International Journal of Statistics and Probability; Vol. 13, No. 1; March 2024
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

X-exponential-G Family of Distributions With Applications

Shahid Mohammad

Correspondence: Department of Mathematics, University of Wisconsin Oshkosh, USA

Received: January 10, 2024 Accepted: February 21, 2024 Online Published: February 29, 2024

doi:10.5539/ijsp.v13n1p40 URL: https://doi.org/10.5539/ijsp.v13n1p40

Abstract

A new family of continuous distributions called the X-exponential-G (XE-G) family is proposed. Explicit expressions are derived for the ordinary and incomplete moments, generating functions, mean deviation about the mean and median, Shannon and Rényi entropies, and order statistics of this new family. Estimation of the parameters of the new family is done using the method of maximum likelihood. Assessment of the performance of the maximum likelihood estimates is carried out through a simulation study using the quantile function of the XE-G distribution. The usefulness of this new family is illustrated by modeling two real datasets.

Keywords: exponential distribution, goodness-of-fit, maximum likelihood estimation, moments, simulation

1. Introduction

Researchers across various fields including technology, healthcare, and finance have constructed numerous statistical distributions over time to describe processes and events over a lifetime. However, as real systems change, classical distributions often fall short in capturing emerging patterns in current data. This need to enhance adaptability in modeling has motivated continuing efforts to adjust well-known distributions or devise completely new ones. Common techniques involve adding extra shape parameters, combining existing distributions, using transformations, and creating multivariate forms. In general, developing and analyzing extended distributions remains an energetic research focus, as time-tested models keep gaining from increased flexibility.

Researchers have constructed new families of statistical distributions by creatively combining properties of existing distributions. Some examples include the Weibull-X (Alzaatreh et al., 2015), Weibull-G (Bourguignon et al., 2014), logistic-G (Torabi et al., 2014), Lomax-G (Cordeiro et al., 2014), logistic-X (Tahir et al., 2016), a new Weibull-G (Tahir et al., 2016), Kumaraswamy-G (Nadarajah et al., 2012), and Gompertz-G (Alizadeh et al., 2017). The goal of creating these new flexible families is to provide statisticians and data scientists with more customized options when selecting an appropriate distribution to model real-world data. By adding new parameters, the resulting families can more accurately fit data than previously exiting distributions. Additional distributions includes bivariate Gumbel-G (Eliwa et al., 2019), shifted Gompertz-G (Eghwerido et al., 2021), transmuted Weibull-G (Alizadeh et al., 2018), beta-generated (beta-G) (Eugene et al., 2002), generalized beta-G (Alexander et al., 2012), gamma-G type 1 (Zografos and Balakrishnan, 2014), gamma-G type 2 (Ristić and Balakrishnan, 2012), and odd-gamma-G type 3 (Torabi and Hedesh, 2012), Kumaraswamy Marshal-Olkin (Alizadeh et al., 2015), alpha power Marshall-Olkin-G (Eghwerido et al., 2021), Kumaraswamy odd log-logistic-G (Alizadeh et al., 2015), transmuted alpha power-G families (Eghwerido et al., 2021), Topp Leone odd Lindley-G (Reyad et al., 2018), and inverse odd Weibull generated (Eghwerido et al., 2020) among various others in the statistical domain.

Within the domain of exponentiated distributions, Gupta and Kandu (2001) is credited with introducing the exponentiated exponential distribution, which has since generated considerable interest and found widespread use across a spectrum of diverse applications. Furthermore, among the notable exponential distributions are the alpha power transformed generalized exponential (Dey et al., 2017), Lindley exponential (Bhati et al., 2015), alpha power Weibull (Nassar et al., 2017), exponentiated Weibull-H (Cordeiro et al., 2017), power Lindley (Ghitany et al., 2013), transmuted exponentiated generalized-G (Yousof et al., 2015), exponentiated transmuted-G (Merovci et al., 2017), exponentiated generalized (Cordeiro et al., 2013), exponentiated half-logistic (Cordeiro et al., 2014), odd generalized exponential (Tahir et al., 2015), and shifted exponential-G (Eghwerido et al., 2022) distributions. Additionally, the category of distributions derived through exponentiation includes the exponentiated gamma, exponentiated Weibull, exponentiated Gumbel, and exponentiated Fréchet distributions, as documented in (Nadarajah and Kotz, 2006).

The paper is organized in the following ways. In Section 2, the X-exponential-G (XE-G) family of distributions is defined, including particular models found within the family. Section 3 discusses two special types of XE-G family of distributions. Mathematical properties, such as moments, moment-generating functions, mean deviation about the mean and median, and entropies, are derived in Section 4. Order statistics and estimation are discussed in Section 5 and Section 6, respectively. Section 7 allows for investigating practical applications rooted in real-world data. Finally, Section 8 provides our concluding remarks on the research presented.

2. The New Family

Let X be a continuous random variable. The cumulative distribution function (cdf) of the exponential distribution function is $1 - e^{-\lambda x}$ for $\lambda, x > 0$. The X-exponential-G (XE-G) family is defined by replacing the parameter λ by $G(x; \boldsymbol{\xi})$, where $G(x; \boldsymbol{\xi})$ represents the cdf of a baseline distribution, and $\boldsymbol{\xi}$ is a vector of baseline parameters. Hence, the cdf of the XE-G distribution is given by

$$F(x) = 1 - e^{-xG(x;\boldsymbol{\xi})}, x > 0. \tag{1}$$

The probability density function (pdf), and hazard rate function corresponding to equation (1) becomes

$$f(x) = \left[xg(x; \boldsymbol{\xi}) + G(x; \boldsymbol{\xi})\right] e^{-xG(x;\boldsymbol{\xi})}, \tag{2}$$

and

$$h\left(x\right) = xg(x; \boldsymbol{\xi}) + G(x; \boldsymbol{\xi}), \tag{3}$$

where $g\left(x; \boldsymbol{\xi}\right)$ represents the pdf of the baseline model.

