International Journal of Advanced and Applied Sciences, 10(6) 2023, Pages: 71-79

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

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

The exponentiated new exponential-gamma distribution: Properties and applications

Lulah Alnaji [1,*], Amani S. Alghamdi [2]

[1]*Department of Mathematics, University of Hafr Al Batin, Hafar Al-Batin, Saudi Arabia*
[2]*Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia*

ARTICLE INFO

Article history:
Received 9 December 2022
Received in revised form
6 April 2023
Accepted 8 April 2023

Keywords:
Exponentiated
Exponential
Gamma
Moment-generating function
r^{th} moment

ABSTRACT

We propose a novel lifetime model by extending the new exponential-gamma distribution to the exponentiated new exponential-gamma distribution. This extension allows for the derivation of a more flexible density function that combines the characteristics of the exponential and gamma distributions. We present various statistical properties of the newly proposed method, including the cumulative function, probability density function, moment-generating function, and moments. Additionally, we discuss the estimation of parameters using maximum likelihood. To compare the performance of our newly developed model with existing probability distributions (gamma, exponential, Lindley, generalized gamma, generalization of the generalized gamma, and new exponential-gamma distribution), we employ model selection criteria such as the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion (AICC), and the Bayesian Information Criterion (BIC). The application of these criteria to different models demonstrates that our proposed model outperforms the other six models across various datasets. For instance, in the first dataset, the AIC, AICC, and BIC values for our model are 366.975, 373.805, and 373.805, respectively, whereas the values for the other six models (exponential, Lindley, generalized gamma, generalization of the generalized gamma) range from 503.012 to 834.327. We conduct simulation studies to assess the efficiency of our proposed model. Furthermore, we apply the proposed method to three real data applications to further examine its effectiveness. It is important to note that the quantile function of the proposed model does not have a closed-form solution, requiring the computation of the quantile function through the Newton-Raphson iterative approach.

1. Introduction

The gamma distribution (α, β) was developed by Thom (1958), it is encountered as a lifetime distribution and has been employed extensively ever since. Its theories are widely studied, applied to different fields of knowledge, and have been developed for a number of distributions. The survival function of gamma distribution has no closed-form expression which is the negative aspect of this acknowledged distribution, moreover, it has a decreasing failure rate (Dahiya and Gurland, 1972). It was extended and generalized by Stacy (1962), to define the probability density function (pdf), which received special attention,

$$f_G(x;\alpha,\theta)=\frac{1}{\Gamma(\alpha)}\theta^{\alpha}x^{\alpha-1}e^{-\theta x};\quad x>0,\ \alpha>0,\ \theta>0 \qquad (1)$$

The gamma distribution is conventionally selected in lifetime data analysis due to monotone risk functions which are favored by this type of data. The pitfall has no closed-form risk function, and sometimes the computation of the numerical integration is necessary. The exponential distribution is a reduced form of gamma distribution when the shape parameter $\beta = 1$, and its survival and hazard functions are straightforward which makes exponential distribution a bit acknowledged compared to a gamma distribution and can be used quite efficaciously to analyze lifetime data in exchange for gamma to overcome such difficulties, which has extended to a number of exponential distributions such as Gupta and Kundu (2001),

* Corresponding Author.
Email Address: laalnaji@uhb.edu.sa (L. Alnaji)
https://doi.org/10.21833/ijaas.2023.06.009
Corresponding author's ORCID profile:
https://orcid.org/0000-0003-3367-192X

Nadarajah and Haghighi (2011), Khan et al. (2017), and Eghwerido et al. (2022) among others. This led to a generalizing gamma distribution (Ghitany, 1998), a generalization of the generalized gamma (Rama and Kamlesh, 2019), and a proposed mixture of gamma density with various density functions such as Abdullahi and Phaphan (2022), Susanto et al. (2022), and Yakubu et al. (2022).

To obtain further flexible density function of the gamma distribution, the exponentiated gamma distribution has been proposed and studied in some of its aspects and statistical properties by Gupta et al. (1998), to defeat the features that render gamma distribution a bit less effective. This development of the model made it a workable model that accommodates both monotonic and non-monotonic failure rates, with (pdf) defined as:

$$f_{EG}(x;\alpha,\theta) = \alpha\theta^2 x e^{-\theta x}\left[1-e^{-\theta x}(\theta x+1)\right]^{\alpha-1}; \quad x>0,\ \alpha>0,\ \theta>0 \tag{2}$$

with cumulative density function (cdf), defined as:

$$F_{EG}(x;\alpha,\theta) = \left[1-e^{-\theta x}(\theta x+1)\right]^{\alpha}; \quad x>0,\ \alpha>0,\ \theta>0 \tag{3}$$

Umar and Yahya (2021) extended the exponentiated gamma distribution defining a new distribution called the new exponential-gamma distribution with pdf as:

$$f_{NEG}(x;\alpha,\theta) = \frac{\theta}{\theta+\Gamma(\alpha)}\left(\theta+\theta^{\alpha-1}x^{\alpha-1}\right)e^{-\theta x}; \quad x>0,\ \alpha>0,\ \theta>0 \tag{4}$$

with cdf, defined as:

$$F_{NEG}(x;\alpha,\theta) = \frac{\theta}{\theta+\Gamma(\alpha)}\left(\left(1-e^{-\theta x}\right)+\theta^{\alpha-1}\int_0^x t^{\alpha-1}e^{-\theta t}dt\right) \tag{5}$$

The new exponential-gamma distribution showed its superiority over some distributions; namely exponential, gamma, Lindley, exponentiated gamma, generalization of the gamma distribution, and a generalization of generalized gamma distribution (Umar and Yahya, 2021). The motivation of this paper is that the pursuit of proposing more efficient and flexible probability distribution continues to exist in the field of probability theory and statistics. This paper focuses on generalizing new exponential-gamma distribution (Umar and Yahya, 2021) to achieve more efficiency. The remaining sections of this article are structured as follows. Section 2 presents the cdf and the corresponding pdf of the exponentiated new exponential-gamma distribution (EEG). Some statistical properties of the new proposed model are presented in Section 3. The simulation studies are conducted to assess the efficiency of the proposed model in Section 4. Section 5 presents applications studies using real data and Section 6 presents the important conclusions from this study.

