
A New Exponential Gompertz Distribution: Theory and Applications

Ibtessam Ali Alsaggaf

*King Abdulaziz University, Department of Statistics, Faculty of Science, Jeddah,
Saudi Arabia*

E-mail: ialsaggaf@kau.edu.sa

<https://orcid.org/0009-0004-6356-7095>

Received 04 September 2024; Accepted 17 October 2024

Abstract

With the rise of numerous phenomena that require interpretation and investigation, developing novel distributions has become an important need. This research introduced a new probability distribution called New Exponential Gompertz distribution based on the new exponential-X family to enhance flexibility and improve performance. The most significant benefit of this novel distribution is that its hazard function could be increasing, decreasing and bathtub which reflects the flexibility of the distribution to fit various applications. Furthermore, its density can adopt a variety of symmetric and asymmetric possible shapes. Some of the theoretical characteristics such as quantile, order statistic and moment are provided. The parameter estimates are derived using five different estimation methods including maximum likelihood, ordinary least square, weighted least square, Cramér-von mises and maximum product of spacing methods. Simulation studies are conducted to assess the effectiveness of the five estimation methods. The maximum likelihood estimate shows the most reliable estimate for estimating parameters since it provides the smallest mean square error While the maximum

Journal of Reliability and Statistical Studies, Vol. 17, Issue 2 (2024), 289–320.

doi: 10.13052/jrss0974-8024.1722

© 2024 River Publishers

product of spacing method is less efficient. The performance of the proposed distribution is assessed through real-world applications in medical, engineering and physics with competitive distributions. The results indicate that the new distribution efficiently represents various types of data compared to other distributions.

Keywords: Gompertz distribution, NEX family, Rényi entropy, maximum likelihood estimation, Ordinary least square method, weighted least square method, Cramér-von mises method, maximum product of spacing method, simulation.

1 Introduction

The Gompertz distribution is a probability distribution employed in survival analysis to model the distribution of lifetimes. It is a classical distribution that represents survival function based on fatality laws. This distribution is critical for predicting many phenomena related to life. It is essential for estimating death rates for humans and fitting financial data. The Gompertz distribution was proposed by [12]. It has been applied for growth models as well as for predicting tumor growth.

The cumulative function (CDF) of Gompertz is

$$F(x) = 1 - e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)}, x > 0; \lambda, \beta > 0. \quad (1)$$

and its probability density function (PDF) is

$$f(x) = \lambda e^{\beta x} e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \quad (2)$$

Although this distribution was good at representing several phenomena in some fields, its ability to represent many real phenomena is still limited because it tends to exhibit an exponentially rising failure rate over lifetime data. These limitations appear in most classical distributions such as exponential and Weibull distributions due to their limited hazard function shapes, which makes them unsuitable for certain applications. Over the years, several improvements and generalizations have been proposed to make the Gompertz distribution more flexible and useful in a variety of applications. [10] introduced the generalized Gompertz distribution with three parameters using the exponentiated method. Following that, [9] expanded [10] model by adding two more shape parameters to the model using the exponentiated generalized technique proposed by [8]. Aside from developing the Gompertz

distribution, [2] provided Gompertz generalized family based on the $T - X$ transformation technique defined by [4]. The Gompertz Normal, Gompertz Beta, Gompertz Gamma, Gompertz Log-Logistic, Gompertz Exponentiated Weibul and Gompertz Lomax [23] all are developed distributions derived from [2].

[1] combined the Exponentiated method in [10] with Gompertz exponential derived from [2] to propose another generalized Gompertz distribution. This model gives more flexibility in modeling survival data. Another generation of Gompertz is the Kumaraswamy-G generalized Gompertz distribution which was proposed by [11]. [24] provided the three parameters Gamma-Gompertz model which was derived by the gamma-X family. In 2024, [6] combined the odd Weibull family with the inverse Gompertz distribution to introduce a new family that merges the features of both distributions. At same year, [15] employed $T - X$ transformation to introduce a new model called the generalized Gompertz-G family. The primary objective of these efforts is to optimize the Gompertz distribution fitting in a way that it can seamlessly handle diverse real-world data across various domains. These generalizations were applied to a variety of data in engineering, medicine, physics, and other fields and various phenomena.

In a new attempt to improve Gompertz’s efficiency, this research aims to utilize an alternative method for generalizing distributions. In this study, a new generalization for the Gompertz distribution is developed by employing the new lifetime exponential-X family (NEX) proposed by [14]. This family is distinguished by its provision of a diverse array of hazard function forms.

The NEX family’s CDF and PDF are provided by:

$$G(x) = 1 - [1 - F(x)] e^{-\theta F(x)}, x > 0; \theta, > 0. \tag{3}$$

$$g(x) = f(x) \left[\frac{1 + \theta [1 - F(x)]}{e^{\theta F(x)}} \right] \tag{4}$$

By integrating the attributes of this family with those of the Gompertz distribution to represent reliability data, we can create a new distribution that encompasses all of these features. The aim of this article is to introduce a novel extension of the Gompertz distribution, known as the New Exponential Gompertz distribution (NEG), based on the NEX distribution family. The NEG distribution has several benefits:

- The NEG distribution expands on the Gompertz distribution by adding new parameters, which enhances its flexibility and allows for a more precise representation of different tail shapes.

- It increases the adaptability of density and hazard rate functions, enabling precise modeling of diverse real-world scenarios.
- It offers greater flexibility than the Gompertz distribution, resulting in a superior fit compared to other distributions.
- The CDF, hazard rate functions, moments, and entropy of the NEG distribution are all provided in closed form, making it a valuable tool for analyzing both complete and censored data.
- The hazard rate function for the NEG distribution exhibits a wide range of forms, including symmetrical and asymmetrical shapes in its density function. This flexibility enables NEG to effectively represent and analyze a diverse array of data from fields such as engineering, medicine, physics, and reliability.

This article is classified as follows: Section 2 describes the NEG using graphical representations. Section 2.1 gives useful expansion for the NEG's CDF and PDF. Section 3 derived statistical properties for NEG. In section 4, five estimation methods are applied to estimate the NEG parameters: maximum likelihood (ML), ordinary least squares (O.LS), weighted least squares (W.LS), Cramér-von Mises (CRM) and maximum product of spacing (MPS). Section 5 presents assessing the performance of the estimation methods using Monte Carlo simulation technique. Section 6 analyzes various datasets of cancer patient data to evaluate the modeling effectiveness of the EFG distribution and compare it performance against competitive distributions. Section 7 concludes with some final remarks.

2 The New Exponential Gompertz Distribution (NEG)

The NEG's CDF and PDF are founded by replacing the $G(x)$ and $g(x)$ in (3) and (4) by (1) and (2) as follows:

$$F(x) = 1 - \left[e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \right] \left[e^{-\theta \left(1 - e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \right)} \right], \quad (5)$$

$$f(x) = \lambda e^{\beta x} e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \left[\frac{1 + \theta \left[e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \right]}{e \left[1 - e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \right]} \right], \quad (6)$$

where $\lambda, \beta, \theta, > 0, x \geq 0$.

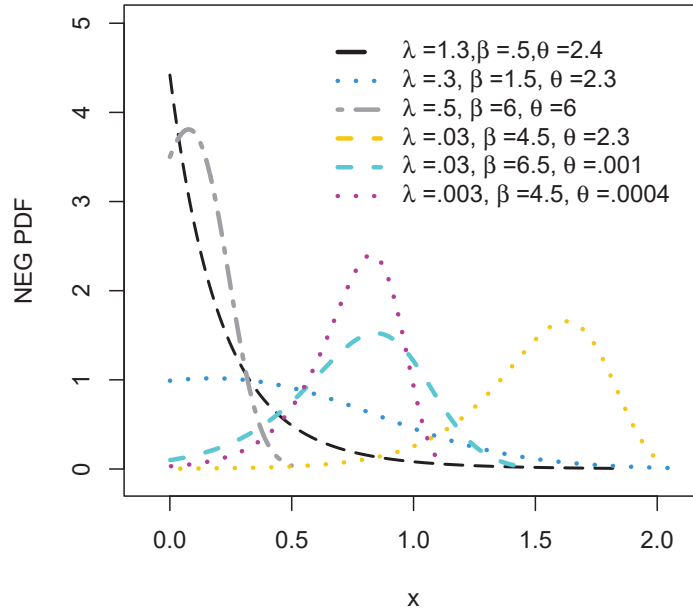


Figure 1 The density plots of $NEG(\lambda, \beta, \theta)$.

The Survival, $S(x)$ of NEG is expressed as

$$S(x) = \left[e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \right] \left[e^{-\theta \left(1 - e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \right)} \right], \tag{7}$$

and its hazard rate functions, $hrf(x)$, is written as

$$hrf(x) = \lambda e^{\beta x} \left[1 + \theta e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \right]. \tag{8}$$

Figures 1 and 2 display several types of curves for NEG’s density and hrf at different combinations of the parameter values. Figure 1 shows the distinct shapes of the density distribution of NEG, as it oscillates from skewed to the right to skewed to the left. It also takes the J-shaped, in addition to the symmetrical shape.