The XE-G distribution does not introduce any additional parameters beyond the baseline distribution $G(x; \boldsymbol{\xi})$. A random variable X following this distribution is denoted as $X \sim$XE-G.

Table 1 includes special models found within the family (1).

Table 1. Distributions and corresponding $G(x; \boldsymbol{\xi})$ functions

Distribution	$G(x; \boldsymbol{\xi})$	$\boldsymbol{\xi}$
Exponential	λ	λ
Weibull	$\frac{1}{x}\left(\frac{x}{\theta}\right)^{\tau}$	(τ, θ)
Rayleigh	$\frac{x}{2\sigma^2}$	σ
Gompertz	$\frac{\eta}{x}\left(e^{bx} - 1\right)$	(η, b)
Lindley	$\ln\left(\frac{\theta+1+\theta x}{\theta+1}\right)^{-\frac{1}{x}}$	θ
Generalized exponential	$\ln\left(1 + x\right)^{\frac{1}{x}}$	$\emptyset$
Nadarajah–Haghighi	$\frac{1}{x}\left[(1 + \gamma x)^{\theta} - 1\right]$	(γ, θ)

3. Special XE Distributions

Here, only two of the many distributions that can arise as XE special models are introduced. Two baseline distributions, namely the inverse paralogistic and Gumbel distributions, are considered, although as many new distributions as desired can be generated.

3.1 The XE-Inverse Paralogistic (XE-IP) Distribution

The pdf and cdf of the inverse paralogistic distribution are, respectively, given by

$$g(x) = \frac{\tau^2 \left(\frac{x}{\theta}\right)^{\tau^2}}{x\left\{1+\left(\frac{x}{\theta}\right)^{\tau}\right\}^{\tau+1}} \quad \text{and} \quad G(x) = \left\{\frac{\left(\frac{x}{\theta}\right)^{\tau}}{1+\left(\frac{x}{\theta}\right)^{\tau}}\right\}^{\tau} \quad , x, \tau, \theta > 0.$$

Then, the pdf, cdf, and hazard rate function of the XE-IP distribution are, respectively, given by

$$f(x) = \left[\frac{\tau^2 \left(\frac{x}{\theta}\right)^{\tau^2}}{\left\{1+\left(\frac{x}{\theta}\right)^{\tau}\right\}^{\tau+1}} + \left\{\frac{\left(\frac{x}{\theta}\right)^{\tau}}{1+\left(\frac{x}{\theta}\right)^{\tau}}\right\}^{\tau}\right] \exp\left[-x\left\{\frac{\left(\frac{x}{\theta}\right)^{\tau}}{1+\left(\frac{x}{\theta}\right)^{\tau}}\right\}^{\tau}\right] \quad , x, \tau, \theta > 0, \tag{4}$$

$$F(x) = 1 - \exp\left[-x\left\{\frac{\left(\frac{x}{\theta}\right)^{\tau}}{1+\left(\frac{x}{\theta}\right)^{\tau}}\right\}^{\tau}\right], \tag{5}$$

and

$$h(x) = \frac{\tau^2 \left(\frac{x}{\theta}\right)^{\tau^2}}{\left\{1+\left(\frac{x}{\theta}\right)^{\tau}\right\}^{\tau+1}} + \left\{\frac{\left(\frac{x}{\theta}\right)^{\tau}}{1+\left(\frac{x}{\theta}\right)^{\tau}}\right\}^{\tau}.$$

Figure 1 illustrates the density and hazard rate function graphs of the XE-IP distribution across different parameter values, presenting a range of patterns that involve increasing or decreasing pattern.

3.2 The XE-Gumbel (XE-Gu) Distribution

The pdf and cdf of the Gumbel distribution are, respectively, given by

$$g(x) = \frac{1}{\sigma}\, e^{-\left[\frac{x-\mu}{\sigma} + e^{-\frac{x-\mu}{\sigma}}\right]} \quad \text{and} \quad G(x) = e^{-e^{-\frac{x-\mu}{\sigma}}} \quad , x \in \mathbb{R}, \mu, \sigma > 0.$$

Then, the pdf, cdf, and hazard rate function of the XE-Gu distribution are, respectively, given by

$$f(x) = \left[1 + \frac{x}{\sigma}\, e^{-\frac{x-\mu}{\sigma}}\right] e^{-\left[e^{-\frac{x-\mu}{\sigma}} + x\, e^{-e^{-\frac{x-\mu}{\sigma}}}\right]} \quad , x, \mu, \sigma > 0, \tag{6}$$

$$F(x) = 1 - \exp\left[-x\, e^{-e^{-\frac{x-\mu}{\sigma}}}\right], \tag{7}$$

and

$$h(x) = \frac{x}{\sigma}\, e^{-\left[\frac{x-\mu}{\sigma} + e^{-\frac{x-\mu}{\sigma}}\right]} + e^{-e^{-\frac{x-\mu}{\sigma}}}.$$

Figure 2 portrays the density and the hazard rate function graphs under various parameter configurations, providing insight into the diverse patterns and trends linked to the XE-Gu distribution.

