Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

The effect of exponentiating generalized models

Hadeel S. Klakattawi *, Aisha A. Khormi

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

ARTICLE INFO

Article history:
Received 1 April 2022
Received in revised form
12 July 2022
Accepted 28 July 2022

Keywords:
T-X family
Exponentiated
Generalized Inverted Kumaraswamy distribution
Gompertz distribution
Maximum likelihood
Monte Carlo simulation

ABSTRACT

A common practice in statistical distribution theory involves exponentiating existing distribution functions to include some extra parameters that increase the flexibility of the distribution. This paper examines the effect of exponentiating some generalized models by adding three extra parameters to their probability distribution. Particularly, a new generalized distribution that is a member of the inverted Kumaraswamy family of distributions is considered. Afterward, three additional parameters are applied to enhance this generalized distribution, which results in a novel distribution referred to as the new generalized exponentiated generalized inverted Kumaraswamy Gompertz distribution (NGEGIKGD). Some of the statistical and mathematical characteristics of this distribution were derived. Additionally, parametric estimation of the new distribution parameters was considered using the maximum likelihood method. Several Monte Carlo simulation studies were conducted in order to explore the usefulness of the estimation method. The proposed distribution is then compared with its corresponding sub-models in order to assess the effects of the exponentiation. Further evaluation of the distribution is accomplished by comparing it to some relative distributions. Specifically, three real-world datasets were analyzed to demonstrate the potentiality of the suggested new modeling approach in enhancing the goodness of fit of the generalized models. Results indicate that exponentiating a generalized model significantly improves its fit compared to the non-exponentiating distributions.

1. Introduction

Statistical distributions play an important role in describing, modeling, analyzing, simulating, and performing inferences in relation to many world phenomena. Specifically, statistical distributions play a key role in understanding data and their characteristics through the implementation of calculations to make critical decisions about a phenomenon based on the results. Thus, the best distribution needs to be found to fit datasets.

Numerous and powerful statistical distributions have been applied to explain and fit a variety of phenomena in different areas. However, some of these classical distributions cannot describe or fit the complex behavior of real data, such as symmetry or strong skewness. Thus, there is a continuous need to find more flexible and adaptable distributions in order to improve data fitting. Recently, many statisticians have been interested in extending the classical distributions and adding one or more extra shape parameters to an existing distribution in order to propose some new family of distributions that will enhance their goodness of fit; for a review, see Gupta et al. (1998), Cordeiro et al. (2013), and Rezaei et al. (2017).

According to the T-X method, a new generator family of distributions has been proposed (Jamal et al., 2019) using the inverted Kumaraswamy distribution as a generator; for more information about this method, see Alzaatreh et al. (2013). The cumulative distribution function (cdf) and the probability density function (pdf) of the new generalized inverted Kumaraswamy family of distributions are obtained as:

$$F(x) = [1-(1-G^{\gamma}(x,\zeta))^{\alpha}]^{\beta}, \alpha,\beta,\gamma > 0, \quad (1)$$

$$f(x) = \alpha\beta\gamma g(x,\zeta)G^{\gamma-1}(x,\zeta)(1-G^{\gamma}(x,\zeta))^{\alpha-1} \times [1-(1-G^{\gamma}(x,\zeta))^{\alpha}]^{\beta-1}, \alpha,\beta,\gamma > 0, \quad (2)$$

where, ζ is the parametric space of the baseline distribution and $G(x,\zeta)$ and $g(x,\zeta)$ are the cdf and pdf of any statistical distribution. The fundamental motivations for obtaining this class of distributions

* Corresponding Author.
Email Address: hklakattawi@kau.edu.sa (H. S. Klakattawi)
https://doi.org/10.21833/ijaas.2022.11.006
Corresponding author's ORCID profile:
https://orcid.org/0000-0001-6617-2081

could be summarised in some points including improving the flexibility of the kurtosis and skewness in comparison with the baseline distribution, constructing heavy-tailed models without exhibiting longer tails for fitting real data, providing various shapes for the pdf of the generated distributions, and enhancing the goodness-of-fit for the given distribution when compared to other competitive distributions.

The Gompertz distribution proposed in Gompertz (1825) is a classical distribution that represents the survival function based on the laws of mortality. The cdf and pdf of the Gompertz distribution are defined, respectively, as:

$$G(x) = 1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}, x \geq 0, \lambda, k > 0, \quad (3)$$

$$g(x) = \lambda e^{kx} e^{-\frac{\lambda}{k}(e^{kx}-1)}, x \geq 0, \lambda, k > 0. \quad (4)$$

Although this Gompertz distribution is one of the most common distributions used to describe lifespans, it must be understood that its hazard rate function is monotonically increasing. In practice, however, it is necessary to take into account cases with non-monotonic increasing functions. Therefore, many authors (Joshi and Kumar, 2020; Eghwerido et al., 2021; Shama et al., 2022) have recently incorporated extra parameters into the Gompertz distribution to overcome this disadvantage.

It is argued in this paper that the benefits of the generalized inverted Kumaraswamy family of distributions can be applied to the Gompertz distribution in order to construct a more flexible distribution for accurately fitting real-world data. Specifically, by replacing $G(x,\zeta)$ in Eq. 1 by the cdf in Eq. 3 and $g(x,\zeta)$ in Eq. 2 by the pdf in Eq. 4 we obtain a new distribution namely, the generalized inverted Kumarswamy Gompertz distribution with a cdf and pdf as follows:

$$G(x) = (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}, \quad x \geq 0, \quad (5)$$

$$\begin{aligned} g(x) &= \alpha\beta\gamma\lambda e^{kx} e^{-\frac{\lambda}{k}(e^{kx}-1)}[1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma-1}\{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha-1} \\ &\times (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta-1}, \quad x \geq 0, \end{aligned} \quad (6)$$

where, $k > 0$ is a scale parameter and λ, α, β, and $\gamma > 0$ are shape parameters.

Adding extra parameters to classical or existing models can be accomplished using the exponentiation method. Recently, taking inspiration from series-parallel-series systems, the authors of Rezaei et al. (2017) proposed a new generalized exponentiated class of distributions with three extra shape parameters. The cdf and pdf of this new class are defined, respectively, as:

$$F(x) = 1 - (1 - \{1 - [1 - G(x,\zeta)]^{a}\}^{b})^{\theta}, \quad (7)$$

$$f(x) = ab\theta g(x,\zeta)[1 - G(x,\zeta)]^{a-1} \times \{1 - [1 - G(x,\zeta)]^{a}\}^{b-1}(1 - \{1 - [1 - G(x,\zeta)]^{a}\}^{b})^{\theta-1}, \quad (8)$$

where, ζ is the parametric space of the baseline distribution, $G(x,\zeta)$ and $g(x,\zeta)$ are the cdf and pdf of any statistical distribution and a, b and $\theta > 0$ are positive real numbers. Different generalizations have been introduced based on the new generalized exponentiated class of distributions; for example, Nasiru et al. (2019) proposed the exponentiated generalized exponential Dagum distribution and De Andrade et al. (2019) proposed the exponentiated generalized extended Gompertz distribution.

Reasons for the adoption of this generator in Eq. 7 include its simplicity. That is, besides the ability of the additional parameters to control both the weights of the data, this cdf is always tractable, as it has no complicated function, and obtaining its inverse is very straightforward. The distributions generated from this family can also be interpreted physically in terms of the series-parallel-series systems and the hierarchical structure with three levels. For the purpose of exploring the usefulness of exponentiating a generalized model by adding the three parameters in Eq. 7, this paper compounds the generalized inverted Kumarswamy Gompertz distribution to the new generalized exponentiated class. Particularly, $G(x,\zeta)$ in Eq.7 is replaced by the cdf in Eq. 5, and $g(x,\zeta)$ in Eq. 8 is replaced by the pdf in Eq. 6, aiming to provide a more flexible, practical and accurate distribution in describing a variety of real life applications, e.g., engineering, reliability, and real-life data. It becomes evident from the application of this novel distribution to a number of different datasets that the exponentiation method plays a very important role in the concept of model flexibility. These reasons highlight the importance of examining further the proposed distribution, namely the new generalized exponentiated generalized inverted Kumaraswamy Gompertz distribution (NGEGIKGD).

This article consists of the following: In Section 2, we propose the new distribution, the NGEGIKGD, and some graphical representations of its density and hazard rate function (hrf) are provided. In Section 3, the expansion of the pdf for the NGEGIKGD is derived. In Section 4, we study some of the statistical properties of the proposed distribution. The distribution parameters are determined using the maximum likelihood (ML) method in Section 5. Section 6, reports the simulation results. In Section 7, three real datasets are presented. Conclusions are provided in Section 8.

2. The new generalized exponentiated generalized inverted Kumaraswamy Gompertz distribution

A random variable X is said to have an NGEGIKGD with eight parameters $k > 0$ as scale parameters and $a, b, \theta, \alpha, \beta, \gamma, \lambda > 0$ as shape parameters if its cdf and pdf are given by the following form:

$$F(x) = 1 - (1 - \{1 - [1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b})^{\theta}, \quad x \geq 0, \tag{9}$$

and,

$$\begin{aligned} f(x) \; &= ab\theta\alpha\beta\gamma\lambda e^{kx} e^{-\frac{\lambda}{k}(e^{kx}-1)}[1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma-1}\{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha-1} \\ &\times (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta-1}[1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a-1} \\ &\times \{1 - [1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b-1} \\ &\times (1 - \{1 - [1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b})^{\theta-1}, \quad x \geq 0. \end{aligned} \tag{10}$$

The survival function (SF) is frequently used to describe the distribution of survival time. Then, the SF of the NGEGIKGD is given by:

$$S(x) = (1 - \{1 - [1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b})^{\theta}, \quad x \geq 0. \tag{11}$$

The hrf of the NGEGIKGD, which is often used in lifespan modeling as it indicates the likelihood of failure, is defined as:

$$\begin{aligned} h(x) \; &= \frac{ab\theta\alpha\beta\gamma\lambda e^{kx} e^{-\frac{\lambda}{k}(e^{kx}-1)}[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma-1}\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha-1}}{1-\{1-[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b}} \\ &\times (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta-1}[1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a-1} \\ &\times \{1 - [1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b-1}. \end{aligned} \tag{12}$$

The pdf and hrf plots for the NGEGIKGD are given in Fig. 1, and Fig. 2, respectively, at some certain values of the distribution's parameters.

