Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

UNIVERSAL WISER PUBLISHER

Research Article

Bias Reduction of Maximum Likelihood Estimation in the Inverse Xgamma Distribution

Ahmed Abdulhadi Ahmed[1], Zakariya Yahya Algamal[1*], Olayan Albalawi[2]

[1] Department of Statistics and Informatics, University of Mosul, Mosul, Iraq
[2] Department of Statistics, Faculty of Science, University of Tabuk, Tabuk, Saudi Arabia
E-mail: zakariya.algamal@uomosul.edu.iq

Received: 17 January 2024; **Revised:** 4 February 2024; **Accepted:** 23 February 2024

Abstract: A distribution for modeling survival data that accounts for flexibility in modeling data with upside-down bathtub-shaped hazard rate functions is the inverse Xgamma distribution (IXG). The maximum likelihood method (MLE) is the most often used technique for parameter estimation of the IXG distribution. Conversely, the MLE is infamously biased for small sample sizes. This motivates us to produce almost unbiased estimators for IXG parameter. More precisely, we minimize MLE biases to the second degree of bias using two techniques for bias correction: bootstrap and analytical approaches. Two actual data applications and Monte Carlo simulations are used to compare the performances of these methods.

***Keywords*:** bias correction, survival analysis, the inverse Xgamma distribution, bootstrap

MSC: 62F10, 62N01

1. Introduction

Survival data analysis can be used to statistically analyze time-to-event data. In survival analysis, the primary outcome of interest is the amount of time until an event of interest occurs. This could be anything from the period until a patient relapse to the likelihood that a machine will malfunction or that a client will depart. The process of modeling and assessing the amount of time until an event of interest using statistical techniques is known as statistical modeling of survival data [1].

An essential component of survival analysis is selecting a statistical distribution to model survival data [2]. The instantaneous failure rate at any given moment is represented by the underlying hazard function, about which different distributions make different assumptions [3–6].The properties of the survival data and the underlying biological or physical processes should direct the choice of distribution. Visual evaluations, domain expertise, and goodness-of-fit tests can all be used to guide the selection of a model.

Applications for the inverse Xgamma distribution include extreme value theory, lifetime modeling, and reliability analysis. It works especially well when modeling variables with big tails, skewness, and a positive value. Regarding the field of survival analysis, the inverse Xgamma distribution was utilized for modeling survival data [7]. Yadav, Maiti and

DOI: https://doi.org/10.37256/cm.5320244311

Saha [7] presented an inverse Xgamma distribution using the idea of inverse distribution to model upside-down bathtub-shaped hazard rate function.

The XG distribution's probability density function (PDF) and cumulative distribution function (CDF) are provided as

$$F(y,\ \theta) = 1 - \frac{\left(1+\theta+\theta y+\frac{\theta^2 y^2}{2}\right)}{1+\theta} \mathrm{e}^{-\theta y};\ \theta > 0,\ y > 0. \tag{1}$$

$$f(y,\ \theta) = \frac{\theta^2}{1+\theta}(1+\frac{\theta y^2}{2})e^{-\theta y};\ \theta > 0,\ y > 0 \tag{2}$$

The random variable $X = (1/Y)$ is said to have an IXG distribution with PDF having the following form if a random variable Y follows the $XG\ (\theta)$ distribution with PDF given in Eq. (1).

$$f(x) = \frac{\theta^2}{x^2(1+\theta)}\left(1+\frac{\theta}{2x^2}\right)e^{-\frac{\theta}{x}};\ \theta > 0,\ x > 0 \tag{3}$$

It is denoted by $X \sim IXG\ (\theta)$, The IXG distribution's CDF is provided by

$$F(x) = \left(1+\frac{\theta}{x(1+\theta)}+\frac{\theta^2}{x^2 2(1+\theta)}\right)e^{-\frac{\theta}{x}};\ \theta > 0,\ x > 0 \tag{4}$$