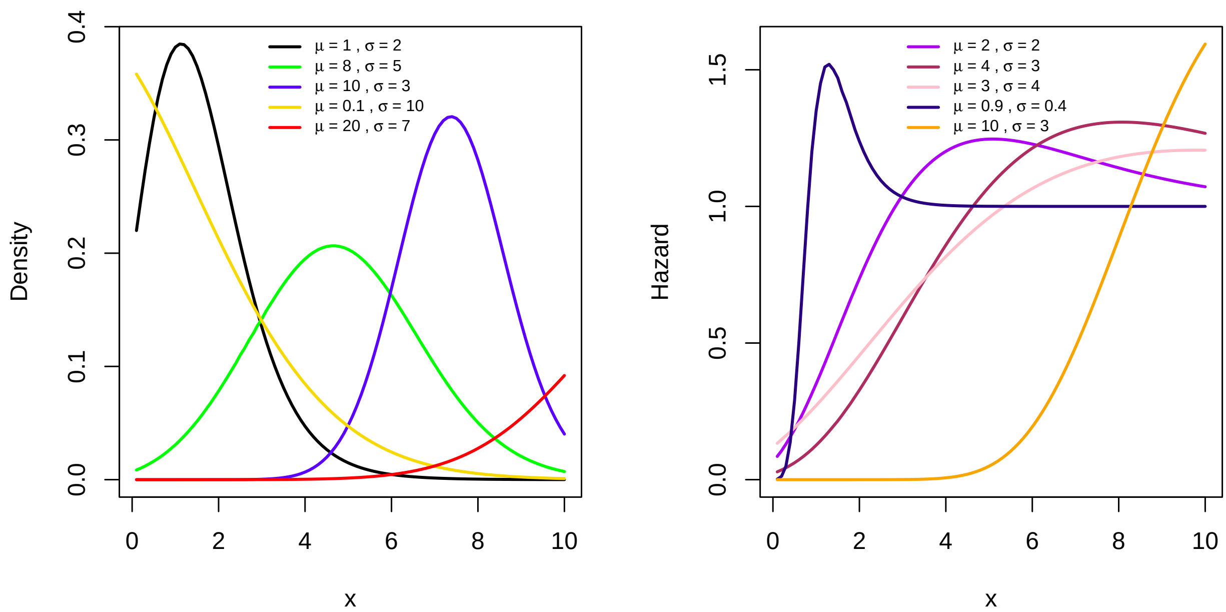

Figure 2. Plots of the XE-Gu densities and hazard rate for some parameter values

4. Mathematical Properties

The properties of the XE-G family of distributions, such as moments, moment generating, mean deviation about the mean and median, as well as entropies, are being derived.

4.1 Moments

Moments are crucial in statistics for understanding distributions and their key features, offering valuable insights in practical applications. For the XE-G distribution, the r^{th} ordinary moment of the random variable X, denoted as $\mu_r^{'}$, can be expressed mathematically as

$$\mu_r^{'} = \int_0^\infty x^r \left[xg(x; \boldsymbol{\xi}) + G(x; \boldsymbol{\xi}) \right] e^{-xG(x;\boldsymbol{\xi})} \, dx. \tag{8}$$

If $Q_G(u)$ represents the quantile function of the baseline distribution $G(x; \boldsymbol{\xi})$, then (8) can be further simplified as

$$\mu_r^{'} = \int_0^1 \left[Q_G(u) \right]^{r+1} e^{-uQ_G(u)} du + \int_0^\infty x^r G(x; \boldsymbol{\xi}) e^{-xG(x;\boldsymbol{\xi})} \, dx. \tag{9}$$

The integral presented above has the potential to be calculated numerically in situations where it cannot be expressed in closed form.

When r is an integer, the expression for the r^{th} central moments of X is as follows

$$\mu_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k \left(\mu_1^{'} \right)^k \mu_{r-k}^{'}, \tag{10}$$

and the r^{th} cumulants, denoted as κ_r of X, can be determined using the equation

$$\kappa_r = \mu_r^{'} - \sum_{k=1}^{r-1} \binom{r-1}{k-1} \kappa_k \, \mu_{r-k}^{'} \quad \text{with} \quad \kappa_1 = \mu_1^{'}. \tag{11}$$

4.2 Incomplete Moments

Incomplete moments are calculated over a specified range rather than over the entire distribution. These partial moments are helpful for curves like the Lorenz curve and the Bonferroni curve.

The r^{th} incomplete moment of XE-G distribution, denoted as $m_r(t)$, can be written as

$$m_r(t) = \int_0^t x^r \left[xg(x;\boldsymbol{\xi}) + G(x;\boldsymbol{\xi}) \right] e^{-xG(x;\boldsymbol{\xi})} \, dx. \tag{12}$$

If $Q_G(u)$ represents the quantile function of the baseline distribution $G(x;\boldsymbol{\xi})$, then (12) can be further simplified as

$$m_r(t) = \int_0^{G(t;\boldsymbol{\xi})} \left[Q_G(u) \right]^{r+1} e^{-uQ_G(u)} du + \int_0^t x^r G(x;\boldsymbol{\xi}) e^{-xG(x;\boldsymbol{\xi})} \, dx. \tag{13}$$

Again, the integral presented above has the potential to be calculated numerically in situations where it cannot be expressed in closed form.

4.3 Moment Generating Function

A moment generating function is a vital tool for finding the moments of a probability distribution. The moment generating function, $M_X(t)$, of the XE-G distribution can be written in the sum form as:

$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mu_n^{'}, \tag{14}$$

where $\mu_n^{'}$ represents the n^{th} moment of the NE-G distribution, which can be obtained from either (8) or (9).