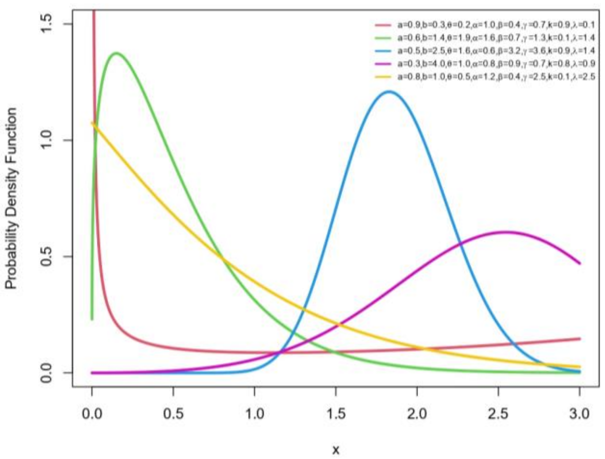

Fig. 1: The pdf's plots of the NGEGIKGD

In Fig. 1, it can be seen that the NGEGIKGD's pdf is decreasing, symmetrical, positively skewed, and negatively skewed. In addition, in Fig. 2, the hrf of the NGEGIKGD takes constant, increasing, decreasing and bathtub shapes which makes it an ideal choice for fitting different hazard behaviors that are more likely to appear in real-world situations such as reliability analysis, human mortality, and biological applications.

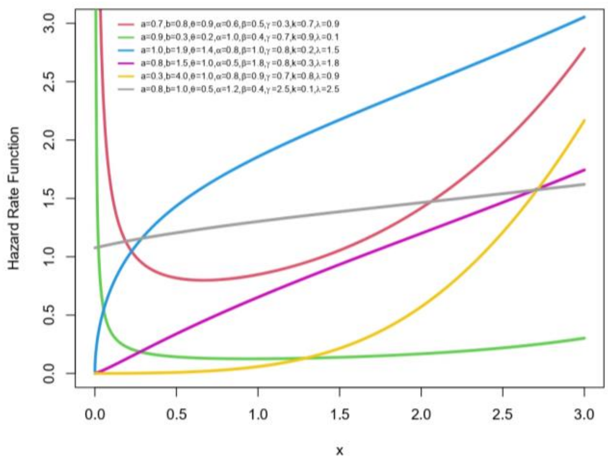

Fig. 2: The hrf's plots of the NGEGIKGD

3. Expansion of the NGEGIKGD's pdf

For a non-negative power, the binomial expansion is defined as:

$$(1-z)^{\eta-1} = \sum_{i=0}^{\infty} (-1)^i \binom{\eta-1}{i} z^i, \quad |z| < 1, \eta > 0. \tag{13}$$

Applying the binomial expansion in Eq. 13 six times to Eq. 10, we obtain:

$$f(x) = ab\theta\alpha\beta\gamma\lambda \sum_{v_1,v_2,v_3,v_4,v_5,v_6=0}^{\infty} (-1)^{v_1+v_2+v_3+v_4+v_5+v_6} \binom{\theta-1}{v_1}\binom{b(v_1+1)-1}{v_2} \times \binom{a(v_2+1)-1}{v_3}\binom{\beta(v_3+1)-1}{v_4}\binom{\alpha(v_4+1)-1}{v_5}\binom{\gamma(v_5+1)-1}{v_6} \times e^{kx} e^{-\frac{\lambda(v_6+1)}{k}e^{kx}} e^{\frac{\lambda}{k}(v_6+1)}.$$

The power series for the exponential function expansion is defined as:

$$e^{-x} = \sum_{\eta=0}^{\infty} \frac{(-1)^{\eta}}{\Gamma(\eta+1)} x^{\eta}. \quad (14)$$

The pdf of the NGEGIKGD is then obtained as:

$$f(x) = ab\theta\alpha\beta\gamma\lambda \sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7=0}^{\infty} V_v \, e^{\frac{\lambda}{k}(v_6+1)} e^{[k(v_7+1)]x} \quad ,x \geq 0, \quad (15)$$

where,

$$V_v = \frac{(-1)^{v_1+v_2+v_3+v_4+v_5+v_6+v_7}}{\Gamma(v_7+1)[\frac{\lambda(v_6+1)}{k}]^{-v_7}} \binom{\theta-1}{v_1}\binom{b(v_1+1)-1}{v_2}\binom{a(v_2+1)-1}{v_3} \times \binom{\beta(v_3+1)-1}{v_4}\binom{\alpha(v_4+1)-1}{v_5}\binom{\gamma(v_5+1)-1}{v_6}. \quad (16)$$

4. Properties of the NGEGIKGD

This section discusses some of the NGEGIKGD's properties.

4.1. The quantile and median

The quantile of the NGEGIKGD is:

$$Q(u) = \frac{1}{k} log\left\{1 - \frac{k}{\lambda} log(z(u))\right\}, \quad (17)$$

where,
$z(u) = 1 - [1 - \{1 - (1 - [1 - \{1 - (1 - u)^{1/\theta}\}^{1/b}]^{1/a})^{1/\beta}\}^{1/\alpha}]^{1/\gamma}$ and u is any value in (0,1).

The median of the NGEGIKGD (Med) is derived by setting $u = 0.5$ in Eq. 17 as:

$$Med = \frac{1}{k} log\left\{1 - \frac{k}{\lambda} log(z(0.5))\right\}, \quad (18)$$

where,
$z(0.5) = 1 - [1 - \{1 - (1 - [1 - \{1 - (0.5)^{1/\theta}\}^{1/b}]^{1/a})^{1/\beta}\}^{1/\alpha}]^{1/\gamma}$.

Therefore, the first quantile (Q(0.25)) and third quantile (Q(0.75)) of the NGEGIKGD are obtained by substituting $u = 0.25$ or $u = 0.75$, respectively.

The interquartile range (IQR) of the NGEGIKGD can be derived as:

$$IQR = Q(0.75) - Q(0.25)$$
$$IQR = \frac{1}{k} log\left[\frac{\{1-\frac{k}{\lambda}log(z(0.75))\}}{\{1-\frac{k}{\lambda}log(z(0.25))\}}\right]. \quad (19)$$

4.2. The Galton skewness and Moors kurtosis

The Galton skewness (GS) (Galton, 1883) measures the symmetry of distribution and is defined as:

$$GS = \frac{Q(\frac{6}{8}) - 2Q(\frac{4}{8}) + Q(\frac{2}{8})}{Q(\frac{6}{8}) - Q(\frac{2}{8})}. \quad (20)$$

The Moors kurtosis (MK) (Moors, 1988) is based on octiles and is defined as:

$$MK = \frac{(Q(\frac{7}{8}) - Q(\frac{5}{8})) + (Q(\frac{3}{8}) - Q(\frac{1}{8}))}{(Q(\frac{6}{8}) - Q(\frac{2}{8}))}. \quad (21)$$

The GS and MK of the NGEGIKGD are given in Fig. 3. From Fig. 3 the NGEGIKGD can be left skewed and for fixed α, the MK is an increasing function of θ.

4.3. The mode

The mode of the NGEGIKGD can be obtained by taking the derivative of the pdf in Eq. 10 with respect to x and equating to zero, $\frac{d}{dx} f(x) = 0$. That is,

$$[ab\theta\alpha\beta\gamma\lambda e^{kx} e^{-\frac{\lambda}{k}(e^{kx}-1)}[1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma-1}\{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha-1} \times (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta-1}[1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a-1} \times \{1 - [1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b-1} \times (1 - \{1 - [1 - (1 - \{1 - [1 - e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b})^{\theta-1}] = 0. \quad (22)$$

Then, the mode can be calculated numerically by solving the nonlinear Eq. 22.