Regarding the hazard rate function of NEG, Figure 2 shows that the hazard function has a wide range of shapes, including increasing, decreasing, J-shaped, and inverse J-shaped. This indicates the high flexibility of the distribution, which indicates the possibility of its compatibility with many phenomena.

A special sub-model of NEG is Gompertz distribution (2) when $\theta = 0$.

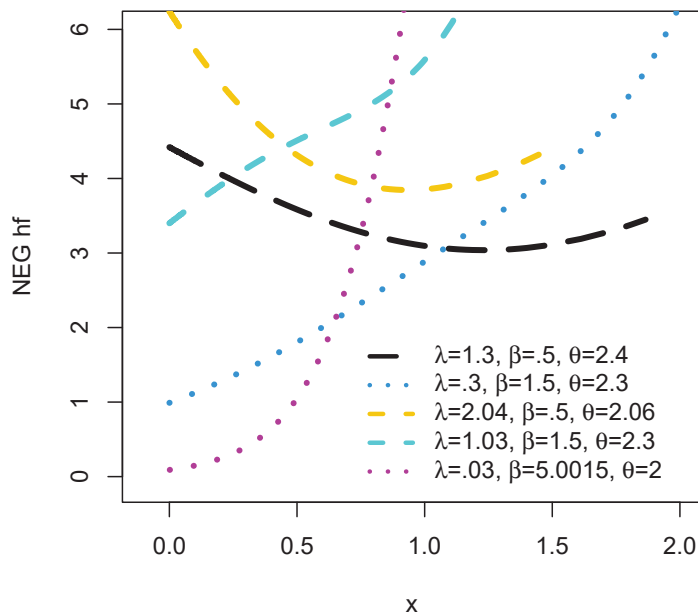


Figure 2 The hazard plots for NEG (λ, β, θ).

2.1 Additional Expression for the NEG’s CDF and PDF

Expansion for the NEG’s CDF and PDF given in Equations (5) and (6), respectively, are delivered in this subsection.

Using exponential expansion, see [22]:

$$e^{-z} = \sum_{k=0}^{\infty} \frac{(-1)^k (z)^k}{k!}, \tag{9}$$

The NEG’s PDF might then be written as

$$F(x) = \sum_{k_1=0}^{\infty} \frac{(-1)^{k_1}}{k_1!} \theta^{k_1} e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \left[1 - e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \right]^{k_1}$$

Further, applying the binomial series (10) for $m > 0$ and $|r| < 1$

$$(1 - r)^n = \sum_{m=0}^{\infty} (-1)^m \binom{n}{m} r^m \tag{10}$$

Thus, the NEG's CDF can be reduced to

$$F(x) = 1 - \eta_1 \left[1 - e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \right]^{k_1 + m_1}, \quad (11)$$

where

$$\eta_1 = \sum_{k_1=0}^{\infty} \sum_{m_1=0}^1 \frac{(-1)^{k_1+m_1}}{k_1!} \binom{1}{m_1} \theta^{k_1} \quad (12)$$

To find the expansion for the NEG'S PDF, the series in Equations (9), (10) and (13) are employed.

$$(1 + u)^l = \sum_{i=0}^{\infty} \binom{l}{i} u^i. \quad (13)$$

Therefore, the NEG's PDF can be reduced to

$$f(x) = \eta_2 e^{\beta(1+k_3)x}, \quad (14)$$

where

$$\eta_2 = \sum_{i_1=0}^1 \sum_{k_2=0}^{\infty} \sum_{m_2=0}^j \sum_{k_3=0}^{\infty} \frac{(-1)^{m_2+k_3}}{k_2!k_3!} \binom{1}{i_1} \binom{n}{k_2} \left(\frac{\lambda^{k_3+1}}{\beta^{k_3}} \right) (m_2 + i_1 + 1)^{k_3} \theta^{(i_1+k_2)} e^{\frac{\lambda}{\beta}(m_2+i_1+1)} \quad (15)$$

3 Statistical Properties of the NEG

Some fundamental properties of NEG such as quantile, moment, order statistic and more are driven in this section.

3.1 Quantile Function and Quartiles

The quantile function (x_p) of the NEG might be written as

$$x_p = \frac{\ln}{\beta} \left[1 - \frac{\beta}{\lambda} \ln \left(\frac{W(\theta e^{\theta}(1-p))}{\theta} \right) \right], \quad 0 < p < 1. \quad (16)$$

where $W[\cdot]$ is the Lambert function.

The median of the NEG distribution can be obtained as

$$x_{0.50} = \frac{\ln}{\beta} \left[1 - \frac{\beta}{\lambda} \ln \left(\frac{W(\theta e^{\theta}(0.5))}{\theta} \right) \right]$$

Hence, the 25th percentile and the 75th percentile of the NEG distribution are given as

$$x_{0.25} = \frac{\ln}{\beta} \left[1 - \frac{\beta}{\lambda} \ln \left(\frac{W(\theta e^{\theta}(0.75))}{\theta} \right) \right],$$

$$x_{0.75} = \frac{\ln}{\beta} \left[1 - \frac{\beta}{\lambda} \ln \left(\frac{W(\theta e^{\theta}(0.25))}{\theta} \right) \right].$$

3.2 Shape Indices

The shape of the NEG can be measured by Galton’s skewness and Moors’ kurtosis [20] which are obtained by applying (16) and respectively given as follows:

$$Skewness = \frac{q_{0.75} - 2q_{0.5} + q_{0.25}}{q_{0.75} - q_{0.25}}, \tag{17}$$

and

$$Kurtosis = \frac{q_{0.875} - q_{0.625} + q_{0.375} - q_{0.125}}{q_{0.75} - q_{0.25}}. \tag{18}$$

For details see [3].

Table 1 shows some values of the mean, median (Q_2), first quartile (Q_1), third quartile (Q_3), skewness, and kurtosis for the NEG, for some different values the parameters λ , β and θ . The table reveals a wide range of values represented by the distribution, with the mean and quartiles spanning from values below one to significantly larger values. Additionally, the measures of skewness indicate that the distribution can exhibit different degrees of skewness, ranging from right-skewed to left-skewed. The kurtosis values also

Table 1 Mean, median, quartiles, skewness and kurtosis of the NEG for various values of λ , β and θ .

λ	β	θ	Q_1	Q_2	Mean	Q_3	Skewness	Kurtosis
1.3	0.5	2.4	0.066	0.159	0.239	0.327	2.44	11.55
0.002	0.05	0.4	37.876	53.678	52.123	66.853	-0.24	2.577
3	2.2	1.6	0.037	0.087	0.125	0.168	1.773	6.641
9	10	3	0.009	0.02	0.027	0.036	1.907	8.149
0.03	2.7	7	0.556	0.84	0.822	1.091	-0.11	2.359
0.05	0.008	0.04	5.263	12.129	16.365	23.118	1.489	5.707
0.003	3.5	2.6	1.321	1.571	1.529	1.793	-0.584	3.444
0.33	0.005	0.004	0.876	1.991	2.907	3.951	1.932	7.791
0.55	0.89	0.27	0.35	0.719	0.824	1.201	0.721	2.909
1.5	3.2	4.5	0.034	0.077	0.107	0.153	1.598	6.368

indicate varying levels of tail width. Specifically, when the lambda parameter is less than 0.01, the distribution tends to be skewed to the left.

3.3 Moments

If X follows the NEG (λ, β, θ) , then the r^{th} moment of X is written as

$$E(x^r) = \int_0^\infty x^r f(x) dx = \eta_2 \int_0^\infty x^r e^{\beta(1+k_3)x} dx, \quad (19)$$

where η_2 is given by (15). Using Laplace transformation, $L_t[f(t)](s) = \int_0^\infty f(t)e^{-st} dt$ where $f(t)$ is defined for $t \geq 0$ [17], with taking $f(t) = t^r$ then $L_t[f(t)](s) = \frac{r!}{s^{r+1}}$, where $s = -\beta(1+k_3)$. Therefore,

$$E(x^r) = \eta_2 r! \left(\frac{-1}{\beta(1+k_3)} \right)^{r+1}$$

Thus, the r^{th} moment is expressed as

$$\mu_r = E(x^r) = \eta_2 \left(\frac{-1}{\beta(1+k_3)} \right)^{r+1} \Gamma(r+1), r \geq 0 \quad (20)$$

Then, the NEG's mean is written as

$$\mu = E(x) = \eta_2 \left(\frac{1}{\beta(1+k_3)} \right)^2 \quad (21)$$

The NEG's variance is determined by

$$\sigma^2 = E(x^2) - \mu^2 = 2\eta_2 \left(\frac{-1}{\beta(1+k_3)} \right)^3 - \mu^2,$$

where η_2 is given by (15).

3.4 Moment Generating Function

If X follows the NEG (λ, β, θ) , then the NEG's moment generating function (MGF) is written as

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x) dx = \eta_2 \int_0^\infty e^{tx} e^{\beta(1+k_3)x} dx, \quad (22)$$

Using Laplace transformation, with taking $f(t) = 1$ then $L_t[f(t)](s) = \frac{1}{s}$, where $s = -(t + \beta(1 + k_3))$, then MGF will be given as

$$M_x(t) = E(e^{tx}) = \eta_2 \left(\frac{-1}{t + \beta(1 + k_3)} \right), \quad (23)$$

where η_2 is given by (15).