2. The exponentiated new exponential-gamma distribution

The pdf of the proposed model is defined as:

$$f_{NEG}(x;\alpha,\theta) = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b\left(\theta+\theta^{\alpha-1}x^{\alpha-1}\right)e^{-\theta x}\left(\left(1-e^{-\theta x}\right)+\frac{1}{\theta}\gamma(\alpha,\theta x)\right)^{b-1} \tag{6}$$

where, $x>0, \alpha>0, \theta>0$. θ and α are the scale and shape parameters, respectively, and b is an additional shape parameter. $\gamma(\alpha,\theta x)$ is the lower incomplete gamma function (DiDonato and Morris, 1986). Fig. 1 exhibits a number of the possible shapes of the (pdf) of EEG distribution for various values of the parameters α,θ and b. The corresponding cdf for this generalization is given as:

$$F(t;\alpha,\theta) = \left[\frac{\theta}{\theta+\Gamma(\alpha)}\left(\left(1-e^{-\theta x}\right)+\theta^{\alpha-1}\zeta\right)\right]^b \tag{7}$$

where, $\zeta = \int_0^x t^{\alpha-1}e^{-\theta t}dt$. When $b=1$, this model reduces to the NEG model (Umar and Yahya, 2021).

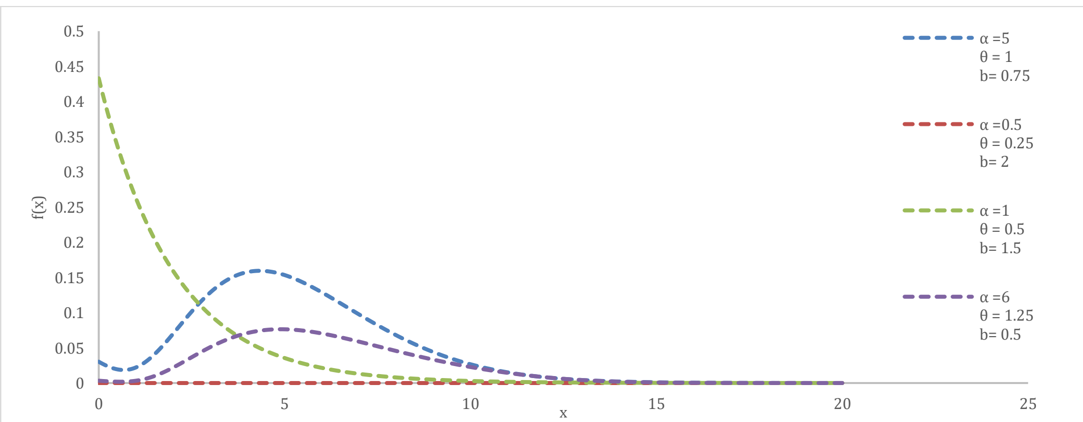

Fig. 1: EEG density function for various values of α, θ, and b

3. Statistical properties

In this section, some statistical properties of the new proposed model are studied, such as moment generating function, r^{th} moment, hazard rate function, the mean residual life function, and maximum likelihood estimation.

3.1. Moment generating function

The moment generating function (mgf) of the random variable X, is defined as follows:

$$M_X(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

Therefore, the mgf of the random variable X, with the pdf of the proposed distribution defined by Eq. 6:

$$M_X(t) = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \int_0^{\infty} e^{tx}\,(\theta + \theta^{\alpha-1}x^{\alpha-1})e^{-\theta x}\{(1 - e^{-\theta x}) + \theta^{\alpha-1}\zeta\}^{b-1}dx \quad (8)$$

where, $\zeta = \int_0^{\infty} t^{\alpha-1}e^{-\theta t}dt$. Using the generalized expansion:

$$(x+y)^s = \sum_{k=0}^{\infty}\binom{s}{k} x^k y^{s-k}; \quad if \quad |x| < |y| \quad and \quad x,y,s \in \mathbb{R} \quad (9)$$

yields,

$$M_X(t) = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \sum_{k=0}^{\infty}\binom{b-1}{k}\theta^{\{(r-k)(\alpha-1)\}} \int_0^{\infty} e^{-(\theta-t)x}(\theta + \theta^{\alpha-1}x^{\alpha-1})(1-e^{-\theta x})^k \zeta^{b-k-1}dx \quad (10)$$

Using the expansion,

$$(1-e^{-\theta x})^k = \sum_{j=0}^{\infty}(-1)^j\binom{k}{j}(e^{-\theta x})^j \quad (11)$$

implies,

$$M_X(t) = C_1\int_0^{\infty} e^{-(\theta-t)x-\theta jx}(\theta + \theta^{\alpha-1}x^{\alpha-1})\zeta^{b-k-1}dx \quad (12)$$

where,

$$C_1 = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}(-1)^j\binom{b-1}{k}\binom{k}{j}\theta^{(b-k-1)(\alpha-1)}$$

Theorem 1: The moment generating function (MGF) of the EEG when α=1 can be obtained as:

$$M_X(t) = \frac{C_2}{m\theta + (\theta(1+j) - t)}$$

where,

$$C_2 = b\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}(-1)^{m+j}\binom{b-1}{k}\binom{b-k-1}{m}\binom{k}{j}\theta^{(2-b-k)}$$