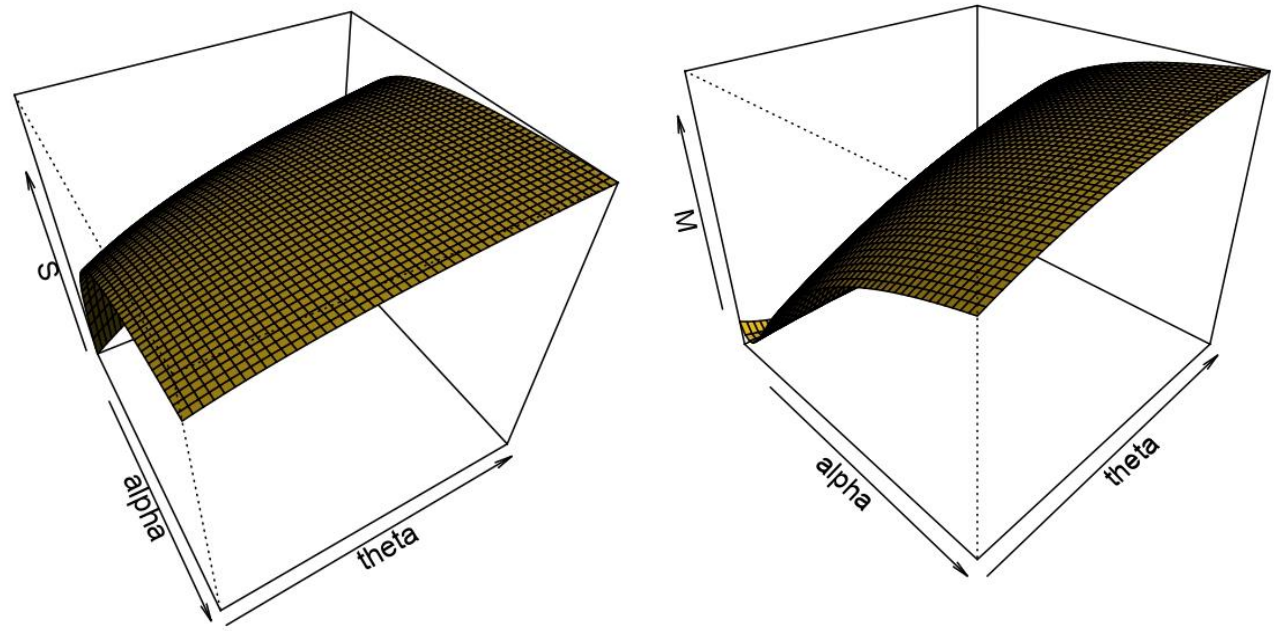

Fig. 3: GS (Left) and MK (Right) for the NGEGIKGD

4.4. The r^{th} moment

The r^{th} moment of the NGEGIKGD can be derived based on El-Gohary et al. (2013), and Khan et al. (2017) as:

$$\begin{aligned}\mu_r = E(x^r) &= \int_0^\infty x^r f(x)\,dx \\ &= ab\theta\alpha\beta\gamma\lambda \sum_{\upsilon_1,\upsilon_2,\upsilon_3,\upsilon_4,\upsilon_5,\upsilon_6,\upsilon_7=0}^{\infty} V_\upsilon \left[\frac{-1}{k(\upsilon_7+1)}\right]^{r+1} e^{\frac{\lambda}{k}(\upsilon_6+1)} \Gamma(r+1),\end{aligned} \tag{23}$$

where, V_υ is defined in Eq. 16. Then, the mean and variance of the NGEGIKGD are respectively, given as:

$$\mu_1 = ab\theta\alpha\beta\gamma\lambda \sum_{\upsilon_1,\upsilon_2,\upsilon_3,\upsilon_4,\upsilon_5,\upsilon_6,\upsilon_7=0}^{\infty} V_\upsilon \left[\frac{-1}{k(\upsilon_7+1)}\right]^2 e^{\frac{\lambda}{k}(\upsilon_6+1)}, \tag{24}$$

$$\sigma^2 = \mu_2 - \mu_1^2 = ab\theta\alpha\beta\gamma\lambda \sum_{\upsilon_1,\upsilon_2,\upsilon_3,\upsilon_4,\upsilon_5,\upsilon_6,\upsilon_7=0}^{\infty} 2V_\upsilon \left[\frac{-1}{k(\upsilon_7+1)}\right]^3 e^{\frac{\lambda}{k}(\upsilon_6+1)} - \mu_1^2. \tag{25}$$

The observed mean, variance, median and IQR of the NGEGIKGD for various values of the parameters are listed in Table 1.

Table 1: The observed mean, variance, median and IQR of the NGEGIKGD for various values of parameters

Case	Parameter								Mean	Variance	Median	IQR
	a	b	θ	α	β	γ	k	λ				
I	0.5	0.4	0.6	0.7	0.9	0.4	0.6	0.9	1.0772	1.4703	0.5170	1.9088
II	0.3	0.8	0.8	1.7	1.9	0.9	1.7	0.7	0.9517	0.2168	0.9357	0.7079
III	0.6	2.0	1.8	4.3	4.8	2.1	4.2	1.7	0.3340	0.0023	0.3314	0.0662
IV	2.4	2.5	2.9	4.5	5.0	2.7	4.6	4.9	0.1416	0.0002	0.1414	0.0183
V	2.5	2.7	3.0	4.7	5.1	3.0	4.8	6.0	0.1268	0.0001	0.1267	0.0149

In Table 1, we generate samples from the NGEGIKGD and calculate the mean, variance, median and IQR by using the built-in function in R program. Thus, Table 1, shows that when the values of the parameters increase the values of the mean, variance and IQR of the NGEGIKGD decrease.

4.5. The NGEGIKGD's moment generating function

The moment generating function of the NGEGIKGD can be derived from the r^{th} moment by using the expansion of $e^{tx} = \sum_{r=0}^{\infty} \frac{t^r x^r}{r!}$ as:

$$\mu_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x)\,dx,$$
$$\mu_x(t) = ab\theta\alpha\beta\gamma\lambda \sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7,r=0}^{\infty} V_v \times \frac{t^r}{r!}\left[\frac{-1}{k(v_7+1)}\right]^{r+1} e^{\frac{\lambda}{k}(v_6+1)}\,\Gamma(r+1). \quad (26)$$

4.6. The NGEGIKGD's characteristic function

A characteristic function is a unique function that can characterize any probability distribution. We can also calculate the characteristic function based on the r^{th} moment of the NGEGIKGD as:

$$\phi_x(t) = E(e^{itx}) = \int_0^\infty e^{itx} f(x)\,dx,$$
$$\phi_x(t) = ab\theta\alpha\beta\gamma\lambda \sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7,r=0}^{\infty} V_v \times \frac{(it)^r}{r!}\left[\frac{-1}{k(v_7+1)}\right]^{r+1} e^{\frac{\lambda}{k}(v_6+1)}\,\Gamma(r+1). \quad (27)$$

where, V_v is defined in Eq. 16.

4.7. Order statistics

Let $X_1, \dots, X_n$ is a random sample (RS) from the NGEGIKGD, where, X_L is the L^{th} order statistics, then the pdf of the L^{th} order statistics is:

$$f_{L:n}(x) = \frac{n!}{(L-1)!\,(n-L)!}[F(x)]^{L-1}[1-F(x)]^{n-L} f(x),$$

where, F(x) and f(x) are defined in Eqs. 9, and 10. By using the binomial expansion in Eq. 13, then we have:

$$[1-F(x)]^{n-L} = \sum_{v_9=0}^{n-L} (-1)^{v_9}\binom{n-L}{v_9}[F(x)]^{v_9}. \quad (28)$$

thus,

$$f_{L:n}(x) = \sum_{v_9=0}^{n-L} \frac{(-1)^{v_9} n!\, f(x)}{(L-1)!\,(n-L)!}\binom{n-L}{v_9} \times [1-(1-\{1-[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b})^{\theta}]^{L+v_9-1},$$

where, f(x) is given by Eq. 15, if we define,

$$V_{Ov} = \frac{n!(-1)^{v_1+v_2+v_3+v_4+v_5+v_6+v_7+v_8+v_9}}{(L-1)!(n-L)!\Gamma(v_8+1)[\frac{\lambda(v_7+1)}{k}]^{-v_8}}\binom{n-L}{v_9}\binom{L+v_9-1}{v_1} \times \binom{\theta(v_1+1)-1}{v_2}\binom{b(v_2+1)-1}{v_3}\binom{a(v_3+1)-1}{v_4} \times \binom{\beta(v_4+1)-1}{v_5}\binom{\alpha(v_5+1)-1}{v_6}\binom{\gamma(v_6+1)-1}{v_7}. \quad (29)$$

then,

$$f_{L:n}(x) = ab\theta\alpha\beta\gamma\lambda \sum_{v_9=0}^{n-L}\sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7,v_8=0}^{\infty} V_{Ov}\, e^{\frac{\lambda}{k}(v_7+1)}\, e^{[k(v_8+1)]x}, \quad x \geq 0. \quad (30)$$

4.8. Rényi entropy

The entropy of a random variable is a measure of uncertainty variation. It can be applied in various applications, including engineering and physics. One popular entropy measure is the Rényi entropy (Abraham and Sankaran, 2006), in which, if X is a random variable, then Rényi entropy can be calculated as follows:

$$I_R(x) = \frac{1}{1-R}\log\left[\int_0^\infty [f(x)]^R dx\right], \quad R > 0, R \neq 1.$$

The Rényi entropy of the NGEGIKGD is:

$$[f(x)]^R = [ab\theta\alpha\beta\gamma\lambda]^R e^{Rkx} e^{-\frac{\lambda R}{k}(e^{kx}-1)}[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{R(\gamma-1)}\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{R(\alpha-1)} \times (1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{R(\beta-1)}[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{R(a-1)} \times \{1-[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{R(b-1)} \times (1-\{1-[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b})^{R(\theta-1)}.$$