3.5 Characteristic Function

The characteristic function of NEG is simply constructed as:

$$\phi_x(t) = E(e^{itx}) = \eta_2 \left(\frac{-1}{it + \beta(1 + k_3)} \right), \quad (24)$$

where η_2 is given by (15).

3.6 Mean Residual Life and Mean Waiting Time

If $X \sim NEG(\lambda, \beta, \theta)$ with $S(t)$ provided in (7), then the mean residual life, $\mu(t)$, is expressed as

$$\mu(t) = \frac{1}{S(t)} \left(E(t) - \int_0^t xf(x)dx \right) - t. \quad (25)$$

If the incomplete moment, $I_{inc} = \int_0^t xf(x)dx$, then

$$I_{inc} = \int_0^t \eta_2 x e^{\beta(1+k_3)x} dx.$$

Setting $y = \beta(1 + m)x$, then simplifying will be obtained,

$$I_{inc} = \left(\frac{\eta_2}{\beta^2(1 + k_3)^2} \right) \int_0^{t\beta(1+k_3)} ye^y dy,$$

Using integration by part by taking $u = y$ and $dv = e^y dy$, the incomplete moment will be obtained as

$$I_{inc} = \left(\frac{\eta_2}{\beta^2(1 + k_3)^2} \right) \left[\beta t(1 + k_3)e^{\beta t(1+k_3)} - e^{\beta t(1+k_3)} + 1 \right], \quad (26)$$

Substituting (21), and (26) in (25), $\mu(t)$ might be rewritten as

$$\mu(t) = \frac{1}{S(t)} \frac{\eta_2}{\beta^2(1 + k_3)^2} \left[e^{\beta t(1+k_3)} - \beta t(1 + k_3)e^{\beta t(1+k_3)} \right] - t, \quad (27)$$

In a similar manner, the mean waiting time, $\bar{\mu}(t)$, might be defined as

$$\bar{\mu}(t) = t - \frac{1}{F(t)} \int_0^t x f(x) dx, \tag{28}$$

where $F(t)$ is provided in (11). Then, $\bar{\mu}(t)$ of the NEG can be found by substituting (11) and (26) in (28) as follows

$$\bar{\mu}(t) = t - \frac{\eta_2}{\beta^2(1+k_3)^2} \left[\frac{\beta t(1+k_3)e^{\beta t(1+k_3)} - e^{\beta t(1+k_3)} + 1}{1 - \eta_1 \left[1 - e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \right]^{k_1+m_1}} \right]$$

3.7 Rényi entropy

$RE_X(\zeta)$ is the Rényi entropy function which given as

$$RE_X(\zeta) = \frac{1}{1-\zeta} \log \left(\int_0^\infty f(x)^\zeta dx \right); \quad \zeta > 0, \zeta \neq 1.$$

Then, applying the NEG's PDF in (6)

$$f(x)^\zeta = \lambda^\zeta e^{\zeta\beta x} e^{-\frac{\zeta\lambda}{\beta}(e^{\beta x}-1)} \left[1 + \theta \left(e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \right) \right]^\zeta \left[e^{-\zeta\theta \left(1 - e^{-\frac{\lambda}{\beta}(e^{\beta x}-1)} \right)} \right].$$

Applying the same approach in Subsection 2.1 and using (9), (13) and (10), then

$$f(x)^\zeta = \eta_2^* e^{(\zeta+k_5)\beta x},$$

where

$$\eta_2^* = \sum_{i_2=0}^{\zeta} \sum_{k_4=0}^{\infty} \sum_{m_3=0}^j \sum_{k_5=0}^{\infty} \frac{(-1)^{j+k_5+m_3}}{j!k_5!} \binom{\zeta}{i_2} \binom{j}{m_3} \left(\frac{\lambda^{\zeta+k_5}}{\beta^{k_5}} \right) (\zeta + m_3 + i_2)^{k_5} \theta^{(i_2+k_4)} e^{\frac{\zeta\lambda}{\beta}}.$$

Using Laplace transformation, since $f(t) = 1$ then $L_t[f(t)](s) = \frac{1}{s}$, where $s = -(\zeta + k_5)\beta$, then the Rényi entropy of the NEG, is then will be reduced to

$$RE_x(\zeta) = \frac{1}{1-\zeta} \log \left[\frac{-\eta_2^*}{(\zeta + k_5)\beta} \right].$$

3.8 Order Statistics

The density function, $f_{j:n}(x)$, of the j^{th} order statistics is given as

$$f_{j:n}(x) = \frac{1}{B(j, n-j+1)} f(x) [F(x)]^{j-1} [1-F(x)]^{n-j}. \quad (29)$$

By employing the binomial series formula in (10) to (29), $f_{j:n}(x)$ can be expressed as

$$f_{j:n}(x) = \frac{1}{B(j, n-j+1)} \sum_{m_3=0}^{n-j} (-1)^{m_3} \binom{n-j}{m_3} f(x) [F(x)]^{m_3+j-1}. \quad (30)$$

By substituting the CDF (11) and PDF of NEG (14) into (30), the PDF of $X_{j:n}$ is

$$f_{j:n}(x) = \frac{[\eta_2 e^{\beta(1+k_3)x}]}{B(j, n-j+1)} \sum_{m_3=0}^{n-j} (-1)^{m_3} \binom{n-j}{m_3} \left[1 - \eta_1 \left[1 - e^{-\frac{\lambda}{\beta}(e^{\beta x} - 1)} \right]^{k_1+m_1} \right]^{m_3+j-1} \quad (31)$$

where η_1 and η_2 are given by (12) and (15), respectively.

4 Estimation Methods

4.1 Maximum Likelihood Estimation (ML)

Assume x_1, \dots, x_n is a random sample from $NEG(\lambda, \beta, \theta)$ with size n . Therefore, the log-likelihood function, (ℓ) , for (λ, β, θ) , is written as

$$\begin{aligned} \ell(\lambda, \beta, \theta) &= \sum_{i=1}^n \ln \left(e^{-\frac{\lambda}{\beta}(e^{\beta x_i} - 1)} \theta + 1 \right) - \theta \sum_{i=1}^n \left(1 - e^{-\frac{\lambda}{\beta}(e^{\beta x_i} - 1)} \right) \\ &\quad - \sum_{i=1}^n \frac{\lambda}{\beta} \left(e^{\beta x_i} - 1 \right) + \beta \sum_{i=1}^n x_i + n \ln(\lambda) \end{aligned} \quad (32)$$

The derivation for the Equation (32) in regard to λ, β and θ , are written as follows:

$$\frac{\partial \ell}{\partial \lambda} = -\frac{\theta}{\beta} \sum_{i=1}^n \left[\frac{(e^{\beta x_i} - 1) e^{-\frac{\lambda}{\beta}(e^{\beta x_i} - 1)}}{\theta e^{-\frac{\lambda}{\beta}(e^{\beta x_i} - 1)} + 1} \right]$$

$$-\frac{\theta}{\beta} \sum_{i=1}^n \left[\left(e^{\beta x_i} - 1 \right) e^{-(e^{\beta x_i} - 1)\lambda} \right] + \frac{n}{\lambda} - \frac{1}{\beta} \cdot \sum_{i=1}^n (e^{\beta x_i} - 1), \quad (33)$$

$$\begin{aligned} \frac{\partial \ell}{\partial \beta} = & \sum_{i=1}^n \left[\frac{\frac{\theta \lambda}{\beta} \left(\frac{(e^{x_i \beta} - 1)}{\beta} - x_i e^{x_i \beta} \right) e^{-\frac{\lambda}{\beta} (e^{x_i \beta} - 1)}}{\theta e^{-\frac{\lambda}{\beta} (e^{x_i \beta} - 1)} + 1} \right] \\ & + \frac{\theta \lambda}{\beta} \sum_{i=1}^n \left[\left(\frac{(e^{x_i \beta} - 1)}{\beta} - x_i e^{x_i \beta} \right) e^{-\frac{\lambda}{\beta} (e^{x_i \beta} - 1)} \right] \\ & - \frac{\lambda}{\beta} \sum_{i=1}^n (x_i e^{x_i \beta}) + \frac{\lambda}{\beta^2} \sum_{i=1}^n (e^{x_i \beta} - 1) + \sum_{i=1}^n x_i, \end{aligned} \quad (34)$$

$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^n \left[\frac{e^{-\frac{\lambda}{\beta} (e^{\beta x_i} - 1)}}{\theta e^{-\frac{\lambda}{\beta} (e^{\beta x_i} - 1)} + 1} \right] + \sum_{i=1}^n \left[e^{-\frac{\lambda}{\beta} (e^{\beta x_i} - 1)} \right] - n, \quad (35)$$

Therefore, the estimate for each parameter can be derived by setting (33), (34) and (35) to zero and numerically solving by Newton–Raphson iteration method which is available in R program. Furthermore, the log-likelihood present in (32) can be alternatively optimized using a non-linear optimization tool.