Proof: Substituting α=1 in 12, we get:

$$M_X(t) = b\theta\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}(-1)^j\binom{b-1}{k}\binom{k}{j}\int_0^{\infty} e^{-x[\theta(1+j)-t]}\left[\int_0^x e^{-\theta t\,dt}\right]^{b-k-1}dx$$

$$= b\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}(-1)^j\binom{b-1}{k}\binom{k}{j}\theta^{(2-b+k)}\int_0^{\infty} e^{-x[\theta(1+j)-t]}\underbrace{\left[1-e^{-\theta t\,dt}\right]^{b-k-1}}_{A}dx$$

By expanding the quantity A in a power series as $A = b\sum_{m=0}^{\infty}\binom{b-k-1}{m}\theta^{-m\,\theta x}$, the MGF of the EEG can be written as follows:

$$M_X(t) = C_2\int_0^{\infty} e^{-[m\theta+(\theta(1+j)-t)]x}\,dx = \frac{C_2}{m\theta + (\theta(1+j)-t)}$$

3.2. Moments

The r^{th} moment about the origin, $E(X^r)$ is defined as:

$$E(X^r) = \int_{-\infty}^{\infty} X^r f(x)\,dx$$

The r^{th} moment of the EEG model follows from Eq. 6 and can be obtained as:

$$E(X^r) = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \int_0^{\infty} x^r\,(\theta + \theta^{\alpha-1}x^{\alpha-1})e^{-\theta x}\{(1 - e^{-\theta x}) + \theta^{\alpha-1}\zeta\}^{b-1}dx \quad (13)$$

Using the generalized expansion (9), yielded:

$$E(X^r) = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \sum_{k=0}^{\infty}\binom{b-1}{k}\theta^{(\alpha-1)(b-k-1)}\int_0^{\infty} x^r e^{-\theta x}(\theta + \theta^{\alpha-1}x^{\alpha-1})(1-e^{-\theta x})^k\zeta^{b-k-1}dx \quad (14)$$

Using the expansion (11) implies:

$$E(X^r) = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}(-1)^j\binom{b-1}{k}\binom{k}{j}\theta^{(\alpha-1)(b-k-1)}\int_0^{\infty} x^r e^{-\theta x-\theta jx}(\theta + \theta^{\alpha-1}x^{\alpha-1})\zeta^{b-k-1}dx \quad (15)$$

Using the expansion $e^{-x} = \sum_{m=0}^{\infty}(-1)^m\frac{x^m}{m!}$ Implies:

$$E(X^r) = C_3\int_0^{\infty} x^{r+m}(\theta + \theta^{\alpha-1}x^{\alpha-1})\zeta^{b-k-1}dx \quad (16)$$

where,

$$C_3 = b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{j+m}}{m!}\binom{b-1}{k}\binom{k}{j}\theta^{(\alpha-1)(b-k-1)}\theta^m(1+j)^m$$

Theorem 2: The rth moment about the origin is defined as follows:

i. $If\ \ \alpha = 1;\ E(X^r) = C_4 \sum_{l=0}^{\infty}(-1)^l \binom{b-k-1}{l} \frac{\Gamma(r+m+1)}{\theta l^{(r+m+1)}}$

ii. $If\ \ \alpha = 1\ and\ r = 2;\ E(X) = C_4 \sum_{l=0}^{\infty}(-1)^l \binom{b-k-1}{l} \frac{\Gamma(m+2)}{\theta l^{(m+2)}}$

iii. $If\ \ \alpha = 1\ and\ r = 2;\ E(X^2) = C_4 \sum_{l=0}^{\infty}(-1)^l \binom{b-k-1}{l} \frac{\Gamma(m+3)}{\theta l^{(m+3)}}$

where, $C_4 = \ \theta^{k-b+2}C_3$.

Proof: Substituting α=1 in 16, gives:

$$E(X^r) = C_3 \int_0^{\infty} \theta x^{r+m} \zeta^{b-k-1}\, dx$$
$$= C_3 \int_0^{\infty} \theta x^{r+m} \left(\int_0^x e^{-\theta l}\right)^{b-k-1} dx$$
$$= C_4 \int_0^{\infty} \theta x^{r+m} \left(1 - e^{-\theta x}\right)^{b-k-1} dx$$

where, $C_4 = \theta^{k-b+2}C_3$. Expanding the quantity $\left(1 - e^{-\theta x}\right)^{b-k-1}$ in a power series as:

$$\left(1 - e^{-\theta x}\right)^{b-k-1} = \sum_{l=0}^{\infty}(-1)^l \binom{b-k-1}{l} e^{\theta l x}$$

The r^{th} moment about the origin when α=1 is written as follows:

$$E(X^r) = C_4 \int_0^{\infty} x^{r+m} \sum_{l=0}^{\infty}(-1)^l \binom{b-k-1}{l} e^{\theta l x}\, dx$$
$$= C_4 \sum_{l=0}^{\infty}(-1)^l \binom{b-k-1}{l} \frac{\Gamma(r+m+1)}{\theta l^{(r+m+1)}}$$

(ii) and (iii) are straightforward, by substituting $r = 1$ in (i).