By applying the binomial expansion in Eq. 13 six times and using the power series expansion in Eq. 14, we obtain:

$$[f(x)]^R = [ab\theta\alpha\beta\gamma\lambda]^R \sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7=0}^{\infty} V_{Rv}\, e^{\frac{\lambda}{k}(v_6+R)} e^{[k(v_7+R)]x}, \quad (31)$$

where,

$$V_{Rv} = \frac{(-1)^{v_1+v_2+v_3+v_4+v_5+v_6+v_7}}{\Gamma(v_7+1)[\frac{\lambda}{k}(v_6+R)]^{-v_7}} \binom{R(\theta-1)}{v_1} \binom{b(v_1+R)-R}{v_2} \binom{a(v_2+R)-R}{v_3} \times \binom{\beta(v_3+R)-R}{v_4} \binom{\alpha(v_4+R)-R}{v_5} \binom{\gamma(v_5+R)-R}{v_6}.$$

thus,

$$\int_0^\infty [f(x)]^R dx = [ab\theta\alpha\beta\gamma\lambda]^R \sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7=0}^{\infty} V_{Rv}\, e^{\frac{\lambda}{k}(v_6+R)} \int_0^\infty e^{[k(v_7+R)]x} dx.$$

if,

$$log \int_0^\infty [f(x)]^R dx = R\, log[ab\theta\alpha\beta\gamma\lambda] + log\left(\sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7=0}^{\infty} V_{Rv}\, e^{\frac{\lambda}{k}(v_6+R)} \int_0^\infty e^{[k(v_7+R)]x} dx \right),$$

the Rényi entropy of the NGEGIKGD is then given as:

$$I_R(x) = \frac{R}{1-R} log[ab\theta\alpha\beta\gamma\lambda] + \frac{1}{1-R} log\left(\sum_{v_1,v_2,v_3,v_4,v_5,v_6,v_7=0}^{\infty} V_{Rv} e^{\frac{\lambda}{k}(v_6+R)} \int_0^\infty e^{[k(v_7+R)]x} dx\right). \quad (32)$$

5. Estimation of the NGEGIKGD parameters

In this section, we estimate the eight unknown parameters for the NGEGIKGD using the ML method. Let $X_1, X_2, \ldots, X_n$ be an RS from the NGEGIKGD with unknown parameters $\Theta = (a, b, \theta, \alpha, \beta, \gamma, k, \lambda)$, then the log-likelihood function, $(logL(\Theta))$ will have the following form:

$$\begin{aligned} logL(\Theta) &= n\, log[ab\theta\alpha\beta\gamma\lambda] + k\sum_{i=1}^{n} x_i - \frac{\lambda}{k}\sum_{i=1}^{n} (e^{kx_i}-1) + (\gamma-1)\sum_{i=1}^{n} log[1-\xi] \\ &+(\alpha-1)\sum_{i=1}^{n} log\{1-[1-\xi]^\gamma\} + (\beta-1)\sum_{i=1}^{n} log(\xi_2) \\ &+(a-1)\sum_{i=1}^{n} log[1-(\xi_2)^\beta] \\ &+(b-1)\sum_{i=1}^{n} log\{1-[1-(\xi_2)^\beta]^a\} \\ &+(\theta-1)\sum_{i=1}^{n} log(1-\{1-[1-(\xi_2)^\beta]^a\}^b), \end{aligned} \quad (33)$$

where, $\xi = e^{-\frac{\lambda}{k}(e^{kx_i}-1)}$ and $\xi_2 = 1-\{1-[1-\xi]^\gamma\}^\alpha$. The maximum likelihood estimates (MLEs) of the NGEGIKGD's parameters can be directly obtained by employing non-linear optimization tool in the R program such as, "optim" or "nlm" which maximizes the $(logL(\Theta))$ in Eq. 33. If the parameters are restricted to be greater than zero, then either the parameters must be transformed, or the constraints used in the optimization tool. For more details, see MacDonald (2014).

6. Simulation studies of the NGEGIKGD

We ran numerous simulation studies for different sample sizes at 25, 50, 100, 200 and 500 with 1000 repetitions to examine the performance of the MLEs for the NGEGIKGD. This Monte Carol simulation was performed with various cases of the true parameter values as follows:

Case I: $a = 0.6, b = 2.8, \theta = 1.3, \alpha = 1.4, \beta = 0.7, \gamma = 0.3, k = 0.8, \lambda = 1.2$.
Case II: $a = 1.8, b = 0.8, \theta = 1.6, \alpha = 0.5, \beta = 1.6, \gamma = 0.6, k = 1.3, \lambda = 0.6$.
Case III: $a = 0.9, b = 1.2, \theta = 0.8, \alpha = 1.7, \beta = 0.5, \gamma = 0.9, k = 0.7, \lambda = 1.4$.

The MLE was obtained using the "optim" function in R, and the samples were generated from Eq. 17, where u is uniformly distributed within [0,1]. For each parameter, we calculated the mean estimates and the coverage rate (CR) of the confidence intervals (CI) of the mean estimates at a 95% confidence level. Additionally, we calculated the mean square error (MSE), the root MSE (RMSE), the standard deviation (SD) and the standard error (SE) using the following relations and the results reported in Tables 2-4, in which,

$$MSE = var(\hat{\theta}) + [Bias(\hat{\theta})]^2 = \frac{1}{N}\sum_{i=1}^{n} (\hat{\theta}-\theta_{tr})^2,$$

where, $Bias = \frac{1}{N}\sum_{i=1}^{n} (\hat{\theta}-\theta_{tr})$ and,

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{n} (\hat{\theta}-\theta_{tr})^2}.$$

In addition,

$$95\%\ CI = mean(\hat{\theta}) \pm 1.96 \times SE,$$

where, $SE = \frac{SD}{\sqrt{n}}$.

Table 2: Simulation study: True parameter, parameter estimates, MSE, RMSE, SD, SE, and CR of 95% CI of the NGEGIKGD for the case I

Sample size	Parameter	True	Estimate	MSE	RMSE	SD	SE	CR of 95% CI
n=25	a	0.6	2.0000	6.1799	2.4859	2.0542	0.4108	(1.1948,2.8053)
	b	2.8	2.1033	6.1113	2.4721	2.3719	0.4744	(1.1735,3.0330)
	θ	1.3	1.2976	3.1793	1.7831	1.7831	0.3566	(0.5986,1.9966)
	α	1.4	2.2797	6.5312	2.5556	2.3994	0.4799	(1.3391,3.2203)
	β	0.7	1.4514	3.3371	1.8268	1.6651	0.3330	(0.7987,2.1041)
	γ	0.3	1.7594	5.2671	2.2950	1.7712	0.3542	(1.0651,2.4537)
	k	0.8	1.2945	1.2488	1.1175	1.0022	0.2004	(0.9016,1.6873)
	λ	1.2	2.2323	6.2846	2.5069	2.2845	0.4569	(1.3368,3.1279)
n=50	a	0.6	1.6445	3.7582	1.9386	1.6331	0.2310	(1.1919,2.0972)
	b	2.8	2.0923	4.6710	2.1613	2.0421	0.2888	(1.5263,2.6584)
	θ	1.3	1.3675	2.7904	1.6704	1.6691	0.2360	(0.9049,1.8302)
	α	1.4	2.0648	4.1769	2.0438	1.9326	0.2733	(1.5291,2.6005)
	β	0.7	1.3463	2.5983	1.6119	1.4767	0.2088	(0.9370,1.7556)
	γ	0.3	1.4147	3.1910	1.7863	1.3958	0.1974	(1.0278,1.8016)
	k	0.8	1.1109	0.7445	0.8628	0.8049	0.1138	(0.8879,1.3340)
	λ	1.2	1.9858	4.0144	2.0036	1.8431	0.2606	(1.4749,2.4966)
n=100	a	0.6	1.5045	3.0626	1.7500	1.4982	0.1498	(1.2108,1.7981)
	b	2.8	2.0031	3.9448	1.9862	1.8193	0.1819	(1.6465,2.3596)
	θ	1.3	1.2864	1.7829	1.3353	1.3352	0.1335	(1.0247,1.5481)
	α	1.4	1.8390	2.9429	1.7155	1.6584	0.1658	(1.5139,2.1640)
	β	0.7	1.2176	1.8327	1.3538	1.2509	0.1251	(0.9724,1.4628)
	γ	0.3	1.2134	2.0637	1.4365	1.1088	0.1109	(0.9961,1.4307)
	k	0.8	1.0266	0.4052	0.6365	0.5948	0.0595	(0.9100,1.1432)
	λ	1.2	1.8873	3.1001	1.7607	1.6210	0.1621	(1.5696,2.2051)
n=200	a	0.6	1.3352	1.7830	1.3353	1.1147	0.0788	(1.1808,1.4897)
	b	2.8	2.1175	3.1305	1.7693	1.6324	0.1154	(1.8912,2.3437)
	θ	1.3	1.2660	1.1265	1.0614	1.0608	0.0750	(1.1190,1.4130)
	α	1.4	1.6893	1.9850	1.4089	1.3789	0.0975	(1.4981,1.8803)
	β	0.7	0.9375	0.8706	0.9331	0.9023	0.0638	(0.8124,1.0625)
	γ	0.3	1.1071	1.5426	1.2420	0.9439	0.0667	(0.9763,1.2380)
	k	0.8	0.9636	0.2487	0.4987	0.4711	0.0333	(0.8983,1.0289)
	λ	1.2	1.6447	1.8945	1.3764	1.3026	0.0921	(1.4642,1.8252)
n=500	a	0.6	1.0727	0.7724	0.8789	0.7409	0.0331	(1.0078,1.1377)
	b	2.8	2.2447	1.9844	1.4087	1.2946	0.0579	(2.1312,2.3581)
	θ	1.3	1.3620	0.9019	0.9497	0.9477	0.0424	(1.2789,1.4452)
	α	1.4	1.4697	0.9836	0.9918	0.9893	0.0442	(1.3829,1.5564)
	β	0.7	0.8310	0.5323	0.7296	0.7177	0.0321	(0.7681,0.8939)
	γ	0.3	0.8385	0.7782	0.8822	0.6987	0.0312	(0.7772,0.8998)
	k	0.8	0.8666	0.1072	0.3274	0.3205	0.0143	(0.8385,0.8947)
	λ	1.2	1.5323	1.0859	1.0421	0.9876	0.0442	(1.4458,1.6189)

