنی توزیع رایلی به خوبی بر روی داده‌های میزان سرعت وزش باد برازش یافته است. پارامتر مجهول λ در مدل رایلی تحت توابع زیان توان دوم خطا و لاینکس نیز تخمین زده شده است. بر پایه این داده‌ها، بازه‌های پیشگویی به‌دست آمده به روش تابع بقا و روش HPD به ترتیب $(۱/۰۱۴۵, ۵/۱۵۶۹)$ و $(۰/۹۰۱۶, ۴/۹۳۱۶)$ است. پر واضح است که طول بازه پیشگویی به‌دست آمده به روش HPD کوتاهتر از طول بازه پیشگویی به روش تابع بقا است. مقدار MSPE تحت توابع زیان توان دوم خطا و لاینکس نیز محاسبه شده است. نتایج نشان داد که برآوردگر بیزی تحت تابع زیان توان دوم خطا بهتر از برآوردگر بیزی تحت تابع زیان لاینکس است. نتایج بر اساس داده‌های واقعی به‌صورت گرافیکی در شکل ۴ و جدول ۳ دیده می‌شود. نتایج به‌دست آمده در بخش داده‌های واقعی، نتایج حاصل از مطالعه شبیه‌سازی را تایید می‌کند.

جدول ۳: مقادیر به‌دست آمده بر اساس داده‌های واقعی مرتبط با میزان سرعت وزش باد

$\widehat{\lambda^{(S)}}$	$\widehat{\lambda^{(L)}}$	p - مقدار	طول بازه (HPD)	طول بازه (تابع بقا)	$MSPE^{(S)}$	$MSPE^{L}_{(a=1)}$	$MSPE^{L}_{(a=-1)}$
۲٫۷۹۶	۲٫۸۰۸	۰٫۰۸۶	۴٫۰۳	۴٫۱۴۲۴	۲٫۳۹۳	۳٫۱۱۳	۴٫۰۵۵

شکل ۳: نمودارهای احتمال P-P و Q-Q بر اساس داده‌های واقعی مرتبط با میزان سرعت وزش باد



شکل ۴: منحنی تابع چگالی احتمال بیزی پیشگوی و هیستوگرام مقادیر تولید شده از آن بر اساس داده‌های واقعی

۶ نتیجه‌گیری

یکی از مسائل مهم در استنباط آماری، به‌دست آوردن اطلاعات از طریق فرایندی است که بتوان با هزینه کمتر و کارایی بیشتر، وضع سیستم‌های آینده جامعه را که با گذشت زمان در حال تغییر هستند پیشگویی و تحلیل کرد. افزایش تعداد داده‌های نمونه همواره راهی است که می‌تواند به آماردانان کمک کند تا شواهد آماری را در جهت استنباط‌های دقیق تر در جامعه هدف، افزایش دهد. به هر حال در بسیاری از آزمایش‌های کاربردی میسر نیست که تعداد داده‌های نمونه را افزایش دهیم و عمدتا تصمیم‌گیری بر اساس تعداد محدود مشاهدات صورت می‌گیرد. در این صورت محقق باید سعی کند تا آنجا که مقدور است اطلاعات دیگری درباره توزیع نمونه‌گیری مشاهدات خود و یا استفاده از اطلاعات قبلی (روش بیزی) راجع به متغیرهای تصادفی آینده به‌دست آورد و یا طرح نمونه‌گیری مناسبی را برگزیند تا بتواند با این روش و با درصد خطای کمتری تصمیم گیری کند. در این مقاله طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی مورد استفاده قرار گرفت. با استفاده از داده‌های رکوردی حاصل از این طرح و در نظر گرفتن مدل رایلی، پیشگویی بیزی مقادیر k‌ـ رکورد آینده به کمک شبیه‌سازی و داده‌های واقعی مرتبط با میزان سرعت وزش باد ارزیابی شد. بررسی‌ها نشان داد نتایج پیشگویی مقادیر mامین k‌ـ رکورد حاصل از یک دنباله آینده بر اساس تابع زیان توان دوم خطا دقیق‌تر از نتایج به‌دست آمده بر اساس تابع زیان لاینکس است. هم چنین طول بازه‌های پیشگویی به‌دست آمده برای مقادیر mامین k‌ـ رکورد حاصل از یک دنباله آینده به روش HPD کوتاهتر و دارای احتمال پوشش بیشتر نسبت به بازه‌های پیشگویی به‌دست آمده بر اساس تابع بقا است. در بررسی این نوع پژوهش از دیدگاه آمار بیزی، خوب است که طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی به‌کار گرفته شود.

مراجع

[۱] افهمی،ب.، مددی، م.، رضاپور، م.، ۱۳۹۴. آزمون نیکویی برازش بر مبنای آنتروپی رکورد از توزیع پارتو تعمیم یافته، مجله علوم آماری، ۹ (۱)، صص. ۴۳-۶۰.

[۲] پاک،ع.، جعفری، ع. ا.، محمودی، م. ر.، ۱۴۰۱. استنباط درباره شاخص عملکرد طول عمر در توزیع نمایی دو پارامتری: داده‌های سانسورشده و رکورد ، پژوهش‌های ریاضی، ۸ (۴)، صص. ۱۹-۴۳.

[۳] گل‌زاده گروی، ا.، نصیری، پ.، صالحی، م.، ۱۴۰۰. مقایسه برآوردها و پیشگویی‌های بیزی تجربی بر اساس طرح نمونه‌گیری معکوس با طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی، مجله علوم آماری، ۱۵ (۱)، صص. ۱۹۳-۲۱۸.