3.3. Hazard rate function

The hazard rate function is defined as:

$$h(x) = \frac{f(x_i;\ \alpha, \theta, b)}{1 - F(x_i;\ \alpha, \theta, b)}$$

where, $f(x)$ and $F(x)$ are the pdf defined by Eq. 6 and the cdf defined by Eq. 7, respectively, hence, the hazard function for the EEG model can be derived as follows:

$$h(x) = \frac{b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b (\theta+\theta^{\alpha-1}x^{\alpha-1})e^{-\theta x}\{(1-e^{-\theta x})+\theta^{\alpha-1}\zeta\}^{b-1}}{1-[(\theta+\Gamma(\alpha))^b - \theta^b\{(1-e^{-\theta x})+\theta^{\alpha-1}\zeta\}]^b}$$
$$= \frac{b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b (\theta+\theta^{\alpha-1}x^{\alpha-1})e^{-\theta x}\{(1-e^{-\theta x})+\theta^{\alpha-1}\zeta\}^{b-1}}{\frac{(\theta+\Gamma(\alpha))^b - \theta^b\{(1-e^{-\theta x})+\theta^{\alpha-1}\zeta\}^b}{(\theta+\Gamma(\alpha))^b}} \quad (17)$$
$$= \frac{b\theta^b(\theta+\theta^{\alpha-1}x^{\alpha-1})e^{-\theta x}\{(1-e^{-\theta x})+\theta^{\alpha-1}\zeta\}^{b-1}}{(\theta+\Gamma(\alpha))^b - \theta^b\{(1-e^{-\theta x})+\theta^{\alpha-1}\zeta\}^b}$$

Fig. 2 shows some possible shapes of hazard rate function of EEG distribution for various values of the parameters α, θ and b.

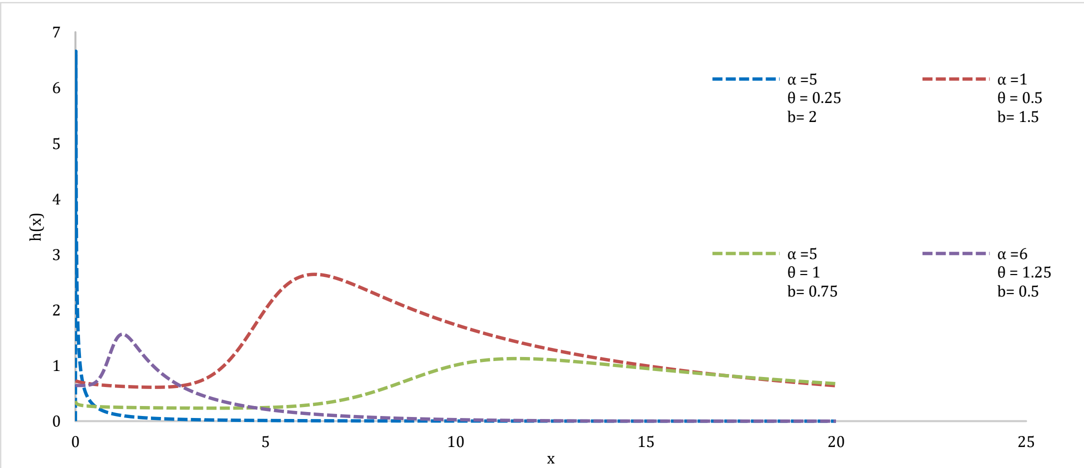

Fig. 2: EEG hazard rate function for various values of α, θ, and b

3.4. The mean residual life function

The mean residual life function is defined as:

$$m(x) = \frac{1}{1-F(x_i;\alpha,\theta,b)} \int_x^{\infty} [1 - F(x_i; \alpha, \theta, b)]dt$$

The mean residual life function of the EEG model follows from Eq. 7 and can be obtained as:

$$m(x) = \frac{1}{1-\left[\frac{\theta}{\theta+\Gamma(\alpha)}\{(1-e^{-\theta x})+\theta^{\alpha-1}\xi\}\right]^b} \int_x^{\infty} 1 - \left[\frac{\theta}{\theta+\Gamma(\alpha)}\{(1 - e^{-\theta x}) + \theta^{\alpha-1}\xi\}\right]^b dt$$
$$= \frac{1}{(\theta+\Gamma(\alpha))^b - \theta^b\{(1-e^{-\theta x})+\theta^{\alpha-1}\xi\}^b} \int_x^{\infty} (\theta + \Gamma(\alpha))^b - \theta^b\{(1 - e^{-\theta x}) + \theta^{\alpha-1}\xi\}^b dt \quad (18)$$

3.5. Maximum likelihood estimation

Let $x_1, x_2, \dots, x_n$ be a random sample from EEG distribution, then the log-likelihood function $l(x; \alpha, \theta, b)$ is defined by,

$$L(x; \alpha, \theta, b) = \prod_{i=1}^{n} f(x_i; \alpha, \theta, b)$$
$$= \prod_{i=1}^{n} b\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^b \left(\theta + \theta^{\alpha-1}x_i^{\alpha-1}\right)e^{-\theta x_i}\left[(1 - e^{-\theta x_i}) + \theta^{\alpha-1}\xi\right]^{b-1}$$
$$= \boldsymbol{b^n}\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^{nb} e^{-\theta\sum_{i=1}^{n}x_i} \prod_{i=1}^{n}\left(\theta + \theta^{\alpha-1}x_i^{\alpha-1}\right)\left[(1 - e^{-\theta x_i}) + \theta^{\alpha-1}\xi\right]^{\boldsymbol{b-1}}$$
$$= \boldsymbol{b^n}\theta^{n\alpha}\left[\frac{\theta}{\theta+\Gamma(\alpha)}\right]^{nb} e^{-\theta\sum_{i=1}^{n}x_i} \prod_{i=1}^{n}\left(\theta^{1-\alpha} + \theta^{-1}x_i^{\alpha-1}\right)\left[(1 - e^{-\theta x_i}) + \theta^{\alpha-1}\xi\right]^{\boldsymbol{b-1}} \quad (19)$$

yielding the log-likelihood,

$$l(x;\alpha,\theta,b) = nlog(b) + n\alpha\log(\theta) + nblog(\theta) - nblog(\theta+\Gamma(\alpha)) - \theta n\bar{x} + \sum_{i=1}^{n}\log(\theta+\theta^{\alpha-1}x_i^{\alpha-1}) + (b-1)\sum_{i=1}^{n}\log\left((1-e^{-\theta x_i})+\theta^{\alpha-1}\xi\right) \quad (20)$$