Table 3: Simulation study: True parameter, parameter estimates, MSE, RMSE, SD, SE, and CR of 95% CI of the NGEGIKGD for the case II

Sample size	Parameter	True	Estimate	MSE	RMSE	SD	SE	CR of 95% CI
n=25	a	1.8	2.2701	6.2408	2.4982	2.4535	0.4907	(1.3083,3.2318)
	b	0.8	1.4830	3.8938	1.9733	1.8513	0.3703	(0.7573,2.2087)
	θ	1.6	1.3090	3.2806	1.8113	1.7877	0.3575	(0.6082,2.0098)
	α	0.5	2.0859	7.6884	2.7728	2.2745	0.4549	(1.1943,2.9775)
	β	1.6	2.0046	4.7450	2.1783	2.1404	0.4280	(1.1655,2.8436)
	γ	0.6	2.1160	6.6097	2.5709	2.0764	0.4153	(1.3020,2.9300)
	k	1.3	1.5563	0.9785	0.9892	0.9554	0.1911	(1.1819,1.9309)
	λ	0.6	1.8387	5.2061	2.2817	1.9162	0.3832	(1.0876,2.5899)
n=50	a	1.8	1.8758	3.8242	1.9556	1.9541	0.2764	(1.3341,2.4174)
	b	0.8	1.4087	2.7563	1.6602	1.5446	0.2184	(0.9806,1.8368)
	θ	1.6	1.3559	2.5606	1.6002	1.5815	0.2237	(0.9175,1.7943)
	α	0.5	1.8817	5.5046	2.3462	1.8962	0.2682	(1.3561,2.4073)
	β	1.6	1.8123	3.3701	1.8358	1.8234	0.2579	(1.3069,2.3178)
	γ	0.6	1.7607	4.0012	2.0003	1.6291	0.2304	(1.3091,2.2123)
	k	1.3	1.4847	0.5732	0.7571	0.7342	0.1038	(1.2812,1.6882)
	λ	0.6	1.7160	3.9789	1.9947	1.6533	0.2338	(1.2577,2.1743)
n=100	a	1.8	1.8479	3.1083	1.7630	1.7624	0.1762	(1.5024,2.1933)
	b	0.8	1.2301	1.8441	1.3580	1.2881	0.1288	(0.9777,1.4826)
	θ	1.6	1.3130	2.1398	1.4628	1.4344	0.1434	(1.0319,1.5941)
	α	0.5	1.5381	3.3258	1.8237	1.4994	0.1499	(1.2442,1.8320)
	β	1.6	1.6755	2.3201	1.5232	1.5213	0.1521	(1.3774,1.9737)
	γ	0.6	1.5979	3.0573	1.7485	1.4358	0.1436	(1.3164,1.8793)
	k	1.3	1.4553	0.4233	0.6507	0.6318	0.0632	(1.3315,1.5791)
	λ	0.6	1.5052	2.7872	1.6695	1.4028	0.1403	(1.2303,1.7802)
n=200	a	1.8	1.7042	2.1592	1.4694	1.4663	0.1037	(1.5009,1.9074)
	b	0.8	1.1401	1.2759	1.1296	1.0771	0.0762	(0.9908,1.2894)
	θ	1.6	1.3853	1.5972	1.2638	1.2454	0.0881	(1.2127,1.5579)
	α	0.5	1.2469	1.8654	1.3658	1.1435	0.0809	(1.0884,1.4054)
	β	1.6	1.5563	1.6188	1.2723	1.2716	0.0899	(1.3801,1.7326)
	γ	0.6	1.3576	1.7555	1.3250	1.0870	0.0769	(1.2070,1.5083)
	k	1.3	1.4521	0.3023	0.5498	0.5284	0.0374	(1.3789,1.5253)
	λ	0.6	1.2914	1.6611	1.2888	1.0877	0.0769	(1.1407,1.4422)
n=500	a	1.8	1.6850	1.4632	1.2096	1.2041	0.0539	(1.5794,1.7905)
	b	0.8	1.0347	0.6973	0.8351	0.8014	0.0358	(0.9645,1.1050)
	θ	1.6	1.4104	1.1820	1.0872	1.0705	0.0479	(1.3165,1.5049)
	α	0.5	1.0071	1.0110	1.0055	0.8683	0.0388	(0.9310,1.0832)
	β	1.6	1.4574	1.0565	1.0279	1.0179	0.0455	(1.3682,1.5467)
	γ	0.6	1.1767	1.0230	1.0115	0.8309	0.0372	(1.1039 ,1.2496)
	k	1.3	1.4369	0.1655	0.4069	0.3831	0.0171	(1.4033,1.4705)
	λ	0.6	1.1383	1.1030	1.0502	0.9018	0.0403	(1.0592,1.2173)

Table 4: Simulation study: True parameter, parameter estimates, MSE, RMSE, SD, SE, and CR of 95% CI of the NGEGIKGD for the case III

Sample size	Parameter	True	Estimate	MSE	RMSE	SD	SE	CR of 95% CI
n=25	a	0.9	2.0940	6.4188	2.5335	2.2346	0.4469	(1.2180,2.9699)
	b	1.2	1.5227	3.2008	1.7891	1.7597	0.3519	(0.8329,2.2125)
	θ	0.8	1.4181	3.3178	1.8215	1.7134	0.3427	(0.7464,2.0897)
	α	1.7	2.5083	7.3049	2.7028	2.5791	0.5158	(1.4973,3.5193)
	β	0.5	1.6240	4.5448	2.1319	1.8115	0.3623	(0.9140,2.3341)
	γ	0.9	1.7071	3.5598	1.8868	1.7054	0.3412	(1.0386,2.3757)
	k	0.7	1.3493	1.9691	1.4032	1.2440	0.2488	(0.8617,1.8370)
	λ	1.4	2.4011	6.7649	2.6009	2.4006	0.4801	(1.4601,3.3421)
n=50	a	0.9	1.6735	3.7569	1.9383	1.7772	0.2513	(1.1809,2.1662)
	b	1.2	1.5327	2.7801	1.6674	1.6338	0.2311	(1.0798,1.9856)
	θ	0.8	1.3964	2.4729	1.5725	1.4551	0.2058	(0.9931,1.7998)
	α	1.7	2.3031	4.9835	2.2324	2.1494	0.3040	(1.7073,2.8989)
	β	0.5	1.4341	3.3775	1.8378	1.5827	0.2238	(0.9954,1.8728)
	γ	0.9	1.4688	2.1821	1.4772	1.3633	0.1928	(1.0909,1.8467)
	k	0.7	1.1684	1.1447	1.0699	0.9619	0.1360	(0.9018,1.4351)
	λ	1.4	2.2361	5.5820	2.3626	2.2097	0.3125	(1.6236,2.8486)
n=100	a	0.9	1.6826	3.5183	1.8757	1.7046	0.1705	(1.3485,2.0167)
	b	1.2	1.3809	1.9080	1.3813	1.3694	0.1369	(1.1125,1.6493)
	θ	0.8	1.3427	2.0346	1.4264	1.3191	0.1319	(1.0842,1.6013)
	α	1.7	1.9708	2.9296	1.7116	1.6901	0.1690	(1.6396,2.3021)
	β	0.5	1.2171	1.9755	1.4055	1.2088	0.1209	(0.9802,1.4541)
	γ	0.9	1.3164	1.3778	1.1738	1.0975	0.1097	(1.1013,1.5315)
	k	0.7	0.9779	0.6137	0.7834	0.7325	0.0732	(0.8343,1.1214)
	λ	1.4	2.0737	3.8517	1.9626	1.8433	0.1843	(1.7124,2.4349)
n=200	a	0.9	1.5121	2.1802	1.4766	1.3439	0.0950	(1.3259,1.6984)
	b	1.2	1.2032	1.3053	1.1425	1.1425	0.0808	(1.0449,1.3615)
	θ	0.8	1.2989	1.6031	1.2661	1.1637	0.0823	(1.1376,1.4602)
	α	1.7	1.9048	2.2610	1.5037	1.4897	0.1053	(1.6983,2.1112)
	β	0.5	1.2219	1.7605	1.3268	1.1132	0.0787	(1.0676,1.3762)
	γ	0.9	1.1785	0.8921	0.9445	0.9025	0.0638	(1.0534,1.3035)
	k	0.7	0.8813	0.3722	0.6101	0.5826	0.0412	(0.8006,0.9620)
	λ	1.7	1.8307	2.5253	1.5891	1.5296	0.1082	(1.6187,2.0427)
n=500	a	0.9	1.2393	1.0661	1.0325	0.9752	0.0436	(1.1539,1.3248)
	b	1.2	1.1233	0.7910	0.8894	0.8860	0.0396	(1.0456,1.2009)
	θ	0.8	1.1990	0.9215	0.9600	0.8731	0.0390	(1.1225,1.2755)
	α	1.7	1.8060	1.5381	1.2402	1.2357	0.0553	(1.6976,1.9143)
	β	0.5	1.0199	0.9553	0.9774	0.8277	0.0370	(0.9473,1.0924)
	γ	0.9	1.1303	0.6142	0.7837	0.7491	0.0335	(1.0647,1.1960)
	k	0.7	0.8107	0.1658	0.4072	0.3918	0.0175	(0.7764,0.8451)
	λ	1.4	1.7118	1.4313	1.1964	1.1550	0.0517	(1.6105,1.8130)

The results reported in Tables 2-4 reveal that the estimates are quite stable and become closer to the actual value of the parameters when n increases. In addition, the MSE, RMSE, SD and SE decrease and the CR become shorter as n increases, so, the ML method provides a good estimate and is appropriate for estimating the parameters for the NGEGIKGD.