[۴] گل‌زاده گروی، ا.، ۱۴۰۲. برآورد پارامتر مدل در تغییرات اقلیمی با بکارگیری داده‌های رکوردی حاصل از طرح نمونه‌گیری مجموعه رتبه‌دار رکوردی، نشریه پژوهش‌های اقلیم شناسی، ۱۴۰۲ (۵۳)، صص. ۵۱-۵۶.

[5] Ahmadi, J., Doostparast, M., and Parsian, A., 2005. Estimation and Prediction in a two-parameter exponential distribution based on k-record values under LINEX loss function. *Comm. Statist. Theory Methods.*, 34 , pp. 795-805.

[6] Ahmadi, J., Jafari Jozani, M., Marchand, E., and Parsian, A., 2009. Prediction of k-records from a general class of distributions under balanced type loss functions. *Metrika.*,70 , pp. 19-33.

[7] Ahmadi, J., MirMostafaee, S.M.T.K., and Balakrishnan, N., 2010. Bayesian Prediction of order statistics based on k-record values from exponential. *Statistics*, 34, pp. 795-805.

[8] Ahmadi, J., Mirmostafaee, S.M.T.K., and Balakrishnan, N., 2012. Bayesian prediction of k-record values based on progressively censored data from exponential distribution. *Journal of Statistical computation and simulation*, 82, pp. 51-62.

[9] Ahsanullah, M., 1995. *Record Statistics*. Nova Science Publishers Inc, Commack, New York.

[10] Aitchison, J., and Dunsmore, I.R., 1975. *Statistical prediction analysis*. Cambridge University Press, Cambridge.

[11] Arnold, B.C., Balakrishnan, N., and Nagaraja, H.N., 1998. *Records*, John Wiley & Sons, New York.

[12] Asgharzadeh, A., Valiollahi, R., Fallah, A., and Nadarajah, S., 2018. Bayesian inference and prediction for normal distribution based on records. *Statistics*, 78, pp. 15-36.

[13] Doostparast, M., and Doostparast, M., 2018. Prediction of corrosions in Gas and Oil pipelines based on the theory of records. *arXiv preprint arXiv:1801.00959*.

[14] Eskandarzadeh, M., Tahmasebi, S., and Afshari, M., 2016. Information measures for Record ranked set samples. *Ciencia e Natura.*, 38 , pp. 554-563.

[15] Ghafouri, S., Habibirad, A., Yousefzadeh, F., and Pakyari, R., 2022. Non-Bayesian Estimation and Prediction under Weibull Interval Censored Data. *Journal of Statistical Research of Iran JSRI*, 17 pp. 171-190.

[16] Golzade Gervi, E., 2024. Inference for the Pareto Type-I distribution using upper record ranked set sampling scheme. *Int. J. Nonlinear Anal. Appl.*, 15 , pp. 115-123.

[17] Golzade Gervi, E., Nasiri, P., and Salehi, M., 2019. Comparision of record tanked set sampling and ordinary records in prediction of future record statistics from an exponential distribution. *Statistical Research and Training Center*, 16, pp. 73-99.

[18] Golzade Gervi, E., Nasiri, P., and Salehi, M., 2021. An overview of Bayesian prediction of future record statistics using upper record ranked set sampling scheme. *Int. J. Nonlinear Anal. Appl.*, 12, pp. 493-507.

[19] Gulati, S., and Padgett, W.J., 2003. *Parametric and Non-Parametric Inference from Record-breaking Data*. Lecture Notes in Statistics., 172 Springer, New York.

[20] Klimczak, M., 2006. Prediction of kth record, Statist. Probab. *Lett.*, 76, pp. 117-127.

[21] Mojiri, A., Waghei, Y., Nili Sani, H.R., and Mohtashami Borzadaran, G.R., 2023. Non-stationary spatial autoregressive modeling for the prediction of lattice data. *Communications in Statistics-Simulation and Computation*, 52, pp. 5714-5726.

[22] Moradi, M., Jabbari Nooghabi, M., and Rounaghi, M.M., 2021. Investigation of fractal market hypothesis and forecasting time series stock returns for Tehran Stock Exchange and London Stock Exchange. *International Journal of Finance Economics*, 26, pp. 662-678.

[23] Nevzorov, V., 2001. Records. Mathematical Theory. *Translation of Mathematical Monographs*, 194, American Mathematical Society, Providence, RI.

[24] Paul, J., and Thomas, P.Y., 2017. Concomitant Record ranked set sampling. *Communications in Statistics -Theory and Methods*, 46, pp. 9518-9540.

[25] Porschan, F., 1963. Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5, pp. 375-383.

[26] Saadati Nik, A., Asgharzadeh, A., and Raqab, M.Z., 2021. Estimation and prediction for a new Paretotype distribution under progressive type-II censoring. *Mathematics and Computers in Simulation*, 190, pp. 508-530.

[27] Safaryian, A.,Arashi, M., and Arabi, R., 2019. Improved Estimators for stress-strength Reliability using Record ranked set sampling scheme. *Communications in statistics-simulation and computation*, 33. doi:10.1080/03610918.2018.1468451.

[28] Salehi, M., and Ahmadi, J., 2014. Record ranked set sampling scheme. *Metron*, 72, pp. 351-365.

[29] Salehi, M., and Ahmadi, J., 2015. Estimation of stress-strength Reliability using Record ranked set sampling scheme from the Exponential distribution. *Filomat.*, 29, pp. 1149-1162.

[30] Salehi, M., Ahmadi, J., and Balakrishnan, N., 2015. Prediction of order Statistics and Record values based on ordered Ranked set sampling. *Statistical Computation and Simulation*, 85, pp. 77-88.

[31] Valiollahi, R.,Asgharzadeh, A., and HKT, Ng., 2022. Prediction for Lindley distribution based on type-II right censored samples. *Journal of the Iranian Statistical Society*, 16, pp. 1-19.