2. Maximum likelihood estimation (MLE)

Suppose that $X = (x_1,\ x_2,\ \ldots,\ x_n)$ be a random sample of size n from the IXG distribution. The log-likelihood function of θ is given by

$$L(\theta,\ \alpha) = n\ln(\theta) - n\ln(\theta+1) - 2\sum_{i=1}^{n}\ln(x_i) + \sum_{i=1}^{n}\ln(1+\frac{\theta}{2x_i^2}) - \theta\sum_{i=1}^{n}\frac{1}{x_i} \tag{5}$$

Maximize Eq. (5) with respect to θ in order to obtain the MLE $(\hat{\theta})$ of θ. We have the following equations:

$$\frac{\partial}{\partial\theta}L(\theta,\ \alpha) = \frac{n}{\theta} + \sum_{i=1}^{n}\frac{1}{2x_i^2+\theta} - \frac{n}{1+\theta} - \sum_{i=1}^{n}\frac{1}{x_i} = 0 \tag{6}$$

Since Eq. (6) are non-linear they cannot be solved analytically. As MLE will be biased for small sample sizes. Findings in deceptive ways, which impacts how phenomena are interpreted in practical applications. This encourages us to think about unbiased estimations, essentially lowering the MLE distribution of these parameters' bias.

3. Bias-corrected MLEs

A statistical method called bias-corrected maximum likelihood estimation (BC-MLE) is used to account for bias in parameter estimations that are derived from MLE. When the average value of the estimates, computed across a large number of samples, differs from the true parameter value, the concept of bias in MLE emerges. In order to give more accurate parameter estimates, bias-corrected approaches try to minimize or completely remove this systematic mistake [8–13].

To evaluate the bias and apply corrections, methods like corrective approach and bootstrap are frequently used [14–17]. This method is useful when bias could compromise the validity of statistical conclusions. In the literature, inspired by these two approaches, a large number of authors tackled BC-MLE issue. Among them are: [18–22].

3.1 *A corrective approach*

Let $L(\tau)$ be the log-likelihood function of a parameter with p dimensions $\tau = (\tau_1,\ \tau_2,\ \ldots,\ \tau_n)$ based on a sample of observations x. The derivatives of the log-likelihood function for $i,\ j = 1,\ 2,\ 3,\ \ldots,\ p$ joint cumulants are given by

$$M_{ij} = E\left[\frac{\partial^2 L(\tau)}{\partial \tau_i \partial \tau_j}\right], \tag{7}$$

$$M_{ijl} = E\left[\frac{\partial^3 L(\tau)}{\partial \tau_i \partial \tau_j d\tau_l}\right], \tag{8}$$

$$M_{ij,l} = E\left[\left(\frac{\partial^2 L}{\partial \tau_i \partial \tau_j}\right)\left(\frac{dL}{d\tau_l}\right)\right], \tag{9}$$

where the derivatives of the joint cumulants are given by

$$M_{ij}^{(l)} = \frac{\partial M_{ij}}{\partial \tau_l} \tag{10}$$

Up to the third order, the log-likelihood function exhibits regularity and good behavior for all derivatives [23].

Cox and Snell [24] demonstrated that the bias of the sth element of the MLE of is when sample data are independent but not necessarily identically distributed

$$Bias(\hat{\tau}_s) = \sum_{i=1}^{p}\sum_{i=1}^{p}\sum_{i=1}^{p} M^{si}M^{jl}\left[\frac{1}{2}M_{ijl} + M_{ij,\ l}\right] + o(n^{-2})s = 1,\ 2,\ \ldots,\ p \tag{11}$$

where M^{ij} is the Fisher information matrix's $(i,\ j)^{th}$ inverse element. Then, Cordeiro and Cribari-Neto [8] noted that in the event where the observations lack independence, the bias expression remains valid. In place of Equation (11) they suggested the easy version that follows as appropriate

$$Bias\ (\hat{\tau}_s) = \sum_{i=1}^{p} M^{si}\sum_{i=1}^{p}\sum_{i=1}^{p}\left[M_{ij}^{(l)} - \frac{1}{2}M_{ijl}\right]M^{jl} + o(n^{-2})s = 1,\ 2,\ \ldots,\ p \tag{12}$$

Given that the elements of the form defined in Equation (10), are absent from Eq. (12), Eq. (12) has a computational advantage over Eq. (11).