4.4 Mean Deviation About Mean and Median

Mean deviations about the mean and median provide valuable insights into the distribution's variability around the central measures. The mean deviation about the mean (δ_1) and about the median (δ_2) are , respectively, given by

$$\delta_1(X) = 2\mu_1^{'} F(\mu_1^{'}) - 2m_1(\mu_1^{'}) \qquad \text{and} \qquad \delta_2(X) = \mu_1^{'} - 2m_1(M),$$

where $\mu_1^{'}$ represents the mean and M denotes the median.

4.5 Shannon and Rényi Entropy

Shannon entropy and Rényi entropy are used to quantify the uncertainty or randomness in probability distributions. Let $X \sim$XE-G. The Shannon entropy of the random variable X, denoted as $H(X)$, is commonly defined as

$$H(x) = -\int_0^{\infty} f(x) \ln(f(x)) \, dx,$$

where $f(x)$ is the probability density function of the random variable X. For the XE-G distribution, this Shannon entropy can be expressed as

$$\begin{aligned} H(x) = &-\int_0^{\infty} (xg(x;\boldsymbol{\xi}) + G(x;\boldsymbol{\xi})) \ln \{(xg(x;\boldsymbol{\xi}) + G(x;\boldsymbol{\xi}))\} e^{-xG(x;\boldsymbol{\xi})} dx \\ &+ \int_0^{\infty} xG(x;\boldsymbol{\xi}) (xg(x;\boldsymbol{\xi}) + G(x;\boldsymbol{\xi})) e^{-xG(x;\boldsymbol{\xi})} dx \end{aligned}$$

By evaluating the second integral in the equation above, we obtain

$$\int_0^\infty xG(x;\boldsymbol{\xi})\,(xg(x;\boldsymbol{\xi})+G(x;\boldsymbol{\xi}))\,e^{-xG(x;\boldsymbol{\xi})}dx = 1.$$

Therefore, the Shannon entropy of the XE-G distribution simplifies to

$$H(x) = 1 - \int_0^\infty (xg(x;\boldsymbol{\xi})+G(x;\boldsymbol{\xi}))\ln\{(xg(x;\boldsymbol{\xi})+G(x;\boldsymbol{\xi}))\}\,e^{-xG(x;\boldsymbol{\xi})}dx \tag{15}$$

The Rényi entropy for a parameter $\nu \neq 1, \nu > 0$ is defined

$$I_R(\nu) = \frac{1}{1-\nu}\ln\int_0^\infty f^\nu(x)dx$$

For the XE-G distribution, this can be expressed as

$$I_R(\nu) = \frac{1}{1-\nu}\ln\int_0^\infty \left[\{xg(x;\boldsymbol{\xi})+G(x;\boldsymbol{\xi})\}^\nu\,e^{-\nu xG(x;\boldsymbol{\xi})}\right]dx \tag{16}$$

Once again, these integrals mentioned above can be evaluated numerically to determine the values of both the Shannon and Rényi entropy for the XE-G distribution.

5. Order Statistics

Let us consider $x_1, x_2, ..., x_n$ be an observed sample of size n drawn from a XE-G distribution with cdf $F(x)$, and pdf $f(x)$. Now let $x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n}$ are the corresponding order statistics of samples. Then the i^{th} order statistic for XE-G distribution is given as

$$f_i(x) = \frac{n!}{(i-1)!\,(n-i)!}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}f(x)\,[F(x)]^{i+j-1}. \tag{17}$$

By substituting the equations (1) and (2) into equation (17), we obtain

$$f_i(x) = \frac{n!}{(i-1)!\,(n-i)!}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}\left[xg(x;\boldsymbol{\xi})+G(x;\boldsymbol{\xi})\right]e^{-xG(x;\boldsymbol{\xi})}\left[1-e^{-xG(x;\boldsymbol{\xi})}\right]^{i+j-1}. \tag{18}$$

Using the binomial expansion, we have

$$\left[1-e^{-xG(x;\boldsymbol{\xi})}\right]^{i+j-1} = \sum_{l=0}^{i+j-1}(-1)^l\binom{i+j-1}{l}e^{-lxG(x;\boldsymbol{\xi})}.$$

Hence equation (18) can be written as

$$f_i(x) = \frac{n!}{(i-1)!\,(n-i)!}\sum_{j=0}^{n-i}\sum_{l=0}^{i+j-1}(-1)^{j+l}\binom{n-i}{j}\binom{i+j-1}{l}\left[xg(x;\boldsymbol{\xi})+G(x;\boldsymbol{\xi})\right]e^{-(l+1)xG(x;\boldsymbol{\xi})}.$$

This expression provides a formula for calculating the pdf of the i^{th} order statistic for the XE-G distribution.

6. Estimation

Consider a random sample $x_1, x_2, \ldots, x_n$ from a XE-G distribution with unknown parameter vector ξ. The log-likelihood function is given by

$$l(\xi) = -\sum_{i=1}^{n} x_i G(x_i; \boldsymbol{\xi}) + \sum_{i=1}^{n} \ln\left[x_i g(x_i; \boldsymbol{\xi}) + G(x_i; \boldsymbol{\xi})\right].$$

To obtain the maximum likelihood estimates (MLEs) of ξ, differentiate $l(\xi)$ with respect to each parameter ξ_j and set the result equal to zero:

$$\frac{\partial l(\xi)}{\partial \xi_j} = -\sum_{i=1}^{n} x_i \frac{\partial G(x_i; \boldsymbol{\xi})}{\partial \xi_j} + \sum_{i=1}^{n} \frac{x_i \frac{\partial g(x_i;\boldsymbol{\xi})}{\partial \xi_j} + \frac{\partial G(x_i;\boldsymbol{\xi})}{\partial \xi_j}}{x_i g(x_i; \boldsymbol{\xi}) + G(x_i; \boldsymbol{\xi})} = 0.$$

This provides a set of nonlinear equations that can be solved numerically for the MLEs of each parameter ξ_j.