7. Applications of the NGEGIKGD

Three real datasets are used in this section to demonstrate the flexibility of the proposed distribution in comparison to other models. The NGEGIKGD is compared with some of its sub-models including

- the Gompertz distribution (GD) with the pdf in Eq. 4,
- the generalized inverted Kumarswamy Gompertz distribution (GIKGD) with the pdf in Eq. 6,
- the exponentiated generalized inverted Kumaraswamy Gompertz distribution (EGIKGD) with the pdf as follows:

$$\begin{aligned} f(x) &= a\alpha\beta\gamma\lambda e^{kx} e^{-\frac{\lambda}{k}(e^{kx}-1)}[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma-1}\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha-1} \\ &\times (1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta-1}[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a-1}, \quad x \geq 0, \end{aligned} \tag{34}$$

and the exponentiated generalized inverted Kumaraswamy Gompertz distribution (EGGIKGD) with, the pdf as follows:

$$\begin{aligned} f(x) &= ab\alpha\beta\gamma\lambda e^{kx} e^{-\frac{\lambda}{k}(e^{kx}-1)}[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma-1}\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha-1} \\ &\times (1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta-1}[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a-1} \\ &\times \{1-[1-(1-\{1-[1-e^{-\frac{\lambda}{k}(e^{kx}-1)}]^{\gamma}\}^{\alpha})^{\beta}]^{a}\}^{b-1}, \quad x \geq 0. \end{aligned} \tag{35}$$

We also compared the NGEGIKGD with a related distribution, which is the Kumaraswamy Weibull BurrXII distribution (KWBXIID) by Hassan and Elgarhy (2016) with the following pdf:

$$\begin{aligned} f(x) &= abc\alpha\beta\sigma\mu^{-c}x^{c-1}e^{-\alpha[(1+(\frac{x}{\mu})^c)^{\sigma}-1]^{\beta}}(1+(\frac{x}{\mu})^c)^{\sigma-1}[(1+(\frac{x}{\mu})^c)^{\sigma}-1]^{\beta-1} \\ &\times [1-e^{-\alpha[(1+(\frac{x}{\mu})^c)^{\sigma}-1]^{\beta}}]^{a-1}\{1-[1-e^{-\alpha[(1+(\frac{x}{\mu})^c)^{\sigma}-1]^{\beta}}]^{a}\}^{b-1}, \quad x > 0. \end{aligned} \tag{36}$$

We calculated the MLEs, the SEs, and the length of confidence intervals (LCIs) of the point estimates at a confidence level of 95% as $LCI = (\text{upperlimit} - \text{lowerlimit})$ for each parameter. Furthermore, we compared the NGEGIKGD with other distributions according to some criteria; the negative log-likelihood function ($-LogL$), Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criterion (CAIC), Hannan-Quinn information criterion (HQIC) and the Kolmogorov-Smirnov (K-S) statistic as $D_n = \sup_x\{|F_n(x) - F(x|\Theta)|\}$ with its P-value. These statistics or criteria are frequently used to evaluate a distribution's performance in modeling a dataset. As a general rule, distributions with a lower value of all these criteria, except the highest value for the P-values, provide the best fit to the data. The results of the analyzed datasets are reported in Tables 5–13 and Figs. 4–6.

7.1. The first dataset

The first dataset reported by Bhaumik et al. (2009), contains the vinyl chloride information obtained from clean upgrading, observing wells in milligrams per Liter (mg/L).

Table 5: Descriptive statistics of the first dataset

Sample size	Mean	Median	IQR	SD	Skewness	Kurtosis
34	1.8790	1.1500	1.9750	1.9526	1.6037	5.0054

Table 6: The goodness-of-fit measures for the first dataset

Distribution	-LogL	AIC	CAIC	BIC	HQIC	K-S	P-value
GD	-76.2349	156.4697	161.5224	159.5224	157.5108	0.5024	7.0068e-08
GIKGD	-79.3854	168.7709	181.4027	176.4027	171.3735	0.5461	3.1092e-09
EGIKGD	-61.1751	134.3502	149.5083	143.5083	137.4734	0.1568	3.7332e-01
EGGIKGD	-60.9433	135.8866	153.5711	146.5711	139.5304	0.1505	4.2473e-01
KWBXIID	-57.3444	128.6887	146.3732	139.3732	132.3324	0.0892	9.4971e-01
NGEGIKGD	-54.5293	125.0585	145.2694	137.2694	129.2228	0.0762	9.8911e-01

Table 7: The MLE, SE (in parentheses) and LCI [in square brackets] for the first dataset

Distribution	MLE, SE in ( ) and LCI in [ ]							
	$\hat{k}$=0.0911	$\hat{\lambda}$=0.1167						
GD	(0.1210)	(0.0217)	-	-	-	-	-	-
	[0.4745]	[0.0851]						
	$\hat{\alpha}$=1.1677	$\hat{\beta}$=0.3438	$\hat{\gamma}$ =0.4664	$\hat{k}$ =0.1658	$\hat{\lambda}$ =0.1462			
GIKGD	(0.2286)	(0.0642)	(0.1104)	(0.1842)	(0.0611)	-	-	-
	[0.8961]	[0.2515]	[0.4329]	[0.7221]	[0.2395]			
	$\hat{a}$ =0.2847	$\hat{\alpha}$=0.8081	$\hat{\beta}$ =0.5551	$\hat{\gamma}$ =0.5911	$\hat{k}$ =0.2354	$\hat{\lambda}$ =0.8753		
EGIKGD	(0.1348)	(0.3928)	(0.2360)	(0.2661)	(0.1011)	(0.4748)	-	-
	[0.5283]	[1.5399]	[0.9252]	[1.0431]	[0.3962]	[1.8613]		
	$\hat{a}$ =0.4624	$\hat{b}$=1.2160	$\hat{\alpha}$ =0.4624	$\hat{\beta}$=0.6087	$\hat{\gamma}$=0.3956	$\hat{k}$ =0.2023	$\hat{\lambda}$=1.1677	
EGGIKGD	(0.4729)	(0.4401)	(0.4290)	(0.2834)	(0.2374)	(0.1161)	(1.0402)	-
	[1.8537]	[1.7251]	[1.6814]	[1.1108]	[0.9306]	[0.4552]	[4.0775]	
	$\hat{a}$ =0.1230	$\hat{b}$ =3.4622	$\hat{\alpha}$ =5.4712	$\hat{\beta}$= 1.1160	$\hat{c}$=6.4271	$\hat{\mu}$=14.8036	$\hat{\sigma}$=5.2755	
KWBXIID	(0.0224)	(1.0669)	(0.3070)	(0.0324)	(0.0129)	(0.1795)	(0.1730)	-
	[0.0879]	[4.1823]	[1.2035]	[0.1270]	[0.0506]	[0.7037]	[0.6784]	
	$\hat{a}$=0.2802	$\hat{b}$ =0.3341	$\hat{\theta}$ =1.2156	$\hat{\alpha}$ =0.2940	$\hat{\beta}$ =1.5928	$\hat{\gamma}$ =2.9319	$\hat{k}$ =0.1102	$\hat{\lambda}$ =2.2058
NGEGIKGD	(0.2129)	(0.2598)	(0.6702)	(0.2204)	(1.2494)	(0.5012)	(0.0373)	(0.0379)
	[0.8347]	[1.0182]	[2.6271]	[0.8641]	[4.8975]	[1.9647]	[0.1460]	[0.1484]

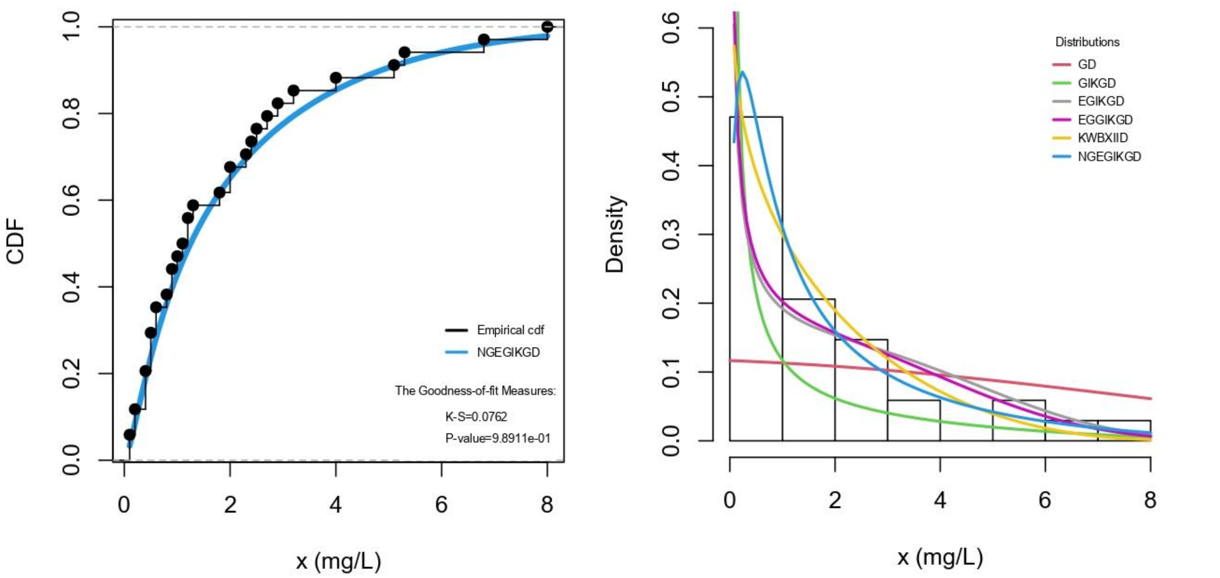

Fig. 4: Comparison of the NGEGIKGD with the other distributions for the first dataset; (Left): Cdf for the NGEGIKGD. (Right): Observed and expected frequencies for each model

7.2. The second dataset

The second dataset discussed in Murthy et al. (2004), represents the failure time between the repairable items.