Then, Let $M = \{-M_{ij}\}$ It is Fisher's information matrix of , and let $a_{ij}^{(l)} = M_{IJ}^{(l)} - \frac{1}{2}M_{ijl}$ they are elements $A^{(l)} = a_{ij}^{(l)}$ matrix for $i,\ j,\ l = 1,\ 2,\ 3,\ \ldots,\ p$. We have $A = \left[A^1 \middle| A^2 \middle| A^3 \middle| \ldots \middle| A^{(P)}\right]$, with $A^{(l)} = \left[a_{ij}^{(l)}\right]$.
Consequently, the matrix representation of $\hat{\tau}$'s bias expression can be expressed as

$$Bias(\hat{\tau}) = M^{-1}A \cdot vec(M^{-1}) + o(n^{-2}). \tag{13}$$

Thus, this shows that the BC-MLE of τ using the corrective approach (CA-MLE), $\hat{\tau}^{CA-MLE}$, is given by

$$\hat{\tau}^{CMLE} = \hat{\tau} - M^{-1}A \cdot vec(M^{-1}) \tag{14}$$

where $\hat{\tau}$ is the MLE of τ, $\hat{M} = M|_{\tau=\hat{\tau}}$, and $\hat{A} = A|_{\tau=\hat{\tau}}$. Whereas the bias of $\hat{\tau}^{CA-MLE}$ is quadratic.

Related to IXG distribution, the derivatives are obtained as

$$\frac{\partial^2}{\partial\theta^2}L(\tau) = -\frac{2n}{\theta^2} - \sum_{i=1}^{n}\frac{1}{(2x_i^2+\theta)^2} + \frac{n}{(1+\theta)^2} = 0 \tag{15}$$

$$\frac{\partial^3}{\partial\theta^3}L(\tau) = \frac{4n}{\theta^3} - \sum_{i=1}^{n}\frac{2}{(2x_i^2+\theta)^3} - \frac{2n}{(1+\theta)^3}$$

Then,

$$A = \left[A^{(1)}\right] = \left[a_{11}^{(1)}\right] \tag{16}$$

with

$$a_{11}^{(1)} = M_{11}^{(1)} - \frac{1}{2}M_{111} \tag{17}$$

where M_{ijl} is defined as by assuming $x = e^y$ and $dx = e^y dy$

$$M_{11} = E\left[\frac{\partial^2}{\partial\theta^2}L(\tau)\right] = \frac{n}{(1+\theta)^2} - \int_0^\infty \frac{(-n\ln(1+\theta)+2n\ln(\theta)-n2\ln(e^y)+n\ln\left(1+\frac{\theta}{2e^y}\right)-\frac{n\theta}{e^y})n}{4e^y\left(1+\frac{\theta}{2e^y}\right)^2}dy - \frac{2n}{\theta^2}$$

$$M_{111} = E\left[\frac{\partial^3}{\partial\alpha^3}L(\tau)\right] = -\frac{2n}{(1+\theta)^3} - \int_0^\infty \frac{(-n\ln(1+\theta)+2n\ln(\theta)-n2\ln(e^y)+n\ln\left(1+\frac{\theta}{2e^y}\right)-\frac{n\theta}{e^y})n}{4e^{2y}\left(1+\frac{\theta}{2e^y}\right)^3}dy + \frac{4n}{\theta^3}$$