4.2 Ordinary Least Square Method (O.LS)

O.LS method is proposed by [25], which is based on the difference between the empirical and theoretical cdf. Suppose a random sample from NEG distribution with size n and $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ are its order statistics. The sum of squares for the difference between the empirical and theoretical cdf of NEG distribution is formulated in Equation (36).

$$O.LS(\Theta) = \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right]^2 \quad (36)$$

where the $F(x_{(k)})$ is the cdf of NEG and $\left(\frac{k}{n+1} \right)$ is the empirical cdf and $i = 1, 2, \dots, n$.

The partial derivation from the Equation (36) in regard to $\Theta = (\lambda, \beta, \theta)$ can be written as follows:

$$\frac{\partial O.LS(\Theta)}{\partial \lambda} = \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \lambda} \quad (37)$$

$$\frac{\partial O.LS(\Theta)}{\partial \beta} = \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \beta} \quad (38)$$

$$\frac{\partial O.LS(\Theta)}{\partial \theta} = \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \theta} \quad (39)$$

where the $\frac{\partial F(x_{(k)}; \Theta)}{\partial \lambda}$, $\frac{\partial F(x_{(k)}; \Theta)}{\partial \beta}$ and $\frac{\partial F(x_{(k)}; \Theta)}{\partial \theta}$ are given as

$$\begin{aligned} \frac{\partial F(x_{(k)}; \Theta)}{\partial \lambda} = & - \left(- \frac{(e^{x\beta} - 1) \theta e^{-\frac{(e^{x\beta} - 1)\lambda}{\beta}}}{\beta} - \frac{e^{x\beta} - 1}{\beta} \right) \\ & - \theta \left(1 - e^{-\frac{(e^{x\beta} - 1)\lambda}{\beta}} \right) - \frac{(e^{x\beta} - 1)\lambda}{e} \end{aligned} \quad (40)$$

$$\begin{aligned} \frac{\partial F(x_{(k)}; \Theta)}{\partial \beta} = & - \left(\theta \left(\frac{\lambda (e^{x\beta} - 1)}{\beta^2} - \frac{x\lambda e^{x\beta}}{\beta} \right) \right. \\ & \left. e^{-\frac{\lambda(e^{x\beta} - 1)}{\beta}} - \frac{x\lambda e^{x\beta}}{\beta} + \frac{\lambda (e^{x\beta} - 1)}{\beta^2} \right) \\ & - \theta \left(1 - e^{-\frac{\lambda(e^{x\beta} - 1)}{\beta}} \right) - \frac{\lambda(e^{x\beta} - 1)}{e} \end{aligned} \quad (41)$$

$$\frac{\partial F(x_{(k)}; \Theta)}{\partial \theta} = - \left(e^{-\frac{(e^{x\beta} - 1)\lambda}{\beta}} - 1 \right) e^{-\frac{(e^{x\beta} - 1)\lambda}{\beta}} \theta - \frac{(e^{x\beta} - 1)\lambda}{e} \quad (42)$$

The O.LS estimates for the parameters Θ can be obtained by minimizing the Equation (36) concerning the $\Theta = (\lambda, \gamma, \theta, \beta, \alpha)$ or by solving the equations from (37) to (39) using numerical techniques available in statistical software.

4.3 Weighted Least Square Method (W.LS)

WLS method is similar to OLS method, which depends on the differences between the empirical and theoretical of the cdf, in addition to the variance of the order statistic as a wight [25]. Therefore, the WLS function for NEG is written as

$$W.LS(\Theta) = \sum_{k=1}^n \omega_k \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right]^2 \quad (43)$$

where $\omega_k = \frac{(n+1)^2(n+2)}{k(n+1-k)}$

The Equation (43) is derived for each parameter in NEG distribution and the derivation equations are given as follows.

$$\frac{\partial W.LS(\Theta)}{\partial \lambda} = \sum_{k=1}^n \omega_k \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \lambda} \quad (44)$$

$$\frac{\partial W.LS(\Theta)}{\partial \beta} = \sum_{k=1}^n \omega_k \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \beta} \quad (45)$$

$$\frac{\partial W.LS(\Theta)}{\partial \theta} = \sum_{k=1}^n \omega_k \left[F(x_{(k)}) - \left(\frac{k}{n+1} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \theta} \quad (46)$$

where the $\frac{\partial F(x_{(k)}; \Theta)}{\partial \lambda}$, $\frac{\partial F(x_{(k)}; \Theta)}{\partial \beta}$ and $\frac{\partial F(x_{(k)}; \Theta)}{\partial \theta}$ are given by (40), (41) and (42).

The W.LS estimation is obtained by minimizing the Equation (43) or by solving the nonlinear equations from (44) to (46) using numerical iterative technique available in any statistical software.

4.4 Cramér-von Mises Method (CRM)

[19] introduced the CRM method for estimation parameters. The function of the CRM method for NEG distribution is presented in Equation (47).

$$CRM(\Theta) = \frac{1}{12n} + \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{2k-1}{2n} \right) \right]^2 \quad (47)$$

The equations below show the partial derivatives of the parameters $\Theta = (\lambda, \beta, \theta)$ from (47). The CRM estimates for NEG parameters are derived

by solving the equations from (48) to (50) using numerical technique or by minimizing (47) using optimization technique available in R Package.

$$\frac{\partial CRM(\Theta)}{\partial \lambda} = \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{2k-1}{2n} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \lambda} \quad (48)$$

$$\frac{\partial CRM(\Theta)}{\partial \beta} = \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{2k-1}{2n} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \beta} \quad (49)$$

$$\frac{\partial CRM(\Theta)}{\partial \theta} = \sum_{k=1}^n \left[F(x_{(k)}) - \left(\frac{2k-1}{2n} \right) \right] \frac{\partial F(x_{(k)}; \Theta)}{\partial \theta} \quad (50)$$

where the $\frac{\partial F(x_{(k)}; \Theta)}{\partial \lambda}$, $\frac{\partial F(x_{(k)}; \Theta)}{\partial \beta}$ and $\frac{\partial F(x_{(k)}; \Theta)}{\partial \theta}$ are given by (40), (41) and (42).

4.5 Maximum Product of Spacing Method (MPS)

[7] proposed MPS method in order to improve the performance of the ML estimator. Let a random sample from NEG distribution with size n and $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ is the corresponding ordered sample. The idea of MPS method is to optimize the geometric mean of spacings, which refers to the variations between the cdf values of adjacent data points. The spacings between neighboring ordered values can be defined as $D_k(\Theta) = F(x_{(k)}; \Theta) - F(x_{(k-1)}; \Theta)$, $\Theta = (\lambda, \gamma, \theta, \beta, \alpha)$, $k = 1, \dots, n+1$.

Therefore, the MPS function for NEG is given as

$$M(\Theta) = \frac{1}{n+1} \sum_{k=1}^{n+1} \log D_k(\Theta)$$

$$M(\Theta) = \frac{1}{n+1} \sum_{k=1}^{n+1} \log [F(x_{(k)}) - F(x_{(k-1)})] \quad (51)$$

The first partial derivative of (51) with respect to $\Theta = (\lambda, \beta, \theta)$, are given as follows. The MPS estimates for NEG parameters are derived by solving the equations from (52) to (54) using numerical technique or via maximizing (51) using optimization technique available in R Package.

$$\frac{\partial M(\Theta)}{\partial \lambda} = \frac{1}{n+1} \sum_{k=1}^{n+1} \left[\frac{F'_\lambda(x_{(k)}) - F'_\lambda(x_{(k-1)})}{F(x_{(k)}) - F(x_{(k-1)})} \right] \quad (52)$$

$$\frac{\partial M(\Theta)}{\partial \beta} = \frac{1}{n+1} \sum_{k=1}^{n+1} \left[\frac{F'_{\beta}(x_{(k)}) - F'_{\beta}(x_{(k-1)})}{F(x_{(k)}) - F(x_{(k-1)})} \right] \quad (53)$$

$$\frac{\partial M(\Theta)}{\partial \theta} = \frac{1}{n+1} \sum_{k=1}^{n+1} \left[\frac{F'_{\theta}(x_{(k)}) - F'_{\theta}(x_{(k-1)})}{F(x_{(k)}) - F(x_{(k-1)})} \right] \quad (54)$$

where the $F'_{\lambda}(\cdot)$, $F'_{\beta}(\cdot)$ and $F'_{\theta}(\cdot)$ are given by (40), (41) and (42).

5 Simulation Study

Five techniques of estimation including ML, O.LS, W.LS, CRM and MPS are performed to calculate the estimation for the NEG parameters using Monte Carlo Simulation. Different sample sizes and three different sets of parameter values are applied.

The samples are drawn from the NEG distribution of size 15, 30, 50, 100 and 200 for the following three parameter sets:

- Set I: ($\lambda = 1.3$, $\beta = 0.5$, $\theta = 2.4$).
- Set II: ($\lambda = 0.002$, $\beta = 0.05$, $\theta = 0.4$).
- Set III: ($\lambda = 3$, $\beta = 2.2$, $\theta = 1.6$).

The sample of size n is generated $N=1000$ times. For each sample, the parameters estimates are calculated using the five estimation methods. Then, the average estimates, bias and the mean squared error (MSE) are calculated for each parameter.