7.3. The third dataset

The third dataset, initially used by Meeker and Escobar (1998), represents the failure times and running times of a larger system.

Table 8: Descriptive statistics of the second dataset

Sample size	Mean	Median	IQR	SD	Skewness	Kurtosis
30	1.5427	1.2350	1.2250	1.1277	1.2955	4.3192

Table 9: The goodness-of-fit measures for the second dataset

Distribution	-LogL	AIC	CAIC	BIC	HQIC	K-S	P-value
GD	-69.8603	143.7205	148.5229	146.5229	144.6170	0.6188	2.1033e-10
GIKGD	-66.0702	142.1404	154.1463	149.1463	144.3816	0.4040	1.1183e-04
EGIKGD	-44.5641	101.1282	115.5354	109.5354	103.8178	0.2147	1.2587e-01
EGGIKGD	-49.8944	113.7888	130.5972	123.5972	116.9266	0.3190	4.4625e-03
KWBXIID	-42.8519	99.7039	116.5122	109.5122	102.8417	0.1160	8.1394e-01
NGEGIKGD	-39.4716	94.9432	114.1528	106.1528	98.5293	0.0635	9.9973e-01

Table 10: The MLE, SE (in parentheses), and LCI [in square brackets] for the second dataset

Distribution	MLE, SE in ( ) and LCI in [ ]							
GD	$\hat{k}$ =0.9193 (0.1348) [0.5286]	$\hat{\lambda}$ =0.0305 (0.0062) [0.0244]	-	-	-	-	-	-
GIKGD	$\hat{\alpha}$=16.0856 (12.1334) [47.5631]	$\hat{\beta}$ =1.0952 (0.5340) [2.0933]	$\hat{\gamma}$ =3.0657 (1.1632) [4.5599]	$\hat{k}$ =0.2418 (0.2350) [0.9214]	$\hat{\lambda}$ =0.1774 (0.2664) [1.0445]	-	-	-
EGIKGD	$\hat{a}$ =0.3202 (0.1316) [0.5160]	$\hat{\alpha}$=0.8887 (0.4759) [1.8655]	$\hat{\beta}$ =0.9695 (1.1330) [4.4415]	$\hat{\gamma}$=0.7254 (0.7892) [3.0935]	$\hat{k}$ =0.6261 (0.1796) [0.7040]	$\hat{\lambda}$ =0.6433 (0.3155) [1.2367]	-	-
EGGIKGD	$\hat{a}$ =0.7455 (0.0411) [0.1613]	$\hat{b}$ =0.3814 (0.0932) [0.3655]	$\hat{\alpha}$ =0.7450 (0.2043) [0.8008]	$\hat{\beta}$ =1.1729 (0.3264) [1.2794]	$\hat{\gamma}$ =0.8446 (0.2963) [1.1616]	$\hat{k}$ =0.6391 (0.3657) [1.4336576]	$\hat{\lambda}$=0.3430 (0.1395) [0.5467]	-
KWBXIID	$\hat{a}$ =6.9831 (5.5968) [21.9396]	$\hat{b}$ =0.5333 (0.3192) [1.2512]	$\hat{\alpha}$ =2.0648 (2.8917) [11.3356]	$\hat{\beta}$ =1.6989 (0.5192) [2.0353]	$\hat{c}$ =3.3006 (1.4014) [5.4936]	$\hat{\mu}$ =0.0135 (0.0142) [0.0558]	$\hat{\sigma}$ =0.0562 (0.0230) [0.0903]	-
NGEGIKGD	$\hat{a}$ =0.4301 (0.4606) [1.8054]	$\hat{b}$=2.3119 (1.2098) [4.7423]	$\hat{\theta}$=0.2741 (0.3238) [1.2692]	$\hat{\alpha}$=2.0623 (0.0161) [0.0632]	$\hat{\beta}$=0.5570 (0.0154) [0.0604]	$\hat{\gamma}$=1.6421 (0.0229) [0.0898]	$\hat{k}$ =0.0855 (0.0039) [0.0155]	$\hat{\lambda}$=2.9868 (0.0155) [0.0612]

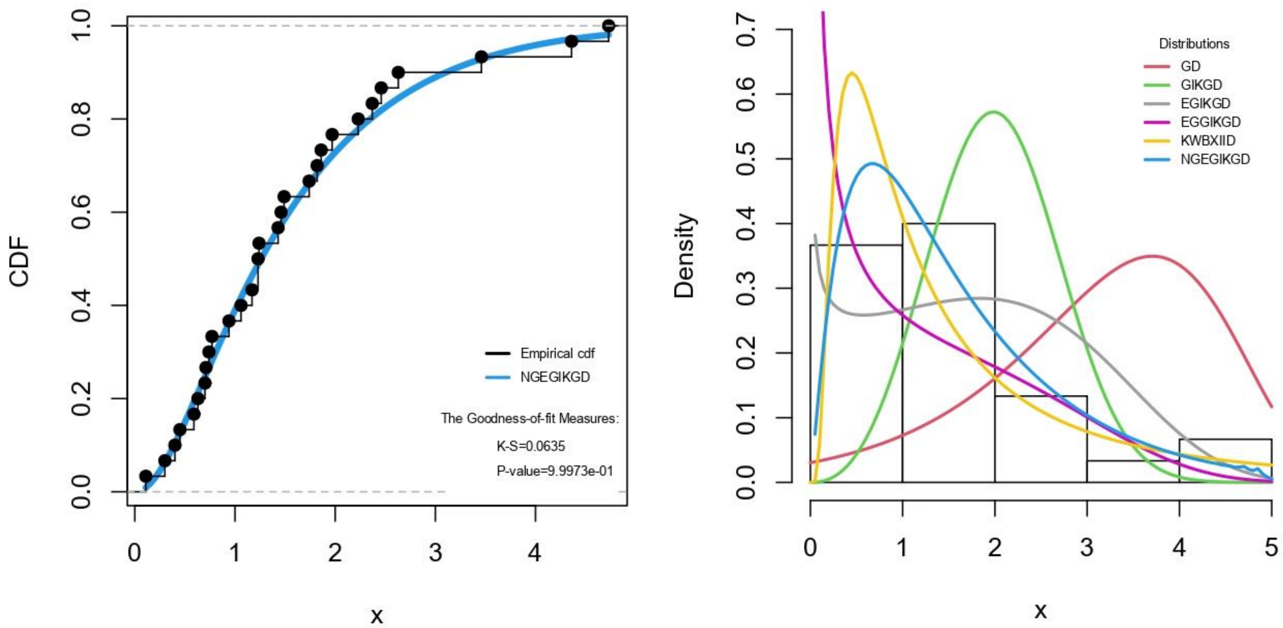

Fig. 5: Comparison of the NGEGIKGD with the other distributions for the second dataset; (Left): Cdf for the NGEGIKGD. (Right): Observed and expected frequencies for each model

Table 11: Descriptive statistics of the third dataset

Sample size	Mean	Median	IQR	SD	Skewness	Kurtosis
30	1.7703	1.9650	2.2950	1.1499	-0.2840	1.4537

Table 12: The goodness-of-fit measures for the third dataset

Distribution	-LogL	AIC	CAIC	BIC	HQIC	K-S	P-value
GD	-71.0404	146.0807	150.8831	148.8831	146.9772	0.3821	0.0003
GIKGD	-44.1388	98.2776	110.2836	105.2836	100.5189	0.2654	0.0292
EGIKGD	-41.8710	95.7420	110.1492	104.1492	98.4315	0.1991	0.1855
EGGIKGD	-37.0195	88.0390	104.8474	97.8474	91.1768	0.2100	0.1418
KWBXIID	-40.4191	94.8382	111.6466	104.6466	97.9760	0.2246	0.0970
NGEGIKGD	-34.2318	84.4636	103.6732	95.6732	88.0497	0.1518	0.4939

Table 13: The MLE, SE (in parentheses) and LCI [in square brackets] for the third dataset