To solve the MLEs for each parameter, derive the derivatives of $l(x;\alpha,\theta,b)$ with respect to (w.r.t) α, θ, and b, set the partial derivatives equal to zero and solve for $\hat{\alpha}$, $\hat{\theta}$ and $\hat{b}$. The first partial derivative of Eq. 20 w.r.t α is defined as:

$$\frac{\partial l}{\partial\alpha} = nlog(\theta) - \frac{nb\Gamma(\alpha)\Gamma'(\alpha)}{\theta+\Gamma(\alpha)} + \sum_{i=1}^{n}\frac{(\theta^{\alpha-1}\log(\theta)x_i^{\alpha-1}+\theta^{\alpha-1}(x^{\alpha-1}\log(x_i))}{\theta+\theta^{\alpha-1}x_i^{\alpha-1}} + (b-1)\sum_{i=1}^{n}\frac{\theta^{\alpha-1}\log(\theta)\xi}{(1-e^{-\theta x_i})+\theta^{\alpha-1}\xi} \quad (21)$$

The first partial derivative of Eq. 20 w.r.t θ follows as:

$$\frac{\partial l}{\partial\theta} = \frac{n\alpha}{\theta} + \frac{nb}{\theta} - \frac{nb}{\theta+\Gamma(\alpha)} - n\bar{x} + \sum_{i=1}^{n}\frac{1+\theta^{\alpha-2}(\alpha-1)x_i^{\alpha-1}}{\theta+\theta^{\alpha-1}x_i^{\alpha-1}} + (b-1)\sum_{i=1}^{n}\frac{x_ie^{-\theta x_i}+\theta^{\alpha-2}(\alpha-1)\xi}{(1-e^{-\theta x_i})+\theta^{\alpha-1}\xi} \quad (22)$$

The first partial derivative of 20 w.r.t b follows as:

$$\frac{\partial l}{\partial b} = \frac{n}{b} + nlog(\theta) - nlog(\theta+\Gamma(\alpha)) + \sum_{i=1}^{n}\log\left((1-e^{-\theta x_i})+\theta^{\alpha-1}\xi\right) \quad (23)$$

It is clear that the partial derivatives of $l(x;\alpha,\theta,b)$ w.r.t the parameters α, θ, and b (21-23) have no explicit analytical solutions. Therefore, it can be solved numerically using the Newton-Raphson iterative method which is an effective method for solving nonlinear system equations. A Newton-Raphson iterative method is implemented in this paper using the function "optim" in the R package (Team, 2020). Thus, the second derivatives w.r.t α, θ, and b are needed at given as follows. The second partial derivative of 20 w.r.t α is defined as:

$$\frac{\partial^2 l}{\partial\alpha^2} = \sum_{i=1}^{n}\frac{\theta^{\alpha-1}\log(\theta)\xi}{(1-e^{-\theta x_i})+\theta^{\alpha-1}\xi} - \frac{n\Gamma(\alpha)\Gamma'(\alpha)}{(\theta+\Gamma(\alpha))} \quad (24)$$

$$\frac{\partial^2 l}{\partial\theta^2} = \frac{n}{\theta} - \frac{n}{\theta+\Gamma(\alpha)} + \sum_{i=1}^{n}\frac{x_ie^{-\theta x_i}+\theta^{\alpha-2}(\alpha-1)\xi}{(1-e^{-\theta x_i})+\theta^{\alpha-1}\xi} \quad (25)$$

$$\frac{\partial^2 l}{\partial b^2} = -\frac{n}{b^2} \quad (26)$$

3.6. Quantile function

The p^{th} quantile function $(0 < p < 1)$ is obtained by inverting the cdf (7), is given by the following relation,

$$e^{\theta x} + \gamma(\alpha,\theta x) = \sqrt[b]{p}(\theta+\Gamma(\alpha)) \quad (27)$$

4. Simulation studies

In this section, we conduct some simulation studies, in order to examine the performance of the parameters α, θ, and b of MLE (19). Random samples from EEG distribution are generated for various sample sizes n. Consider the random variable X given by the relation,

$$e^{\theta x_i} + \gamma(\alpha,\theta x_i) = \sqrt[b]{p}(\theta+\Gamma(\alpha)) \quad (28)$$

For this setting, we consider two cases as follows:

- Case 1: Assume the true parameters are $\alpha = 5$, $\theta = 0.7$, and $b = 1$ for sample sizes $n = 100, 1000, 50{,}000$, and $100{,}000$.
- Case 2: Assume the true parameters are $\alpha = 6$, $\theta = 0.6$, and $b = 0.1$ for sample sizes $n = 100, 1000, 50{,}000$, and $100{,}000$.

In both cases, the MLE $\hat{\alpha}$, $\hat{\theta}$ and $\hat{b}$ are denoted generally by $\hat{\eta}$ for all parameters. The accuracy of $\hat{\eta}$ are measured for α, θ, and b by bias and root mean square error (RMSE), defined as:

$$bias(\hat{\eta}) = E(\hat{\eta}) - \eta \quad (29)$$

$$RMSE(\hat{\eta}) = \sqrt{E(\hat{\eta}-\eta)^2} \quad (30)$$

The simulation results of case 1 and case 2 are exhibited in Tables 1 and 2, which show the estimates of each parameter ($\hat{\alpha}$, $\hat{\theta}$, $\hat{b}$). It is observed that the values of both measures (Bias and RMSE) decrease as the sample size increases.