Table 2 to 4 display the result of the simulation. The tables show ML, O.LS, W.LS, CRM and MPS estimate values for each parameter and corresponding bias and MSE. The results in the tables show that, as the sample size increases, the MSE generally decreases for all five methods, indicating improved accuracy with more data. Additionally, parameter estimates themselves tend to converge towards the true values. The performance of the methods O.LS, W.LS, CRM and MPS are similar in terms of the MSE values. However, the MPS is considered less efficient as it has large MSE values at some parameter estimation. Out of all the estimators, the ML estimator has the lowest MSE values. Consequence of this, ML is the most reliable choices for estimating NEG parameters.

Table 2 Parameter estimation, MSE and Bias from five different methods (Set I).

		Set I: ($\lambda=1.3, \beta=0.5, \theta=2.4$)				
n	Par.	ML	O.LS	W.LS	CRM	PMS
15	$\hat{\lambda}$	2.0315	2.0965	2.0287	1.9939	1.9659
	Bias	0.9380	1.3056	1.28572	1.2125	1.2459
	MSE	1.1828	1.7872	1.6709	1.7411	1.5485
	$\hat{\beta}$	1.3751	0.3436	0.5011	1.4022	0.1726
	Bias	1.2476	2.0975	1.9022	2.3226	1.3904
	MSE	1.8491	2.9236	2.5878	3.5417	1.7854
	$\hat{\theta}$	1.2539	2.7795	3.1570	2.5601	7.7718
	Bias	1.219	2.1898	2.5882	2.0646	7.4249
	MSE	1.3448	3.3806	4.4675	3.3194	17.3591
30	$\hat{\lambda}$	1.9242	1.9454	1.8694	1.9698	1.7217
	Bias	0.7500	1.1428	1.1266	1.0991	1.1011
	MSE	0.9908	1.4746	1.4421	1.5316	1.3023
	$\hat{\beta}$	0.9335	0.3171	0.4352	0.7756	0.0186
	Bias	0.7803	1.3455	1.1854	1.3583	0.9379
	MSE	1.1328	1.7181	1.5599	1.8139	1.1376
	$\hat{\theta}$	1.4466	2.6318	2.9034	2.2815	8.0858
	Bias	1.0605	1.9492	2.1671	1.6949	7.3997
	MSE	1.2061	2.6649	3.1662	2.3049	15.5624
50	$\hat{\lambda}$	1.8082	1.8812	1.7861	1.9048	1.6348
	Bias	0.6432	1.04919	1.0433	1.0238	1.0042
	MSE	0.8075	1.2911	1.2959	1.2896	1.1727
	$\hat{\beta}$	0.7495	0.3324	0.3750	0.6132	0.0746
	Bias	0.5563	0.9945	0.8434	0.9970	0.7198
	MSE	0.7799	1.2617	1.09143	1.3045	0.8739
	$\hat{\theta}$	1.6319	2.4849	2.8776	2.2484	6.6048
	Bias	0.9846	1.7485	2.0619	1.5776	5.7493
	MSE	1.1212	2.1691	2.7127	1.9714	12.0264
100	$\hat{\lambda}$	1.7262	1.8554	1.7409	1.8517	1.5115
	Bias	0.5724	0.9762	0.9498	0.9502	0.8544
	MSE	0.7389	1.2304	1.1533	1.2174	1.0234
	$\hat{\beta}$	0.6012	0.3280	0.3522	0.4700	0.1605
	Bias	0.3733	0.6932	0.5741	0.6822	0.5071
	MSE	0.4806	0.8657	0.7269	0.8684	0.6213
	$\hat{\theta}$	1.8060	2.3345	2.6560	2.2417	5.0841
	Bias	0.8842	1.5409	1.7784	1.4465	3.9542
	MSE	1.0341	1.7584	2.0669	1.6683	8.1207
200	$\hat{\lambda}$	1.4833	1.8526	1.7399	1.8444	1.4345
	Bias	0.3315	0.9244	0.9126	0.8983	0.7449
	MSE	0.4550	1.1442	1.1119	1.1219	0.9028

continued on next page

Table 2 Parameter estimation (Sete I) – (continue)

		Set I: ($\lambda=1.3, \beta=0.5, \theta=2.4$)				
n	Par.	ML	O.LS	W.LS	CRM	PMS
	$\hat{\beta}$	0.5674	0.3237	0.3183	0.3989	0.2195
	Bias	0.2551	0.5019	0.3972	0.4914	0.3731
	MSE	0.3263	0.6166	0.4919	0.6083	0.4558
	$\hat{\theta}$	2.1560	2.1987	2.5417	2.1440	4.0119
	Bias	0.6403	1.4297	1.6477	1.3605	2.6925
	MSE	0.7821	1.5585	1.8109	1.4949	4.8025

Table 3 Parameter estimation, MSE and Bias from five different methods (Set II)

		Set II: ($\lambda=0.002, \beta=0.05, \theta=0.4$)				
n	Par.	ML	O.LS	W.LS	CRM	PMS
15	$\hat{\lambda}$	0.0022	0.0032	0.0031	0.0026	0.0033
	Bias	0.0005	0.0022	0.0022	0.0018	0.0021
	MSE	0.0009	0.0029	0.0029	0.0025	0.0029
	$\hat{\beta}$	0.0496	0.0459	0.0466	0.0527	0.0439
	Bias	0.0007	0.0139	0.0132	0.0142	0.0119
	MSE	0.0033	0.0177	0.0167	0.01913	0.0146
	$\hat{\theta}$	0.4000	0.8008	0.8341	0.7595	0.7724
	Bias	7.1×10^{-6}	0.9919	1.0302	0.9396	0.9411
	MSE	2.9×10^{-5}	1.1804	1.2266	1.1158	1.1818
30	$\hat{\lambda}$	0.0021	0.0029	0.0028	0.0025	0.0028
	Bias	0.0003	0.0018	0.0018	0.0016	0.0017
	MSE	0.0005	0.0024	0.0022	0.0019	0.0022
	$\hat{\beta}$	0.04980	0.0468	0.0475	0.0502	0.0457
	Bias	0.0002	0.0103	0.0098	0.0101	0.0089
	MSE	0.0016	0.0128	0.0119	0.0128	0.0109
	$\hat{\theta}$	0.4000	0.7592	0.8273	0.7509	0.9064
	Bias	1.4×10^{-6}	0.9801	1.0462	0.9429	1.0548
	MSE	9.8×10^{-6}	1.1479	1.2502	1.1083	1.3762
50	$\hat{\lambda}$	0.0021	0.0027	2.1517	0.0026	0.0027
	Bias	0.0002	0.0016	0.6994	0.0015	0.0016
	MSE	0.0004	0.0020	1.1158	0.0019	0.0021
	$\hat{\beta}$	0.0498	0.0476	0.3849	0.0489	0.0464
	Bias	0.0001	0.0082	0.1398	0.0083	0.0077
	MSE	0.0011	0.0099	0.1966	0.0103	0.0092
	$\hat{\theta}$	0.4000	0.8056	1.3172	0.6798	0.9524
	Bias	9.7×10^{-7}	1.0022	0.6025	0.8615	1.0896
	MSE	9.5×10^{-6}	1.2436	0.8795	1.0358	1.5024
100	$\hat{\lambda}$	0.0020	0.0026	0.0025	0.0025	0.0024

continued on next page

Table 3 Parameter estimation (Set II) – (continue)

Set II: ($\lambda=0.002, \beta=0.05, \theta=0.4$)							
n	Par.	ML	O.LS	W.LS	CRM	PMS	
	Bias	0.0001	0.0015	0.0014	0.0014	0.0014	
	MSE	0.0002	0.0018	0.0018	0.0017	0.0018	
	$\hat{\beta}$	0.0499	0.0476	0.0483	0.0487	0.0477	
	Bias	0.0001	0.0069	0.0066	0.0067	0.0062	
	MSE	0.0002	0.0083	0.0079	0.0081	0.0075	
	$\hat{\theta}$	0.3999	0.7149	0.7637	0.6930	1.0069	
	Bias	4.9×10^{-7}	0.8727	0.9011	0.8338	1.0845	
	MSE	2.2×10^{-6}	1.0413	1.1381	0.9928	1.5686	
	200	$\hat{\lambda}$	0.0021	0.0026	0.0026	0.0025	0.0024
	Bias	8.8×10^{-5}	0.0014	0.0013	0.0013	0.0013	
	MSE	1.4×10^{-4}	0.0016	0.0017	0.0016	0.0016	
	$\hat{\beta}$	0.0499	0.0476	0.0479	0.0481	0.0480	
	Bias	3.6×10^{-5}	0.0058	0.0057	0.0057	0.0051	
	MSE	2.6×10^{-4}	0.0069	0.0068	0.0068	0.0064	
	$\hat{\theta}$	0.3999	0.6235	0.6042	0.6210	0.9270	
Bias	2.4×10^{-7}	0.7702	0.7480	0.7534	0.9689		
MSE	1.2×10^{-6}	0.9273	0.9533	0.9147	1.5688		