6.1 Simulation Study

It is necessary to check the performance of the maximum likelihood estimators (MLEs) of the XE-IP distribution. For this purpose, we generate $N = 1000$ samples of sizes $n = 25, 50, \ldots, 1000$ using the quantile function of the XE-IP distribution. The biases and MSEs are computed using the formulas

$$\text{Bias}(\boldsymbol{\Theta}) = \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}\right) \quad \text{and} \quad \text{MSE}(\boldsymbol{\Theta}) = \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}\right)^2,$$

for $\boldsymbol{\Theta} = (\tau, \theta)$, respectively.

The graphical representations in Figures 3-5 reveal essential insights from our simulation results. Firstly, the estimates are stable as the sample size (n) increases. The MSEs consistently decrease toward zero with larger sample sizes, while the absolute and estimated biases decrease. This observation confirms the performance and consistency of the MLEs for the parameters of the XE-IP distribution.

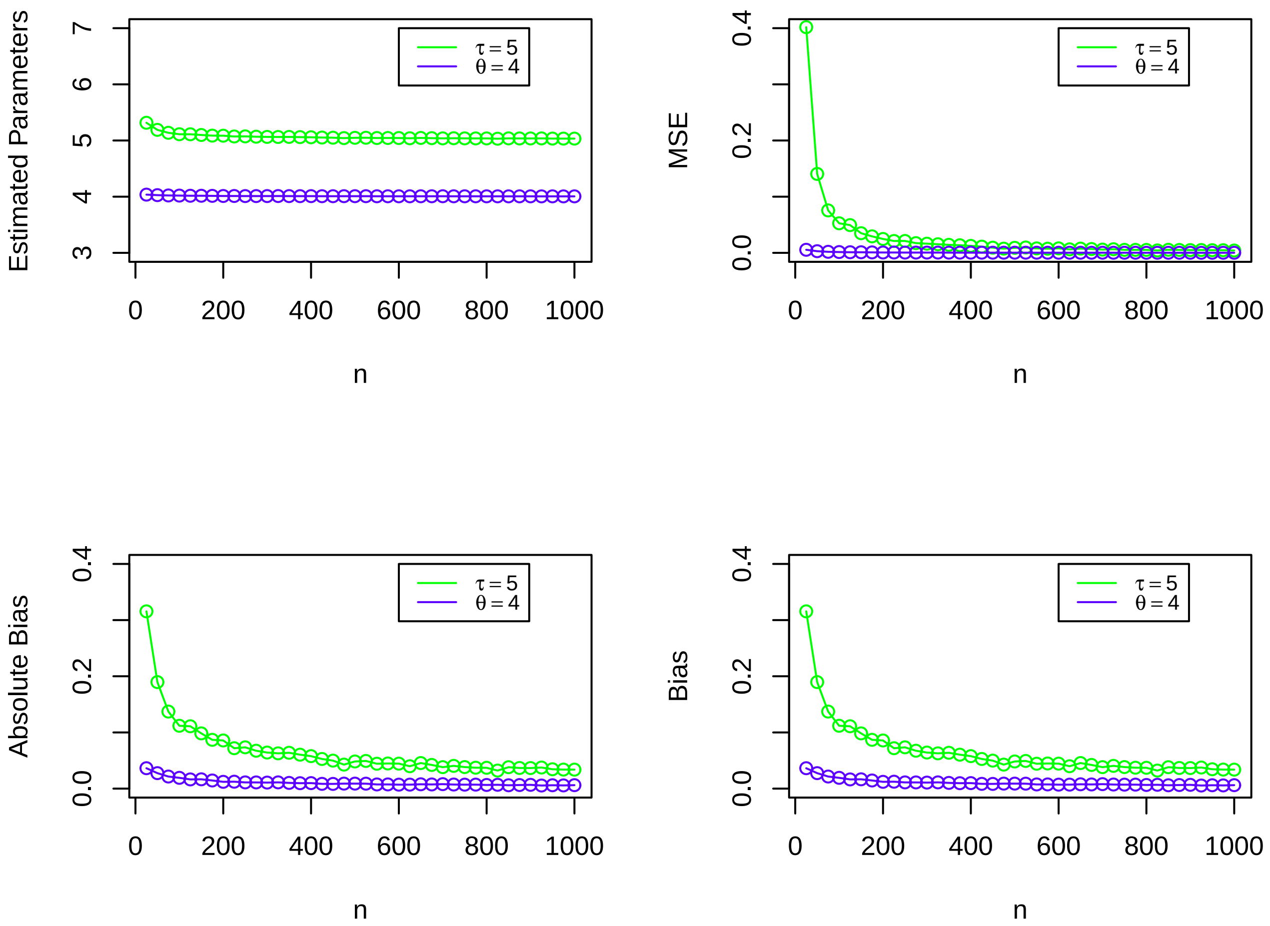

Figure 3. Plots of the estimated parameters, MSEs, absolute biases, and biases for $\tau = 5$ and $\theta = 4$