5. Real data applications

This section demonstrates the effectiveness of the EEG distribution through three different real data sets for further examining the new model in compression to some related distributions; namely gamma, exponential, Lindley, generalized gamma (GG) (Stacy, 1962), the generalization of the generalized gamma (GGG) (Rama and Kamlesh, 2019), and the new exponential-gamma distribution (NEG) (Umar and Yahya, 2021).

More specifically, the MLE parameters values of EEG distribution are computed using R language, then compared to those values coming from the distributions mentioned previously. The Akaike Information Criterion (AIC) (Akaike, 1974), the Akaike Information Criterion corrected (AICC) (Hurvich and Tsai, 1993), and Bayesian Information Criterion (BIC) (Schwarz, 1978) are applied in order to choose the best model among these various models. The model with minimum values of these criteria is determined to be the best model.

5.1. The exceedances of the Wheaton River flood dataset

This real-life dataset was discussed in Umar and Yahya (2021), which concerns the exceedances of Wheaton River flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada, for the years 1958-1984. It is analyzed in this paper for the purpose of illustrating the effectiveness of EEG distribution compared to other related distributions mentioned previously. The observations of this dataset are reported in many papers such as Ekhosuehi and Opone (2018), Urama et al. (2021), and Ikechukwu and Eghwerido (2022), and represented in Table 3.

Table 1: Case 1: The true parameters are $\alpha = 5$, $\theta = .7$ and $b = 1$ for sample sizes $n = 100, 1000, 50{,}000$ and $100{,}000$

Sample size	Parameter	MLE	Bias	RMSE
	α	5.1076002	0.1076002	0.01076002
n = 100	θ	0.3185989	-0.3814011	0.03814011
	b	1.5677371	0.5677371	0.05677371
	α	5.0460404	0.04604045	0.001455927
n = 1000	θ	0.2279778	-0.47202217	0.014926652
	b	1.64447597	0.64447588	0.020380117
	α	5.0512606	0.05126065	0.0002292446
n = 50, 000	θ	0.2449188	-0.45508123	0.0020351851
	b	1.6332835	0.63328351	0.0028321299
	α	5.048409	0.04840879	0.000153082
n = 100, 000	θ	0.237933	-0.46206705	0.001461184
	b	1.636762	0.63676167	0.002013617

Table 2: Case 2: The true parameters are $\alpha = 6$, $\theta = 0.6$ and $b = 0.1$ for sample sizes $n = 100, 1000, 50{,}000$ and $100{,}000$

Sample size	Parameter	MLE	Bias	RMSE
	α	6.2153842	0.21538415	0.021538415
n = 100	θ	0.6324872	0.03248721	0.003248721
	b	0.2057230	0.10572304	0.010572304
	α	6.232313	0.2323128	0.0073463757
n = 1000	θ	1.017915	0.4179152	0.013215639
	b	0.205408	0.1054080	0.003333295
	α	6.2278300	0.2278300	0.0010188866
n = 50, 000	θ	0.9289968	0.3289968	0.0014713185
	b	0.2054345	0.1054345	0.000471517
	α	6.2260943	0.2260943	0.00071497302
n = 100, 000	θ	0.9234960	0.3234960	0.0010229841
	b	0.2054497	0.1054497	0.0003334613

The parameter estimates are reported in Table 4 for the new proposed distribution and some related distributions. The Akaike information criterion (AIC), the Akaike Information Criterion corrected (AICC), and Bayesian Information Criterion (BIC) are used to assess the best model. The results are reported in Table 5. According to the values of AIC, AICC, and BIC, EGG performs well and exceeded the other models.

Table 3: The exceedances of Wheaton River flood data

1.7	2.2	14.4	1.1	0.4	20.6	5.3	0.7	1.9	13.0
12.0	9.3	1.4	18.7	8.5	25.5	11.6	14.1	22.1	1.1
2.5	14.4	1.7	37.6	0.6	2.2	39.0	0.3	15.0	11.0
7.3	22.9	1.7	0.1	1.1	0.6	9.07	1.7	7.0	20.1
0.4	2.8	14.1	9.9	10.4	10.7	30.0	3.6	5.6	30.8
13.3	4.2	25.5	3.4	11.9	21.5	27.6	36.4	2.7	64.0
1.5	2.5	27.4	1.0	27.1	20.2	16.8	5.3	9.7	27.5
2.5	27.0								

This dataset can be read from left to right

Table 4: Parameters estimation of the EEG and the six other existing distributions for the exceedances of Wheaton River flood data

Distribution	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{b}$
EXP.	…	0.0819	…	…	…
GAMMA	0.8383	0.0687	…	…	…
GG	0.4768	0.0085	1.4921	…	…
GGG	0.4768	0.0250	1.4919	0.4859	…
LINDLEY	…	0.0819	…	…	…
NEG	1.4616	0.0755	…	…	…
EEG	1.3289	0.01119	…	…	0.1727

Table 5: Goodness-of-Fit test results of the EEG and the six other existing distributions for the exceedances of Wheaton River flood data

Distribution	-2logLik	AIC	AICC	BIC
EXP.	504.256	506.256	506.313	508.533
GAMMA	502.689	506.689	506.863	506.965
GG	502.131	508.131	508.479	506.408
GGG	502.131	510.131	510.711	506.408
LINDLEY	830.051	832.051	832.108	834.327
NEG	499.012	503.012	503.186	503.289
EEG	360.975	366.975	367.328	373.805

5.2. The remission time of the bladder cancer patients dataset

This dataset was collected from bladder cancer patients and reported by Lee and Wang (2003). It has been studied in detail by Ieren et al. (2019), Ijaz et al. (2020), and Kayid (2022). The real-life data contains a set of remission times (in months) from 128 bladder cancer patients which is represented in Table 6. It is used here to examine the new proposed model in comparison to some related models. The parameter estimates are reported in Table 7 for our distribution and the other related distributions. The results of the goodness of fit test are based on the AIC, AICC, and BIC Criteria as illustrated in Table 8. It is clear that EEG distribution with minimum results of all three criteria in compression to the six distributions.