Table 4 Parameter estimation, MSE and Bias from five different methods. Set III

Set III: ($\lambda=3, \beta=2.2, \theta=1.6$)						
n	Par.	ML	O.LS	W.LS	CRM	PMS
15	$\hat{\lambda}$	3.9322	4.3582	4.2387	4.1921	3.8708
	Bias	1.4051	2.5085	2.5473	2.4046	2.2388
	MSE	2.4254	3.5843	3.6015	3.5199	2.8925
	$\hat{\beta}$	3.8863	1.6668	1.9058	3.6254	1.2220
	Bias	2.6651	4.0830	3.7259	4.3677	2.8072
	MSE	3.6928	5.5261	5.4303	6.3735	3.5693
	$\hat{\theta}$	1.0210	2.0309	2.3391	2.0002	6.5198
	Bias	0.9275	1.7119	2.0173	1.7641	6.1664
	MSE	1.2258	2.9075	3.4730	3.3698	19.1609
30	$\hat{\lambda}$	3.6988	3.9450	3.7604	3.8534	3.5659
	Bias	1.2542	2.1257	2.0994	2.0490	1.8882
	MSE	1.5303	2.7173	2.7858	2.6625	2.1896

continued on next page

Table 4 Parameter estimation (Set III) – (continue)

		Set III: ($\lambda=3, \beta=2.2, \theta=1.6$)				
n	Par.	ML	O.LS	W.LS	CRM	PMS
50	$\hat{\beta}$	2.9954	1.5525	1.7536	2.5062	1.3431
	Bias	1.6886	2.6041	2.2638	2.6468	1.8356
	MSE	2.3045	3.3284	2.9826	3.5060	2.3013
	$\hat{\theta}$	1.1767	1.9432	2.1947	1.8704	3.8317
	Bias	0.7988	1.4772	1.6692	1.4056	3.3037
	MSE	0.9342	1.8631	2.3645	1.9858	8.6487
	$\hat{\lambda}$	3.5703	3.8595	3.7238	3.8216	3.3875
	Bias	1.1732	1.9669	1.9595	1.8917	1.6775
	MSE	1.4371	2.2982	2.2877	2.2546	1.9340
	$\hat{\beta}$	2.5994	1.6458	1.7512	2.1884	1.4693
	Bias	1.1939	1.9547	1.6447	1.9264	1.3618
	MSE	1.5871	2.4575	2.0605	2.4781	1.6699
100	$\hat{\theta}$	1.3045	1.7989	1.9632	1.7068	2.6875
	Bias	0.7949	1.3445	1.4894	1.2569	2.0296
	MSE	0.9902	1.5895	1.7479	1.5092	4.1732
	$\hat{\lambda}$	3.3079	3.8279	3.6028	3.8005	3.1955
	Bias	0.8619	1.8947	1.8400	1.8255	1.5130
	MSE	1.0235	2.1977	2.0587	2.0985	1.7332
	$\hat{\beta}$	2.3920	1.6920	1.7822	1.9717	1.6057
	Bias	0.7654	1.4484	1.2057	1.4104	0.9854
	MSE	1.0026	1.7729	1.4632	1.7393	1.1895
	$\hat{\theta}$	1.5261	1.7063	1.9394	1.6455	2.4444
	Bias	0.6711	1.2482	1.3998	1.1866	1.6472
	MSE	0.8241	1.3583	1.5782	1.3023	2.4673
200	$\hat{\lambda}$	3.2362	3.7742	3.6352	3.7493	3.1166
	Bias	0.8826	1.8009	1.7777	1.7573	1.4572
	MSE	1.0504	2.0324	1.9913	1.9899	1.6681
	$\hat{\beta}$	2.2067	1.7354	1.7847	1.8816	1.6800
	Bias	0.5499	1.0961	0.9211	1.0674	0.7436
	MSE	0.7088	1.3263	1.1181	1.3020	0.9055
	$\hat{\theta}$	1.6225	1.6818	1.8469	1.6535	2.4265
	Bias	0.7027	1.1868	1.2991	1.1479	1.5679
	MSE	0.8863	1.2611	1.4379	1.2183	2.1053

6 Real-life Data Applications

Five real-world dataset from different disciplines in medicine, engineering, and physics are employed to illustrate the flexibility and efficiency of the NEG as a lifetime model. The first dataset comprised the survival times of 121 breast cancer patients who were treated at a major hospital between 1929 and 1938 [18]. The second dataset included 36 patients and recorded the periods of remission in months for individuals with bladder cancer [13]. This third dataset consisted of the lifetime (in years) of 40 blood cancer (leukemia) patients from one of the Ministry of Health hospitals in Saudi Arabia [5]. The fourth dataset was provided by [16]. The data involved 59 items which reflects the electrical relay failure time (in hours) for a test conductor. The fifth dataset describes the fracture toughness for different materials [21].

The NEG model is compared with the exponential (E), Rayleigh (R), Gompertz (G), generalized Gompertz distribution (GG) [10], Exponentiated Gompertz Exponential (EGoG) [1].

The performance of the NEG is assessed using several goodness of fit indexes (GoF). The negative likelihood value ($-\ell$), Akaike information criterion (AIC), corrected AIC (CAIC), Bayesian information criterion (BIC), Kramér-von Mises (W^*) test statistic, Anderson-Darling (AD^*) test statistic, Kolmogorov-Smirnov (KS) test statistic and associated p -value are employed in order to compare the proposed test with the competitive models. A model

Table 5 The Performance of NEG for Dataset I (Breast cancer data).

Distributions	NEG	GG	EGoE	Gompertz	R	E
Estimates	$\hat{\lambda} = 0.008$ (0.004)	$\hat{\lambda} = 0.029$ (0.009)	$\hat{\lambda} = 1.430$ (0.401)	$\hat{\lambda} = 0.016$ (.002)	$\hat{\lambda} = 41.113$ (1.869)	$\hat{\lambda} = 0.515$ (0.086)
	$\hat{\beta} = 0.014$ (0.003)	$\hat{\beta} = 0.002$ (0.004)	$\hat{\beta} = 2.380$ (0.867)	$\hat{\beta} = 0.011$ (0.002)		
	$\hat{\theta} = 0.845$ (0.767)	$\hat{\theta} = 1.752$ (0.398)	$\hat{\theta} = 0.726$ (0.106)	$\hat{\gamma} = 0.006$ (0.001)		
$-\ell$	579.218	580.974	583.599	583.765	600.989	585.128
AIC	1164.44	1167.949	1175.197	1171.530	1203.98	1172.255
CAIC	1168.629	1172.142	1180.789	1174.325	1205.378	1173.653
BIC	1172.823	1176.336	1186.38	1177.121	1206.776	1175.051
W^*	0.100	0.099	0.329	0.332	1.539	0.459
AD^*	0.734	0.832	1.942	2.5082	11,381	2.693
KS	0.057	0.078	0.105	0.098	0.198	0.120
p -value	0.828	0.447	0.136	0.194	.0001	0.060

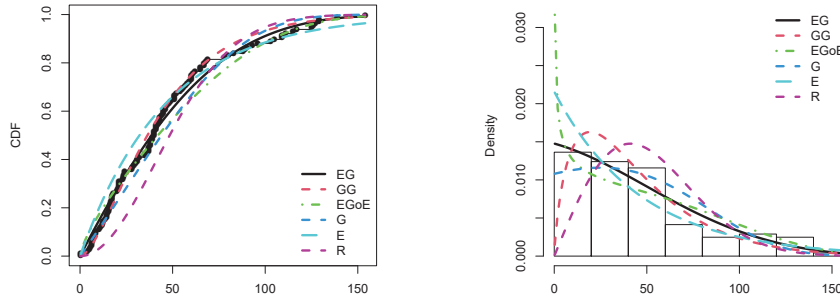


Figure 3 Empirical CDFs and PDFs for Dataset II (Breast cancer data).

Table 6 The Performance of NEG for Dataset II (Bladder cancer data).

Distributions	NEG	GG	EGoE	Gompertz	R	E
Estimates	$\hat{\lambda} = 0.015$ (0.014)	$\hat{\lambda} = 0.779$ (0.619)	$\hat{\lambda} = 5.891$ (2.145)	$\hat{\lambda} = 2.369$ (0.164)	$\hat{\lambda} = 1.536$ (0.128)	$\hat{\lambda} = 0.515$ (0.086)
	$\hat{\beta} = 1.051$ (0.217)	$\hat{\beta} = 0.108$ (0.394)	$\hat{\beta} = 12.172$ (12.995)	$\hat{\beta} = 0.004$ (0.002)		
	$\hat{\theta} = 6.161$ (6.569)	$\hat{\theta} = 2.445$ (1.390)	$\hat{\theta} = 1.425$ (0.742)			
			$\hat{\gamma} = 0.049$ (0.004)			
$-\ell$	47.522	53.619	48.645	66.089	51.399	59.857
AIC	101.044	113.238	105.291	136.179	104.797	121.714
CAIC	103.419	115.614	108.458	137.762	105.589	122.505
BIC	105.79	117.989	111.625	139.346	106.381	123.297
W*	0.084	0.299	0.159	1.089	0.172	0.656
AD*	0.547	1.662	0.886	11.758	1.121	3.408
KS	0.102	0.213	0.153	0.333	0.162	0.230
<i>p</i> -value	0.845	0.212	0.366	0.001	0.301	0.044

is considered to be the best in representing data when the values associated with these statistics are less than those of competing models.