Distribution	MLE, SE in ( ) and LCI in [ ]							
GD	$\hat{k}$=2.7452 (0.1470) [0.5764]	$\hat{\lambda}$=0.0026 (0.0009) [0.0035]	-	-	-	-	-	-
GIKGD	$\hat{\alpha}$=0.0469 (0.0116) [0.0456]	$\hat{\beta}$=0.7227 (0.1552) [0.6086]	$\hat{\gamma}$=2.2372 (0.0601) [0.2355]	$\hat{k}$ =0.1038 (0.0046) [0.0180]	$\hat{\lambda}$=10.3699 (0.0344) [0.1349]	-	-	-
EGIKGD	$\hat{a}$=0.1578 (0.0583) [0.2285]	$\hat{\alpha}$=0.2387 (0.0977) [0.3829]	$\hat{\beta}$=0.1309 (0.0660) [0.2587]	$\hat{\gamma}$=11.3132 (0.1731) [0.6786]	$\hat{k}$ =0.2175 (0.0009) [0.0034]	$\hat{\lambda}$=8.5949 (0.0223) [0.0876]	-	-
EGGIKGD	$\hat{a}$=5.5073 (4.9746) [19.5005]	$\hat{b}$=0.2480 (0.5677) [2.2254]	$\hat{\alpha}$=15.0891 (3.5360) [13.8612]	$\hat{\beta}$=0.7324 (1.3300) [5.2136]	$\hat{\gamma}$=3.4494 (0.0681) [0.2669]	$\hat{k}$ =1.0680 (0.2101) [0.8237]	$\hat{\lambda}$=0.0133 (0.0108) [0.0423]	-
KWBXIID	$\hat{a}$=0.1904 (0.0509) [0.1997]	$\hat{b}$ =1.7932 (1.4993) [5.8771]	$\hat{\alpha}$ =0.6073 (0.3247) [1.2728]	$\hat{\beta}$=3.6473 (0.0441) [0.1727]	$\hat{c}$=1.4256 (0.1521) [0.5963]	$\hat{\mu}$ =18.2797 (0.0628) [0.2462]	$\hat{\sigma}$=7.1158 (0.2281) [0.8940]	-
NGEGIKGD	$\hat{a}$ =0.0963 (0.1001) [0.3923]	$\hat{b}$ =0.4818 (0.4705) [1.8442]	$\hat{\theta}$ =0.6496 (0.5398) [2.1160]	$\hat{\alpha}$=0.6983 (0.4094) [1.6049]	$\hat{\beta}$ =0.1877 (0.2500) [0.9801]	$\hat{\gamma}$ =6.3527 (0.0528) [0.2072]	$\hat{k}$ =1.5033 (0.0403) [0.1578]	$\hat{\lambda}$ =0.6086 (0.0402) [0.1577]

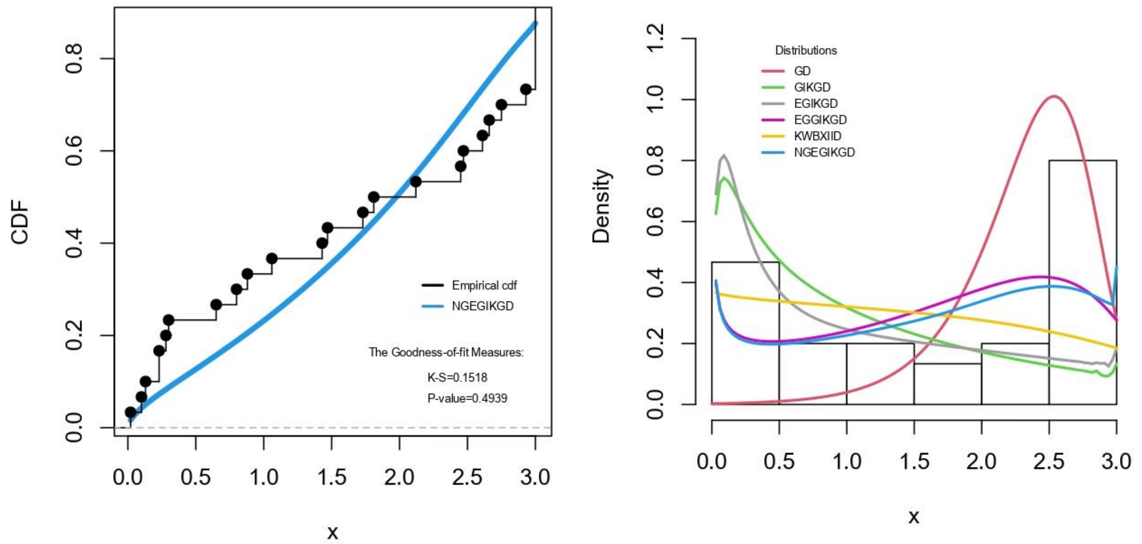

Fig. 6: Comparison of the NGEGIKGD with the other distributions for the third dataset; (Left): Cdf for the NGEGIKGD. (Right): Observed and expected frequencies for each model

The MLE and SE of the distribution parameters for all datasets are listed in Tables 7, 10, and 13. Furthermore, Tables 6, 9, and 12 show that the measures AIC, CAIC, BIC, HQIC, and K-S of the NGEGIKGD are the smallest, also, the NGEGIKGD has the best P-values. Moreover, based on Figs. 4–6, it is clear that the NGEGIKGD provides the closest fits to the actual distribution of the analyzed datasets. As a result, the NGEGIKGD is the best model for the investigated real-life datasets compared to competitive distributions.

8. Conclusions

This study investigated the effect of a new modeling approach that adds three shape parameters to a generalized distribution by introducing the NGEGIKGD. The proposed NGEGIKGD was inspired by the notion that generalization allows for greater flexibility in analyzing practical data. The NGEGIKGD's hazard rate function takes a variety of forms, which supports its application in modeling different hazard behaviors in real-world scenarios, such as human mortality and biological applications. Several useful statistical and mathematical properties of the NGEGIKGD were obtained. Additionally, we estimated the NGEGIKGD's parameters, and the performance of the estimators were examined via various simulation studies. Finally, the usefulness and flexibility of the NGEGIKGD were illustrated by means of three real-life datasets. The suggested NGEGIKGD with the extra parameters is capable of providing a better fit than several lifetime models.

We expect that this generalization will lead to further lifetime and reliability analysis applications.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Abraham B and Sankaran PG (2006). Renyi's entropy for residual lifetime distribution. Statistical Papers, 47(1): 17-29. https://doi.org/10.1007/s00362-005-0270-y

Alzaatreh A, Lee C, and Famoye F (2013). A new method for generating families of continuous distributions. Metron, 71(1): 63-79. https://doi.org/10.1007/s40300-013-0007-y

Bhaumik DK, Kapur K, and Gibbons RD (2009). Testing parameters of a gamma distribution for small samples. Technometrics, 51(3): 326-334. https://doi.org/10.1198/tech.2009.07038

Cordeiro GM, Ortega EM, and da Cunha DC (2013). The exponentiated generalized class of distributions. Journal of Data Science, 11(1): 1-27. https://doi.org/10.6339/JDS.201301_11(1).0001

De Andrade TA, Chakraborty S, Handique L, and Gomes-Silva F (2019). The exponentiated generalized extended Gompertz distribution. Journal of Data Science, 17(2): 299-330. https://doi.org/10.6339/JDS.201904_17(2).0004

Eghwerido JT, Nzei LC, and Agu FI (2021). The alpha power Gompertz distribution: Characterization, properties, and applications. Sankhya A, 83(1): 449-475. https://doi.org/10.1007/s13171-020-00198-0

El-Gohary A, Alshamrani A, and Al-Otaibi AN (2013). The generalized Gompertz distribution. Applied Mathematical Modelling, 37(1-2): 13-24. https://doi.org/10.1016/j.apm.2011.05.017

Galton F (1883). Inquiries into human faculty and its development. Macmillan, New York, USA. https://doi.org/10.1037/14178-000

Gompertz B (1825). 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. Philosophical Transactions of the Royal Society of London, 115: 513-583. https://doi.org/10.1098/rstl.1825.0026

Gupta RC, Gupta PL, and Gupta RD (1998). Modeling failure time data by Lehman alternatives. Communications in Statistics-Theory and Methods, 27(4): 887-904. https://doi.org/10.1080/03610929808832134

Hassan AS and Elgarhy M (2016). Kumaraswamy Weibull-generated family of distributions with applications. Advances and Applications in Statistics, 48(3): 205-239. https://doi.org/10.17654/AS048030205

Jamal F, Arslan Nasir M, Ozel G, Elgarhy M, and Mamode Khan N (2019). Generalized inverted Kumaraswamy generated family of distributions: Theory and applications. Journal of Applied Statistics, 46(16): 2927-2944. https://doi.org/10.1080/02664763.2019.1623867

Joshi RK and Kumar V (2020). Lindley Gompertz distribution with properties and applications. International Journal of Statistics and Applied Mathematics, 5(6): 28-37. https://doi.org/10.22271/maths.2020.v5.i6a.610

Khan MS, King R, and Hudson IL (2017). Transmuted generalized Gompertz distribution with application. Journal of Statistical Theory and Applications, 16(1): 65-80. https://doi.org/10.2991/jsta.2017.16.1.6

MacDonald IL (2014). Does Newton–Raphson really fail? Statistical Methods in Medical Research, 23(3): 308-311. https://doi.org/10.1177/0962280213497329 **PMid:24837788**

Meeker WQ and Escobar LA (1998). Statistical methods for reliability data. Wiley, New York, USA.

Moors JJA (1988). A quantile alternative for kurtosis. Journal of the Royal Statistical Society (Series D), 37: 25-32. https://doi.org/10.2307/2348376

Murthy DP, Xie M, and Jiang R (2004). Weibull models. John Wiley and Sons, Hoboken, USA.

Nasiru S, Mwita PN, and Ngesa O (2019). Exponentiated generalized exponential Dagum distribution. Journal of King Saud University-Science, 31(3): 362-371. https://doi.org/10.1016/j.jksus.2017.09.009

Rezaei S, Marvasty AK, Nadarajah S, and Alizadeh M (2017). A new exponentiated class of distributions: Properties and applications. Communications in Statistics-Theory and Methods, 46(12): 6054-6073. https://doi.org/10.1080/03610926.2015.1116579

Shama MS, Dey S, Altun E, and Afify AZ (2022). The Gamma–Gompertz distribution: Theory and applications. Mathematics and Computers in Simulation, 193: 689-712. https://doi.org/10.1016/j.matcom.2021.10.024