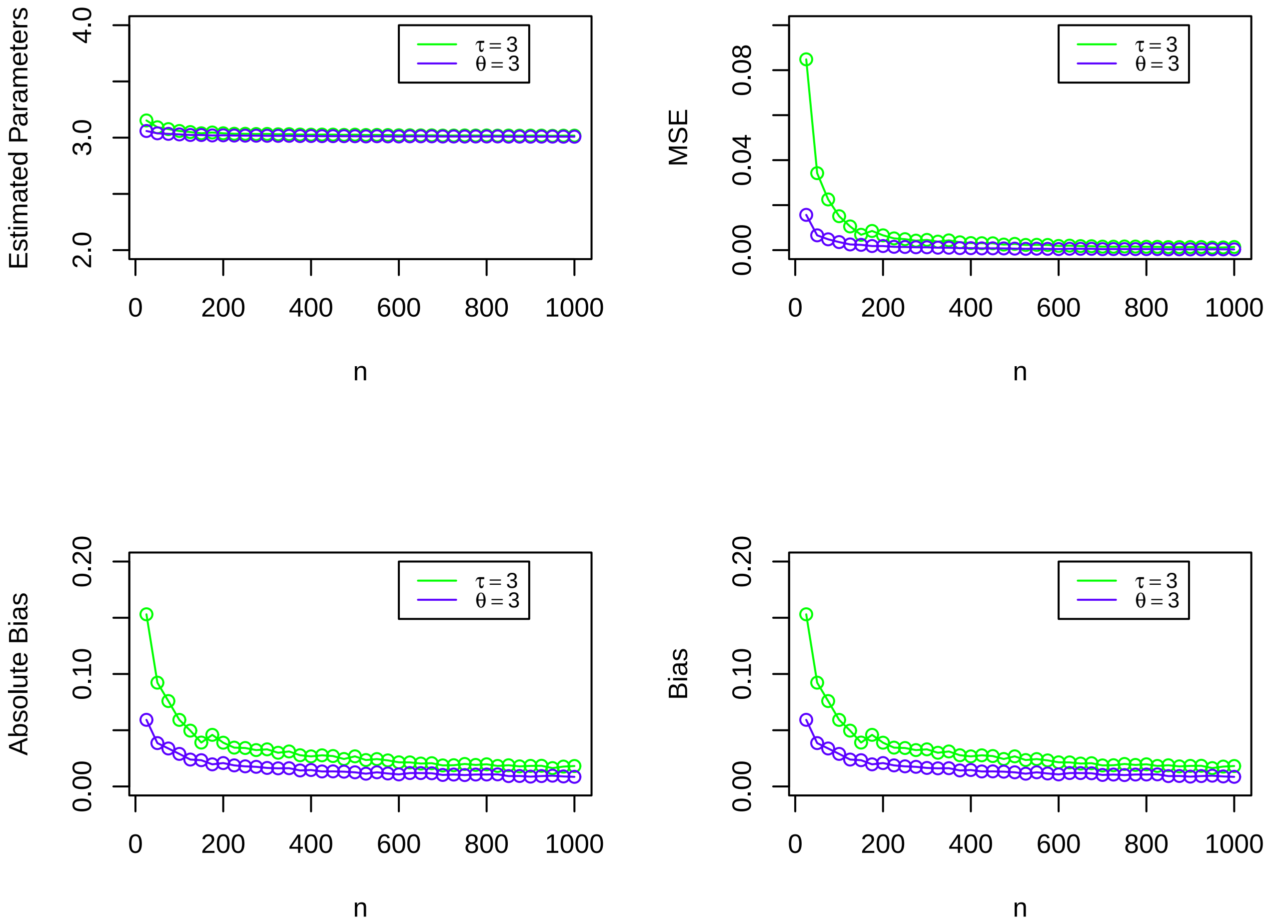

Figure 4. Plots of the estimated parameters, MSEs, absolute biases, and biases for $\tau = 3$ and $\theta = 3$

7. Applications

The significance of the XE-G distribution is investigated using two authentic datasets. These datasets are described as follows:

Data Set 1. Life time data.

This dataset contains information about the durations of relief experienced by a cohort of 20 patients who were administered a specific analgesic treatment. The original data was recorded by Gross and Clark (1975, p. 105). The dataset consists of values representing the time of relief measured in minutes. These time intervals serve as indicators of the efficacy of the analgesic treatment. The recorded relief times are as follows: 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0. The descriptive statistics of the data set is provided in Table 2.

Table 2. Descriptive statistics for data set 1

n	Mean	Median	Min	Max	Variance	Skewness	Kurtosis
20	1.9	1.7	1.1	4.1	0.4958	1.5924	2.3465