The Tables from 5 to 9 report the MLEs and the standard error corresponding with the GoF measurements of NEG and competitive distributions. The tables show that, the performance of the classical distributions fluctuated depending on the type of data. The exponential, Rayleigh, Gompertz distributions did not fit all the data, as they showed a significant *p*-value < 0.05. As for modern distributions, such as GG and EGoG, the *p*-value > 0.05, which means that these distributions are able to represent the data. Although these models fit the five datasets (*p*-value > 0.05), the NEG has the highest *p*-value among all the fitted models. Moreover, the Tables demonstrate that the

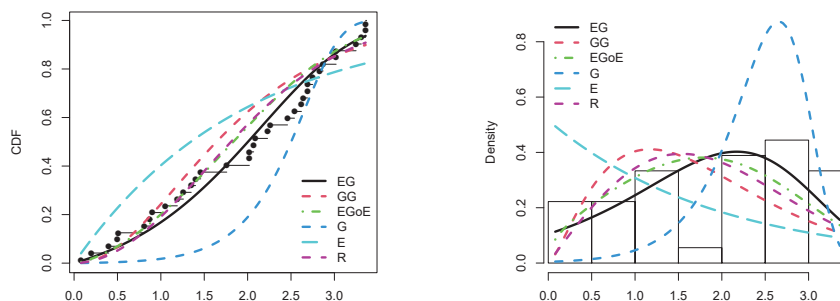


Figure 4 Empirical CDFs and PDFs for Dataset II (Bladder cancer data).

Table 7 The Performance of NEG for Dataset III (Blood cancer data).

Distributions	NEG	GG	EGoE	gompertz	R	E
Estimates	$\hat{\lambda} = 0.039$ (0.075)	$\hat{\lambda} = 0.033$ (0.031)	$\hat{\lambda} = 0.062$ (0.044)	$\hat{\lambda} = 0.503$ (0.113)	$\hat{\lambda} = 2.415$ (0.191)	$\hat{\lambda} = 0.318$ (0.050)
	$\hat{\beta} = 0.812$ (0.367)	$\hat{\beta} = 0.859$ (0.223)	$\hat{\beta} = 1.821$ (0.710)	$\hat{\beta} = 0.096$ (0.031)		
	$\hat{\theta} = 0.123$ (1.594)	$\hat{\theta} = 0.965$ (0.366)	$\hat{\theta} = 0.949$ (0.299)			
			$\hat{\gamma} = 0.482$ (0.140)			
$-\ell$	65.778	65.810	65.881	68.281	70.806	85.778
AIC	137.556	137.620	139.762	140.563	143.612	173.556
CAIC	140.089	140.154	143.139	142.252	144.456	174.401
BIC	142.623	142.687	146.517	143.940	145.301	175.245
W*	0.027	0.034	0.047	0.118	0.250	1.084
AD*	0.211	0.282	0.358	0.719	1.297	5.479
KS	0.066	0.079	0.083	0.112	0.158	0.300
<i>p-value</i>	0.995	0.944	0.994	0.701	0.277	0.001

NEG achieved the minimum values for all GoF indices. This indicates that the NEG distribution suits all five datasets better than the other competitive distributions.

The estimated PDF and CDF of NEG and the other models are presented in Figures from 3 to 7. The histogram represents the empirical density for the data and the dot black line represents the empirical CDF for the data. The Figures demonstrate that NEG closely aligns with the actual distribution of the datasets under examination. Therefore, when contrasted with alternative distributions, the NEG model emerges as the most suitable choice for the analyzed data.

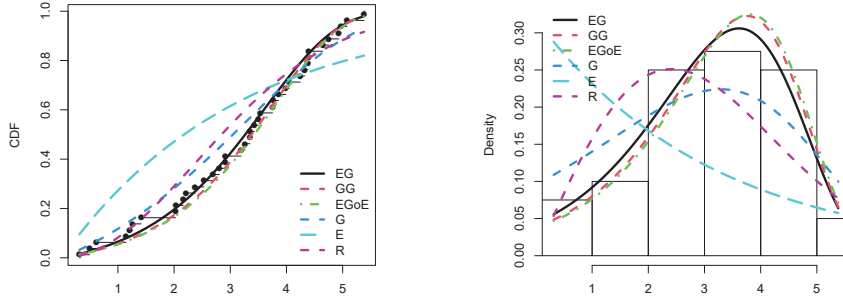


Figure 5 Empirical CDFs and PDFs for Dataset III (Blood cancer data).

Table 8 The pPERFORMANCE of NEG for Dataset IV (electrical relay failure time). (in hours) for a test conductor.

Distributions	NEG	GG	EGoE	Gompertz	R	E
Estimates	$\hat{\lambda} = 0.003$ (0.001)	$\hat{\lambda} = 0.004$ (0.002)	$\hat{\lambda} = 0.019$ (0.018)	$\hat{\lambda} = 0.674$ (0.034)	$\hat{\lambda} = 5.064$ (0.329)	$\hat{\lambda} = 0.143$ (0.019)
	$\hat{\beta} = 0.636$ (0.038)	$\hat{\beta} = 0.615$ (0.049)	$\hat{\beta} = 0.001$ (0.003)	$\hat{\beta} = 0.003$ (0.001)		
	$\hat{\theta} = 1.015$ (0.615)	$\hat{\theta} = 0.890$ (0.165)	$\hat{\theta} = 15.847$ (5.936)			
			$\hat{\gamma} = 21.560$ (12.855)			
$-\ell$	115.614	117.938	117.006	118.151	137.412	173.641
AIC	237.228	241.877	242.013	240.303	276.825	349.281
CAIC	240.345	244.993	246.168	242.380	277.864	350.319
BIC	243.461	248.109	250.323	244.458	278.902	351.359
W*	0.165	0.235	0.264	0.332	1.708	3.431
AD*	1.107	1.461	1.672	1.714	1.708	3.431
KS	0.108	0.130	0.114	0.162	0.313	0.430
p-value	0.465	0.248	0.399	0.080	1.2×10^{-05}	1.9×10^{-10}

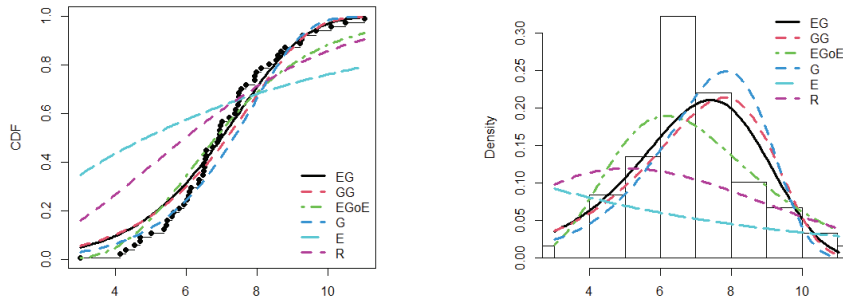


Figure 6 Empirical CDFs and PDFs for Dataset IV (Electrical failure time).

Table 9 The Performance of NEG for the Dataset V (Fracture toughness).

Distribution	NEG	GG	EGoE	Gompertz	R	E
Estimates	$\hat{\lambda} = 0.002$ (0.0003) $\hat{\beta} = 1.192$ (0.044) $\hat{\theta} = 0.843$ (0.416)	$\hat{\lambda} = 0.151$ (.096) $\hat{\beta} = 0.451$ (0.135) $\hat{\theta} = 5.030$ (2.649)	$\hat{\lambda} = 0.011$ (0.005) $\hat{\beta} = 0.960$ (0.356) $\hat{\theta} = 1.157$ (0.215) $\hat{\gamma} = 1.008$ (0.336)	$\hat{\lambda} = 1.213$ (0.040) $\hat{\beta} = 0.003$ (0.001)	$\hat{\lambda} = 3.141$ (0.144)	$\hat{\lambda} = 0.231$ (0.021)
$-\ell$	170.142	170.294	170.719	173.534	221.035	293.275
AIC	346.283	346.588	349.437	351.068	444.069	588.551
CAIC	350.452	350.757	354.996	353.847	445.459	589.941
BIC	354.621	354.926	360.554	356.626	446.848	591.330
W*	0.089	0.103	0.113	0.246	3.421	6.716
AD*	0.547	0.818	0.690	1.305	17.274	32.261
KS	0.065	0.076	0.068	0.105	0.299	0.396
<i>p-value</i>	0.701	0.501	0.632	0.143	1.2×10^{-9}	1.1×10^{-16}

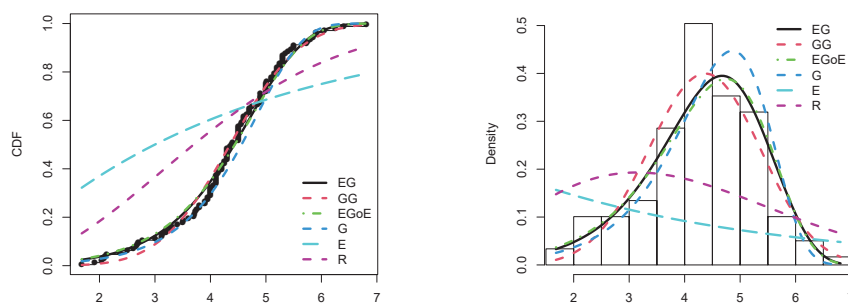


Figure 7 Empirical CDFs and PDFs for Dataset V (Fracture toughness).