Table 6: The remission times (in months) from 128 bladder cancer patients' data

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.20	2.23
3.52	4.98	6.97	9.02	13.29	0.40	2.26	3.57	5.06	7.09
9.22	3.80	25.74	0.50	2.46	3.64	5.09	7.26	9.47	14.24
25.82	0.51	2.54	3.70	5.17	7.28	9.74	14.76	26.31	0.81
2.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64	3.88	5.32
7.39	10.34	14.83	34.26	0.90	2.69	4.18	5.34	7.59	10.66
15.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01
1.19	2.75	4.26	5.41	7.63	17.12	46.12	1.2	2.83	4.33
5.49	7.66	11.25	17.14	79.05	1.35	2.87	5.62	7.87	11.64
17.36	1.40	3.02	4.34	5.71	7.93	1.46	18.10	11.79	4.40
5.85	8.26	11.98	19.13	1.76	3.25	4.50	6.25	8.37	12.02
2.02	13.31	4.51	6.54	8.53	12.03	20.28	2.02	3.36	12.07
6.76	21.73	2.07	3.36	6.93	8.65	12.63	22.69.		

This dataset can be read from left to right

Table 7: Parameters estimation of the EEG and the six other existing distributions for the bladder cancer data

Distribution	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{b}$
EXP.	...	0.1068	...	...	...
GAMMA	1.1723	7.9914	...	...	...
GG	3.7101	1.2952	0.5211	...	...
GGG	4.9913	2.9708	0.4451	0.4402	...
LINDLEY	...	0.1960	...	...	...
NEG	1.1929	0.1247	...	...	...
EEG	1.3201	0.0223	...	...	0.2144

Table 8: Goodness-of-fit test results of the EEG and the six other existing distributions for the bladder cancer data

Distribution	-2logLik	AIC	AICC	BIC
EXP.	828.680	830.680	830.728	833.532
GAMMA	826.700	830.700	830.796	836.404
GG	821.720	827.720	827.914	836.276
GGG	821.860	829.860	830.185	841.268
LINDLEY	838.940	840.940	840.9717	843.792
NEG	826.800	830.800	830.896	836.504
EEG	622.060	628.060	628.254	636.616

5.3. The COVID-19 dataset

This dataset is about COVID-19 data from 15 April to 30 June 2020 in The United Kingdom for 76 days. These numbers indicate the death rate due to drought and have been used by Mubarak and Almetwally (2021). The data are represented in Table 9. Table 10 shows the estimates of the parameters of the EEG distribution and the other related distributions. By looking at Table 11, it is observed that the EEG distribution performs better since it has minimum AIC, BIC, and AICC compared to the other distributions.

Table 9: COVID-19 data from 15 April to 30 June 2020 in The United Kingdom of 76 days

0.0587	0.0863	0.1165	0.1247	0.1277	0.1303	0.1652
0.2079	0.2395	0.2751	0.2845	0.2992	0.3188	0.3317
0.3446	0.3553	0.3622	0.3926	0.3926	0.4110	0.4633
0.4690	0.4954	0.5139	0.5696	0.5837	0.6197	0.6365
0.7096	0.7193	0.7444	0.85907	1.0438	1.0602	1.1305
1.1468	1.1533	1.2260	1.2707	1.3423	1.4149	1.5709
1.6017	1.6083	1.6324	1.6998	1.8164	1.8392	1.8721
1.9844	2.1360	2.3987	2.4153	2.5225	2.7087	2.7946
3.3609	3.3715	3.7840	3.9042	4.1969	4.3451	4.4627
4.6477	5.3664	5.4500	5.7522	6.4241	7.0657	7.4456
8.2307	9.6315	10.1870	11.1429	11.2019	11.4584.	

This dataset can be read from left to right

Table 10: Parameters estimation of the EEG and the six other existing distributions for the COVID-19 data

Distribution	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{b}$
EXP.	...	0.4103	...	...	...
GAMMA	0.8016	3.0393	...	...	...
GG	6.0458	5.3023	0.3169	...	...
GGG	5.6524	2.7935	0.3289	5.5294	...
LINDLEY	...	0.6578	...	...	...
NEG	0.7844	0.3402	...	...	...
EEG	0.6433	0.2156	...	...	0.6409

6. Conclusion

This paper introduces and establishes a new distribution called the exponentiated new exponential-gamma distribution, which serves as a generalization of the new exponential-gamma distribution. We provide expansions for several statistical properties of this newly proposed distribution, including the moment generating function, r^{th} moment, hazard rate function, and mean residual life function. To estimate the numerical values of the parameters of the EEG distribution, we employ the maximum likelihood estimation method.

Table 11: Goodness-of-fit test results of the EEG and the six other existing distributions for the COVID-19 data

Distribution	-2logLik	AIC	AICC	BIC
EXP.	287.400	289.409	289.454	291.731
GAMMA	284.820	288.820	288.986	293.482
GG	279.640	285.635	285.968	292.632
GGG	279.680	287.680	288.243	297.003
LINDLEY	301.898	303.898	303.979	306.229
NEG	285.100	289.092	289.256	293.761
EEG	272.380	278.380	278.713	285.372

To evaluate the performance of the new model, we conduct simulation studies with various sample sizes. Additionally, we apply the EEG distribution to real-life datasets. The results of these applications demonstrate that the EEG distribution yields superior fits compared to the exponential, gamma, Lindley, generalization of the generalized gamma distribution, and new exponential-gamma distribution. This superiority of the EEG distribution over the other six distributions confirms its effectiveness and robustness.