Data Set 2. Covid-19 Canada

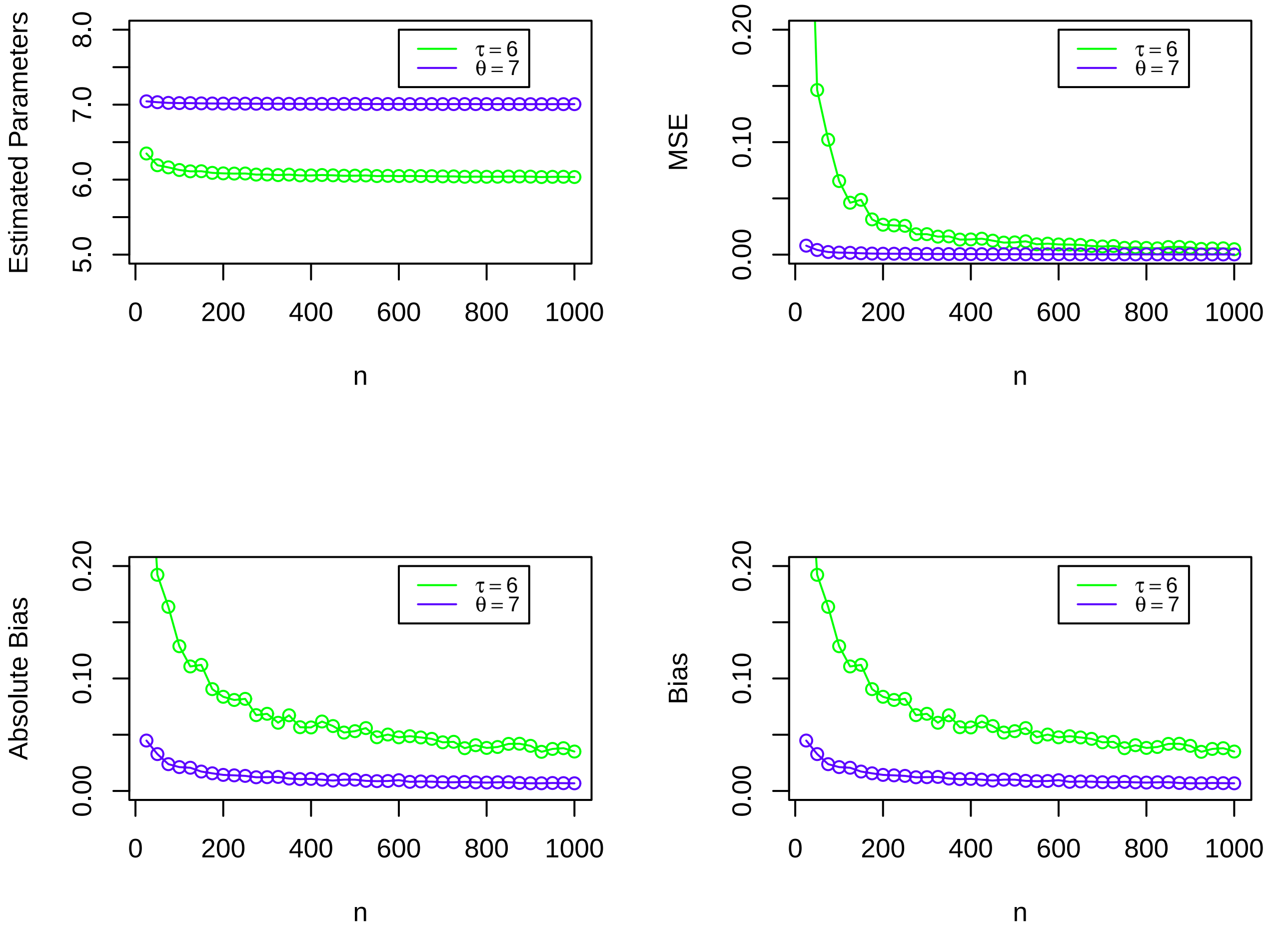

Figure 5. Plots of the estimated parameters, MSEs, absolute biases, and biases for $\tau = 6$ and $\theta = 7$

This dataset pertains to the COVID-19 mortality rate in Canada from April 10th to May 15th, 2020. The source of this data is the World Health Organization (WHO) and it is available on the WHO Covid-19 website's main page. This specific dataset has been examined and analyzed in separate studies conducted by Altmetwally et al. (2021) and Klakattawi et al. (2022). The descriptive statistics of the data set is provided in Table 3.

Table 3. Descriptive statistics for data set 2

n	Mean	Median	Min	Max	Variance	Skewness	Kurtosis
36	3.2816	3.1777	1.5157	6.8686	0.9971	1.1637	2.8146

To estimate the unknown parameters of the distributions, the standard maximum likelihood method is used through the nlmixed procedure in SAS (www.sas.com). The performance of the XE-G distribution, specially with XE-IP and XE-Gu is compared to several other established distributions: Lindley exponential, exponential exponential, exponentiated Gumbel, alpha power transformed generalized exponential (α-PTGE), odd exponential Fréchet (OE-Fr), alpha power Weibull, power Lindley, and exponentiated Weibull.

The selection of each comparative model is based on goodness-of-fit criteria using likelihood functions, including the negative log-likelihood (NLL), Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike's information (CAIC), and Hannan-Quinn information (HQIC). Lower values of these indicate a better model fit. Additionally, distribution function-based statistics like Kolmogorov-Smirnov (K-S), Anderson-Darling (A-D), and Cramér-Von

Mises (C-M), along with their p-values, is considered. For these statistics, lower values and higher p-values point to a better fit.

Tables 4 and 5 present the MLEs alongside their corresponding standard errors (SEs), as well as statistical values and p-values for the distribution models applied to both dataset 1 and dataset 2. Figures 6 and 7 illustrate the estimated pdf, cdf, hazard rate, Kaplan-Meier survival curve, P-P plots, and Q-Q plots for the NE-IP distribution concerning dataset 1 and dataset 2.

Upon examination of Tables 4 and 5, it becomes evident that among the range of models under consideration, the XE-IP and XE-Gu models consistently yields the most favorable results with the lowest values of statistics such as NLL, AIC, BIC, CAIC, and HQIC in both datasets. Additionally, the K-S, A-D, and C-M statistics associated with the XE-G models are notably lower, leading to higher p-values for both datasets. These observations strongly indicate that the XE-G models offers the most appropriate fit to both datasets in comparison to the alternative models. These conclusions are further substantiated by the supporting evidence found in Figures 6 and 7.