Therefore, the bias MLE of IXG distribution is given by

$$Bias\left(\hat{\theta}\right) = M^{-1}AVec(M^{-1}) + o(n^{-2}). \tag{18}$$

And then,

$$\hat{\theta}_{CA-MLE} = \hat{\theta}_{MLE} - Bias(\hat{\theta}). \tag{19}$$

4. Bootstrap approach

An alternative method based on the parametric bootstrap resampling methodology is used to produce second-order bias-corrected estimators [3]. Let $X = (x_1, x_2, x_3, \ldots, x_n)$ be a random sample of size n from the random variable X with the distribution function F. By generating Bindependent bootstrap samples from distribution function F, the estimated bias of the MLE of $\widehat{\tau}$ is

$$Bias(\hat{\tau}_{MLE}) = \frac{1}{B}\sum_{j=1}^{B}(\hat{\tau}^*_{j,\ MLE} - \hat{\tau}_{MLE}) \tag{20}$$

Where $\hat{\tau}^*_j$ is the MLE of τ from the j^{th} bootstrap sample generated from the IXG distribution. Then, the bias-corrected bootstrap approach (BC-Boot) [9, 10] is defined as

$$\hat{\tau}_{BC-Boot} = 2\hat{\tau}_{MLE} - \frac{1}{B}\sum_{j=1}^{B}\hat{\tau}^*_{j,\ MLE}. \tag{21}$$

5. Simulation results

This simulation study's objective is to assess how well the several estimators of the IXG distribution's parameters: MLE, CA-MLE, and BC-Boot perform. The IXG distribution was used to generate samples with sizes $n = 10,\ 30,\ 50,$ and 100, with parameters $\theta = 0.8$, $\theta = 1.2$, and $\theta = 3$. Each case was generating under Monte Carlo samples with 5,000 times and 1,000 bootstrap samples in each time. The root mean square error (RMSE) and bias of the estimations, as well as their accuracy, are assessed, which they defined in Eq. (28) and Eq. (29), respectively, are reported. All results of the averaged biases and RMSE are summarized in Tables 1-3.

$$Bias(\hat{\tau}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\tau}_{i,\ BC-MLE} - \hat{\tau}_{MLE}) \tag{22}$$

$$RMSE(\tau) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\tau}_{i,\ BC-MLE} - \hat{\tau}_{MLE})^2} \tag{23}$$

From Tables 1-3, there are a few conclusions that can be reached.

For all the simulations considered, the MLE estimators of θ seem to be biased in the positive direction. This illustrates how, in general, they overstate the parameter θ value, particularly in cases where the sample size is small. Furthermore, when the real value of the parameter θ is equal to or larger than 0.8, the MLE estimators frequently exhibit a positive bias, that is, they continuously overestimate the true value of the parameter θ for various sample sizes.

The BC-Boot and CA-MLE estimators outperformed the MLE estimators of θ in terms of Bias and RMSE in all of simulations for different sample sizes. Further, the BC-Boot of θ outperformed the CA-MLE in terms of RMSE. Conversely, CA-MLE attained better performance than BC-Boot for θ in terms of Bias.

Naturally, all of the analyzed estimators' biases and RMSEs will decrease as sample size n increases. This is mostly because most estimators in statistical theory perform better as sample size n increases. As previously stated, for small sample numbers, both CA-MLE and BC-Boot show extremely significant reductions in bias and RMSE. For instance, from Table 1, in the case of $n = 10$, it can be seen that the reduction in RMSE of both CA-MLE and BC-Boot was about 17.03% and 17.41% for θ lower than that of the MLE. On the other hand, the reduction for the same case of both CA-MLE and BC-Boot in terms of Bias was 75.02% and 65.91% for θ lower than that of the MLE, respectively.

Finally, although the two approached, CA-MLE and BC-Boot, are equally efficient, the BC-Boot is computationally easier than the CA-MLE.