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

Abdullahi I and Phaphan W (2022). Some properties of the new mixture of generalized gamma distribution. Lobachevskii Journal of Mathematics, 43(9): 2349-2359. https://doi.org/10.1134/S1995080222120022

Akaike H (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6): 716–723. https://doi.org/10.1109/TAC.1974.1100705

Dahiya RC and Gurland J (1972). Goodness of fit tests for the gamma and exponential distributions. Technometrics, 14(3): 791-801. https://doi.org/10.1080/00401706.1972.10488967

DiDonato AR and Morris Jr AH (1986). Computation of the incomplete gamma function ratios and their inverse. ACM Transactions on Mathematical Software, 12(4): 377-393. https://doi.org/10.1145/22721.23109

Eghwerido JT, Nzei LC, and Zelibe SC (2022). The alpha power extended generalized exponential distribution. Journal of Statistics and Management Systems, 25(1): 187-210. https://doi.org/10.1080/09720510.2021.1872692

Ekhosuehi N and Opone F (2018). A three parameter generalized Lindley distribution: Properties and application. Statistica, 78(3): 233-249.

Ghitany ME (1998). On a recent generalization of gamma distribution. Communications in Statistics-Theory and Methods, 27(1): 223-233. https://doi.org/10.1080/03610929808832662

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

Gupta RD and Kundu D (2001). Generalized exponential distribution: Different method of estimations. Journal of Statistical Computation and Simulation, 69(4): 315-337. https://doi.org/10.1080/00949650108812098

Hurvich CM and Tsai CL (1993). A corrected Akaike information criterion for vector autoregressive model selection. Journal of Time Series Analysis, 14(3): 271-279. https://doi.org/10.1111/j.1467-9892.1993.tb00144.x

Ieren TG, Kromtit FM, Agbor BU, Eraikhuemen IB, and Koleoso PO (2019). A power Gompertz distribution: Model, properties and application to bladder cancer data. Asian Research Journal of Mathematics, 15(2): 1-14. https://doi.org/10.9734/arjom/2019/v15i230146

Ijaz M, Asim SM, Farooq M, Khan SA, and Manzoor S (2020). A gull alpha power Weibull distribution with applications to real and simulated data. PLOS ONE, 15(6): e0233080. https://doi.org/10.1371/journal.pone.0233080 **PMid:32530965 PMCid:PMC7292407**

Ikechukwu AF and Eghwerido JT (2022). Transmuted shifted exponential distribution and applications. Journal of Statistics and Management Systems, 25(4): 857-878. https://doi.org/10.1080/09720510.2021.1958517

Kayid M (2022). Applications of bladder cancer data using a modified log-logistic model. Applied Bionics and Biomechanics, 2022: 6600278. https://doi.org/10.1155/2022/6600278 **PMid:35154377 PMCid:PMC8828332**

Khan MS, King R, and Hudson IL (2017). Transmuted generalized exponential distribution: A generalization of the exponential distribution with applications to survival data. Communications in Statistics-Simulation and Computation, 46(6): 4377-4398. https://doi.org/10.1080/03610918.2015.1118503

Lee ET and Wang J (2003). Statistical methods for survival data analysis. Volume 476, John Wiley and Sons, Hoboken, USA. https://doi.org/10.1002/0471458546

Mubarak AES and Almetwally EM (2021). A new extension exponential distribution with applications of COVID-19 data. Financial and Business Research Journal, 22(1): 444-460. https://doi.org/10.21608/jsst.2021.51484.1178

Nadarajah S and Haghighi F (2011). An extension of the exponential distribution. Statistics, 45(6): 543-558. https://doi.org/10.1080/02331881003678678

Rama S and Kamlesh KS (2019). A generalization of generalized gamma distribution. International Journal of Computational and Theoretical Statistics, 6(1): 33-42. https://doi.org/10.12785/ijcts/060105

Schwarz G (1978). Estimating the dimension of a model. The Annals of Statistics, 6(2): 461-464. https://doi.org/10.1214/aos/1176344136

Stacy EW (1962). A generalization of the gamma distribution. The Annals of Mathematical Statistics, 33(3): 1187-1192. https://doi.org/10.1214/aoms/1177704481

Susanto I, Iriawan N, and Kuswanto H (2022). On the Bayesian mixture of generalized linear models with gamma-distributed responses. Econometrics, 10(4): 32. https://doi.org/10.3390/econometrics10040032

Team RC (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Thom HC (1958). A note on the gamma distribution. Monthly Weather Review, 86(4): 117-122. https://doi.org/10.1175/1520-0493(1958)086<0117:ANOTGD>2.0.CO;2

Umar MA and Yahya WB (2021). A new poison-exponential-gamma distribution for modelling count data with applications. https://doi.org/10.2139/ssrn.3971601

Urama KU, Onyeagu SI, and Eze FC (2021). The transmuted Kumaraswamy pareto distribution. Earthline Journal of Mathematical Sciences, 6(2): 325-358. https://doi.org/10.34198/ejms.6221.325358

Yakubu OM, Mohammed YA, and Imam A (2022). A mixture of gamma-gamma, loglogistic-gamma distributions for the analysis of heterogenous survival data. International Journal of Mathematical Research, 11(1): 1-9. https://doi.org/10.18488/24.v11i1.2924