7 Conclusions

In this paper, the NEG model is generated from Gompertz distribution based on NEX family. This distribution is designed to give greater adaptability and fit data from real life. The NEG’s density and hazard functions of NEG possess appealing shapes that can be used to fit various data patterns. Number of mathematical properties of the NEG model are described in detail including exact formulations for the density, moments, quantile order statistics and Rényi entropy. The ML, O.LS, W.LS, CRM and PMS estimations of the parameters are derived as well as assessed via simulation studies. ML is the most reliable estimate for estimating NEG parameters since it provides the smallest MSE compared to other estimates. The O.LS, W.LS, CRM, and MPS

estimates exhibit similar performance regarding MSE values. While, MPS is considered as less efficient due to its higher MSE values in certain parameter estimations. Five different applications in medical, engineering and physics are used to assess the proposed model. The foremost eminent perspective is that NEG is a more suitable model than several competing models since it provides the smallest value for several GoF criteria. This illustrates that NEG exceeds all other competitor models in regard to execution and flexibility when conducted on different datasets. Based to these results, the NEG model could be useful for modeling other types of data in various area.

In future work, we plan to extend the application of the NEG model to additional domains and data types to further assess its versatility and robustness. This will include exploring its performance in other fields beyond medical, engineering, and physics, such as finance and social sciences, to evaluate its adaptability to diverse datasets. We will also investigate the potential of integrating the NEG model with other advanced estimation techniques, including Bayesian approaches, to improve parameter estimation and model fitting. Additionally, we aim to develop methods for handling complex censoring mechanisms and assess the model's performance with various types of lifetime data. Enhancing the model's computational efficiency and conducting real-world case studies will also be a focus, with the goal of refining the model's applicability and providing more comprehensive insights into its practical utility across different applications.

Appendix

```
#####
#PDF
dEG=function(x, a, b, th)
{d=exp(-(a/b)*(exp(b*x)-1))
ff=a*exp(b*x)*d*((1+(th*d))/exp(th*(1-d)))
return(ff)
}
#####
#CDF
pEG=function(q, a, b, th)
{
x=q
d=exp(-(a/b)*(exp(b*x)-1))
```

```

F=1-(d*exp(-th*(1-d)))
return (F)
}
#####
# quantile NEG
qEG=function (n, a, b, th)
{
u=runif (n, 0, 1)
z=th*exp ( th )*(1-u)
d=log ( lambertWp(z) / th )
F=log (1 -((b/a)*d)) / b
return (F)
}
#####
# Likelihood
lh=function (par, x){
  a=par [1] ; b=par [2] ; th=par [3]
  L=sum ( log (dEG(x, a, b, th)) )
  return (-L)}
#####
## Simulation
n=15; #n=(30,50,100,200)

NS=1000
opt_th=c ()
opt_a=c ()
opt_b=c ()

start=c (a, b, th)
#####
sample=qEG(n*1000, a=a, b=b, th=th)
x= matrix ( sample, nrow = NS, ncol=n)

for ( i in 1:NS) #i=1
{
opt= optim ( par = start, fn = lh, x = x[i,] );
opt_a[i]=opt$par [1]
opt_b[i]=opt$par [2]
opt_th[i]=opt$par [3]
}

```

```

}
#####
MLH_a.h = mean(opt_a)
MLH_b.h = mean(opt_b)
MLH_th.h = mean(opt_th)
MLH.EG.h=c(MLH_a.h,MLH_b.h,MLH_th.h)
#### #### ####
Bias_a.h = mean(abs(opt_a - a))
Bias_b.h = mean(abs(opt_b - b))
Bias_th.h = mean(abs(opt_th - th))
Bias.EG.h=c(Bias_a.h,Bias_b.h,Bias_th.h)
#### #### ####
MSE_a.h = sqrt(mean((opt_a - a)^2))
MSE_b.h = sqrt(mean((opt_b - b)^2))
MSE_th.h = sqrt(mean((opt_th - th)^2))
MSE.EG.h=c(MSE_a.h,MSE_b.h,MSE_th.h)
#####
}

```

References

- [1] A. Ademola, J. Adeyeye, M. Khaleel, and A. Olubisi. Exponentiated gompertz exponential (egoe) distribution: Derivation, properties and applications. *Istatistik Journal of the Turkish Statistical Association*, 13(1):12–28, 2021.
- [2] M. Alizadeh, G. M. Cordeiro, L. G. B. Pinho, and I. Ghosh. The gompertz-g family of distributions. *Journal of statistical theory and practice*, 11:179–207, 2017.
- [3] A. Alzaatreh, C. Lee, and F. Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63–79, 2013.
- [4] A. Alzaatreh, C. Lee, and F. Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63–79, 2013.
- [5] M. Atallah, M. Mahmoud, and B. Al-Zahrani. A new test for exponentiality versus nbunmgf life distributions based on laplace transform. *Quality and Reliability Engineering International*, 30(8):1353–1359, 2014.
- [6] L. A. Baharith. New generalized weibull inverse gompertz distribution: Properties and applications. *Symmetry*, 16(2):197, 2024.

- [7] R. Cheng and N. Amin. Maximum product-of-spacings estimation with applications to the lognormal distribution. *Math report*, 791, 1979.
- [8] G. M. Cordeiro, E. M. Ortega, and D. C. da Cunha. The exponentiated generalized class of distributions. *Journal of Data Science*, 11(1):1–27, 2013.
- [9] T. A. De Andrade, S. Chakraborty, L. Handique, and F. Gomes-Silva. The exponentiated generalized extended gompertz distribution. *Journal of Data Science*, 17(2):299–330, 2019.
- [10] A. El-Gohary, A. Alshamrani, and A. N. Al-Otaibi. The generalized gompertz distribution. *Applied mathematical modelling*, 37(1-2):13–24, 2013.
- [11] Z. K. Ezmareh and G. Yari. Kumaraswamy-g generalized gompertz distribution with application to lifetime data. *International Journal of Industrial Engineering*, 33(4):1–22, 2022.
- [12] B. Gompertz. Xxiv. on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. frs &c. *Philosophical transactions of the Royal Society of London*, (115):513–583, 1825.
- [13] R. Hibatullah, Y. Widyaningsih, and S. Abdullah. Marshall-olkin extended power lindley distribution with application. *J. Ris. and Ap. Mat*, 2(2):84–92, 2018.
- [14] X. Huo, S. K. Khosa, Z. Ahmad, Z. Almaspoor, M. Ilyas, and M. Aamir. A new lifetime exponential-x family of distributions with applications to reliability data. *Mathematical Problems in Engineering*, 2020:1–16, 2020.
- [15] J. Y. Kajuru, H. D. Garba, A. M. Suleiman, and A. F. Ibrahim. The generalized gompertz-g family of distributions: Statistical properties and applications. *UMYU Scientifica*, 3(1):120–128, Mar. 2024.
- [16] A. A. Khalaf et al. [0, 1] truncated exponentiated exponential gompertz distribution: Properties and applications. In *AIP Conference Proceedings*, volume 2394. AIP Publishing, 2022.
- [17] G. A. Korn and T. M. Korn. *Mathematical handbook for scientists and engineers: definitions, theorems, and formulas for reference and review*. Courier Corporation, 2000.
- [18] E. T. Lee and J. Wang. *Statistical methods for survival data analysis*, volume 476. John Wiley & Sons, 2003.
- [19] P. Macdonald. Estimation procedure for mixture distributions. *Journal of the royal society series B- statistical methodology*, 33(2):326–, 1971.

- [20] J. Moors. A quantile alternative for kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 37(1):25–32, 1988.
- [21] S. Nadarajah and S. Kotz. On the alternative to the weibull function. *Engineering fracture mechanics*, 74(3):451–456, 2007.
- [22] M. Nassar, A. Alzaatreh, M. Mead, and O. Abo-Kasem. Alpha power weibull distribution: Properties and applications. *Communications in Statistics-Theory and Methods*, 46(20):10236–10252, 2017.
- [23] P. E. Oguntunde, M. A. Khaleel, M. T. Ahmed, A. O. Adejumo, O. A. Odetunmbi, et al. A new generalization of the lomax distribution with increasing, decreasing, and constant failure rate. *Modelling and Simulation in Engineering*, 2017, 2017.
- [24] M. Shama, S. Dey, E. Altun, and A. Z. Afify. The gamma–gompertz distribution: Theory and applications. *Mathematics and Computers in Simulation*, 193:689–712, 2022.
- [25] J. J. Swain, S. Venkatraman, and J. R. Wilson. Least-squares estimation of distribution functions in johnson’s translation system. *Journal of Statistical Computation and Simulation*, 29(4):271–297, 1988.

Biography

Ibtesam Ali Alsaggaf received his B.Sc. and M.Sc. degrees from the Statistics department at King Abdulaziz University and Ph.D. degrees from the School of Mathematics at Universiti Sains Malaysia, Penang, in 2013. She is currently an assistant professor at the Statistics Department at King Abdulaziz University. She has published articles in International journals. Her fields of interest are Distribution theory and modeling.