Table 1. Average RMSE and Bias when $\theta = 0.8$

n		MLE	CA-MLE	BC-Boot
10	RMSE	0.3951	0.3278	0.3263
	Bias	0.2138	0.0534	0.0729
30	RMSE	0.3349	0.2676	0.2661
	Bias	0.2084	0.0594	0.0675
50	RMSE	0.3071	0.2385	0.237
	Bias	0.207	0.0643	0.0658
100	RMSE	0.263	0.1959	0.1942
	Bias	0.2061	0.0737	0.0781

Table 2. Average RMSE and Bias when $\theta = 1.2$

n		MLE	CA-MLE	BC-Boot
10	RMSE	0.4512	0.3835	0.382
	Bias	0.2695	0.0609	0.1285
30	RMSE	0.3906	0.3233	0.3218
	Bias	0.264	0.0555	0.1232
50	RMSE	0.3628	0.2942	0.2929
	Bias	0.2627	0.0541	0.1215
100	RMSE	0.3188	0.2516	0.2499
	Bias	0.2618	0.0531	0.12

Table 3. Average RMSE and Bias when $\theta = 3$

n		MLE	CA-MLE	BC-Boot
10	RMSE	0.5045	0.4372	0.4357
	Bias	0.3232	0.1146	0.1225
30	RMSE	0.4443	0.3772	0.3755
	Bias	0.3178	0.1092	0.1169
50	RMSE	0.4165	0.3479	0.3464
	Bias	0.3164	0.1078	0.1152
100	RMSE	0.3724	0.3053	0.3036
	Bias	0.3155	0.1069	0.1122

6. Real data application

In this part, we use two real datasets with a small sample to demonstrate the usefulness of the suggested bias-corrected estimators for the IXG distribution. The first dataset representing the life time failure of 18 electronic device [10]. This data was further analyzed by M. Wang and Wang [10]. The second dataset represents the tubes to show leak under 120 psi stress level [11]. The sample size of this data is 30.This data was further analyzed by Çetinkaya and Bulut [17, 25–27].

The estimated values for the IXG distribution's parameters are shown in Table 4. Table 4 shows that the MLE technique overestimates this parameter since the estimates of from the CA-MLE and BC-Boot are smaller than the MLE estimate.

The analysis of the IXG distribution pdf in relation to Table 4 for θ values of both datasets is shown in Figures 1 and 2, respectively. We suggest using CA-MLE and BC-Boot estimates for both datasets because the density shape based on the MLE method may be deceptive, as this Figure illustrates.

Table 4. Point estimates of the θ of IXG distribution for both data set used

	Electronic device	Show leak data
	θ	θ
MLE	1.1021	0.6841
CA-MLE	1.0446	0.6722
BC-Boot	1.0359	0.6536

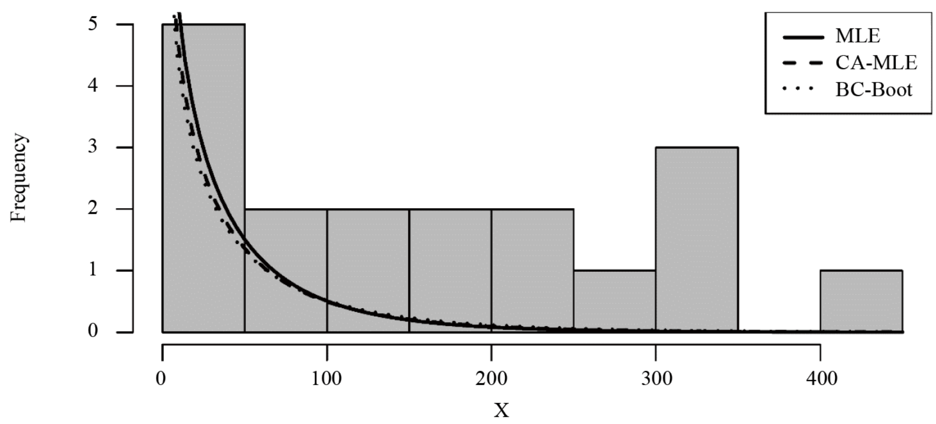

Figure 1. Fitted density functions estimated of the first dataset

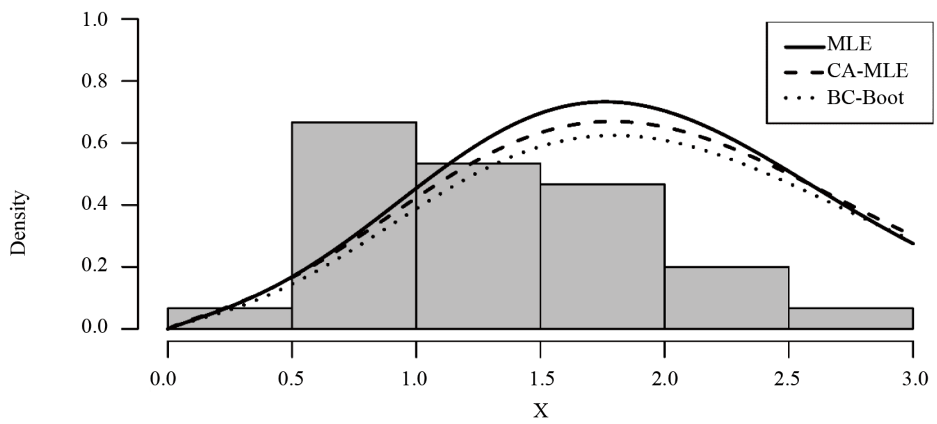

Figure 2. Fitted density functions estimated of the second dataset

7. Conclusion

In order to obtain straightforward closed-form equations for the second-order biases of the MLE of the parameters of the IXG distribution, the corrective method was proposed in this paper. Namely: CA-MLE and BC-Boot. Comparing the newly suggested estimators to the MLE, they converge to their genuine value much more quickly, as evidenced by their biases being of order $O(n^{-2})$as opposed to $O(n^{-1})$for the MLE. in terms of bias and RMSE, the suggested approaches exceed the MLE, as demonstrated by the numerical data, making them highly appealing. The suggested bias-corrected estimators are highly advised, particularly in cases when the sample size is not large.

Conflict of interest

The authors declare there is no conflict of interest at any point with reference to research findings.

References

[1] Yao DS, Chen WX, Long CX. Parametric estimation for the simple linear regression model under moving extremes ranked set sampling design. *Applied Mathematics-A Journal of Chinese Universities*. 2021; 36(2): 269-277.

[2] Bibani AA, Algamal ZY. Survival function estimation for fuzzy gompertz distribution with neutrosophic date. *International Journal of Neutrosophic Science*. 2023; 21(3): 137-142.

[3] Alanaz MM, Algamal ZY. Neutrosophic exponentiated inverse Rayleigh distribution: Properties and application. *International Journal of Neutrosophic Science*. 2023; 21(4): 36-43.

[4] Alanaz MM, Mustafa MY, Algamal ZY. Neutrosophic lindley distribution with application for alloying metal melting point. *International Journal of Neutrosophic Science*. 2023; 21(4): 65-71.

[5] Algamal ZY, Alobaidi NN, Hamad AA, Alanaz MM, Mustafa MY. Neutrosophic Beta-Lindley distribution: Mathematical properties and modeling bladder cancer data. *International Journal of Neutrosophic Science*. 2024; 23(2): 186-194.

[6] Mustafa MY, Algamal ZY. Neutrosophic inverse power lindley distribution: A modeling and application for bladder cancer patients. *International Journal of Neutrosophic Science*. 2023; 21(1): 216-223.

[7] Yadav AS, Maiti SS, Saha M. The inverse xgamma distribution: Statistical properties and different methods of estimation. *Annals of Data Science*. 2019; 8(2): 275-293.

[8] Cordeiro GM, Cribari-Neto F. *An Introduction to Bartlett Correction and Bias Reduction*. Heidelberg: Springer; 2014.

[9] Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993; 80(1): 27-38.

[10] Zhang J. Bias correction for the maximum likelihood estimate of ability. *Statistics*. 2005; 2005(2): i-39.

[11] Afify AZ, Aljohani HM, Alghamdi AS, Gemeay AM, Sarg AM, Martinucci B. A new two-parameter burr-hatke distribution: Properties and bayesian and non-bayesian inference with applications. *Journal of Mathematics*. 2021; 2021(3): 1-16.

[12] Al-Shomrani AA. An improvement in maximum likelihood estimation of the Burr XII distribution parameters. *AIMS Mathematics*. 2022; 7(9): 17444-17460.

[13] Al-Shomrani AA. An improvement in maximum likelihood estimation of the gompertz distribution parameters. *Journal of Statistical Theory and Applications*. 2023; 22(1-2): 98-115.

[14] Stošić BD, Cordeiro GM. Using maple and mathematica to derive bias corrections for two parameter distributions. *Journal of Statistical Computation and Simulation*. 2019; 79(6): 751-767.

[15] Arrué J, Arellano-Valle RB, Calderín-Ojeda E, Venegas O, Gómez HW. Likelihood based inference and bias reduction in the modified skew-t-normal distribution. *Mathematics*. 2023; 11(5): 881-894.

[16] Arrué J, Arellano-Valle RB, Gómez HW. Bias reduction of maximum likelihood estimates for a modified skew-normal distribution. *Mathematics*. 2016; 86(15): 2967-2984.

[17] Çetinkaya C, Bulut YM. Bias-reduced and heuristics parameter estimations for the inverse power lindley distribution. *International Journal of Modelling and Simulation*. 2022; 43(5): 600-611.

[18] Lagos Álvarez B, Jiménez Gamero MD. A note on bias reduction of maximum likelihood estimates for the scalar skew t distribution. *Journal of Statistical Planning and Inference*. 2012; 142(2): 608-612.

[19] Lemonte AJ, Cribari-Neto F, Vasconcellos KLP. Improved statistical inference for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis*. 2007; 51(9): 4656-4681.

[20] Ling X, Giles DE. Bias reduction for the maximum likelihood estimator of the parameters of the generalized rayleigh family of distributions. *Communications in Statistics-Theory and Methods*. 2014; 43(8): 1778-1792.

[21] Schwartz J, Giles DE. Bias-reduced maximum likelihood estimation of the zero-inflated poisson distribution. *Communications in Statistics-Theory and Methods*. 2016; 45(2): 465-478.

[22] Zhang J. Reducing bias of the maximum-likelihood estimation for the truncated pareto distribution. *Statistics*. 2013; 47(4): 792-799.

[23] Kosmidis I, Lunardon N. Empirical bias-reducing adjustments to estimating functions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 2024; 86(1): 62-89.

[24] Cox DR, Snell EJ. A general definition of residuals. *Journal of the Royal Statistical Society: Series B*. 1968; 30(2): 248-265.

[25] Ahmed AA, Algamal ZY, Albalawi O. Bias reduction of maximum likelihood estimation in exponentiated teissier distribution. *Frontiers in Applied Mathematics and Statistics*. 2024; 10: 1-7. Available from: https://doi.org/10.3389/fams.2024.1351651.

[26] Hammood NM, Rashad NK, Algamal ZY. Neutrosophic topp-leone extended exponential distribution modeling with application for bladder cancer patients. *International Journal of Neutrosophic Science*. 2025; 25(1): 239-245.

[27] Al-Saqal OE, Hadied ZA, Algamal ZY. Modeling bladder cancer survival function based on neutrosophic inverse gompertz distribution. *International Journal of Neutrosophic Science*. 2025; 25(1): 75-80.