Table 4. MLEs and statistics of the distributions for data set 1

Distribution	MLEs (SEs)	NLL	AIC	BIC	CAIC	HQIC	K-S (p-value)	A-D (p-value)	C-M (p-value)
XE-IP	$\tau = 4.2099(0.971)$ $\theta = 1.2162(0.092)$	15.348	34.7	36.7	35.4	35.1	0.096 (0.992)	0.138 (0.999)	0.022 (0.995)
XE-Gu	$\mu = 1.6564(0.127)$ $\sigma = 0.4142(0.118)$	15.410	34.8	36.8	35.5	35.2	0.099 (0.990)	0.162 (0.998)	0.025 (0.992)
Lindley exponential	$\lambda = 2.2337(0.437)$ $\theta = 37.547(25.29)$	16.259	36.5	38.5	37.2	36.9	0.134 (0.863)	0.310 (0.930)	0.048 (0.895)
Exponential exponential	$\alpha = 36.683(25.26)$ $\lambda = 2.2352(0.436)$	16.261	36.5	38.5	37.2	36.9	0.134 (0.863)	0.311 (0.929)	0.048 (0.895)
Exponentiated Gumbel	$\alpha = 0.3021\ (0.282)$ $\mu = 1.3025\ (0.182)$ $\sigma = 0.1908\ (0.138)$	15.44	36.9	39.9	38.4	37.5	0.100 (0.988)	0.187 (0.994)	0.032 (0.972)
α-PTGE	$\lambda = 1.7347\ (0.642)$ $\alpha = 0.1233\ (0.272)$ $\beta = 26.171\ (19.87)$	15.83	37.7	40.6	39.2	38.2	0.109 (0.973)	0.216 (0.986)	0.030 (0.979)
OE-Fr	$\lambda = 4.2991\ (6.662)$ $\theta = 2.8382\ (1.760)$ $\tau = 1.5421\ (0.689)$	16.99	40.0	43.0	41.5	40.6	0.167 (0.629)	0.501 (0.744)	0.091 (0.634)
Alpha power Weibull	$\alpha = 0.0156\ (0.034)$ $\theta = 2.9204\ (0.430)$ $\tau = 3.6240\ (0.564)$	18.71	43.4	46.4	44.9	44.0	0.158 (0.698)	0.679 (0.575)	0.102 (0.581)
Power Lindley	$\alpha = 2.2530(0.307)$ $\beta = 0.3445(0.099)$	20.432	44.9	46.9	45.6	45.3	0.188 (0.482)	1.037 (0.338)	0.175 (0.323)
Exponentiated Weibull	$c = 2.7870(0.427)$ $\lambda = 0.1216(0.056)$	20.586	45.2	47.2	45.9	45.6	0.185 (0.501)	1.083 (0.316)	0.183 (0.304)

Table 5. MLEs and statistics of the distributions for data set 2

Distribution	MLEs (SEs)	NLL	AIC	BIC	CAIC	HQIC	K-S (p-value)	A-D (p-value)	C-M (p-value)
XE-Gu	μ = 3.7660(0.253)	47.552	99.1	102.3	99.5	100.2	0.101	0.375	0.060
	σ = 1.4215(0.253)						(0.858)	(0.872)	(0.817)
XE-IP	τ = 2.7619(0.376)	47.609	99.2	102.4	99.6	100.3	0.105	0.427	0.070
	θ = 2.8259(0.240)						(0.818)	(0.821)	(0.756)
α-PTGE	λ = 1.4781 (0.198)	47.427	100.9	105.6	101.6	102.5	0.102	0.397	0.064
	α = 30.032 (69.32)						(0.851)	(0.851)	(0.789)
	β = 23.512 (19.30)								
Exponentiated	α = 1.8834 (1.334)	47.898	101.8	106.5	102.5	103.5	0.105	0.476	0.080
Gumbel	μ = 3.3548 (0.674)						(0.826)	(0.771)	(0.693)
	σ = 1.1121 (0.395)								
Exponential	λ = 1.1984(0.158)	48.513	101.0	104.2	101.4	102.1	0.124	0.673	0.114
exponential	α = 29.072(12.58)						(0.642)	(0.581)	(0.521)
Lindley	λ = 1.1980(0.158)	48.519	101.0	104.2	101.4	102.1	0.124	0.674	0.115
exponential	θ = 29.979(12.61)						(0.640)	(0.580)	(0.520)
OE-Fr	λ = 24.538 (50.42)	48.220	102.4	107.2	103.2	104.1	0.108	0.555	0.095
	θ = 10.710 (12.25)						(0.791)	(0.690)	(0.612)
	τ = 1.0518 (0.528)								
Alpha power	α = 0.0061 (0.013)	48.457	102.9	107.7	103.7	104.6	0.120	0.545	0.093
Weibull	θ = 5.0339 (0.538)						(0.674)	(0.700)	(0.622)
	τ = 4.2977 (0.514)								
Power Lindley	α = 2.5035(0.260)	50.169	104.3	107.5	104.7	105.4	0.146	0.853	0.150
	β = 0.0833(0.030)						(0.429)	(0.443)	(0.390)
Exponentiated	c = 3.3136(0.380)	51.474	106.9	110.1	107.3	108.1	0.150	1.143	0.198
Weibull	λ = 0.0139(0.008)						(0.392)	(0.290)	(0.272)

8. Conclusion

This study proposes a new family of distributions called the X-exponential-G (XE-G) family, in which inverse paralogistic (IP) and Gumbel (Gu) distributions are used as baseline distributions. The XE-G family have been defined mathematically, with derivations of their statistical properties. To demonstrate the effectiveness of this new approach, the XE-IP and XE-Gu distributions are fitted to real-world datasets and compared against other competitive models. Results across several goodness-of-fit measures showed that these distributions provided superior fits to the data compared to other models. Future investigations into the XE-G family could explore the introduction of new members with diverse baseline distributions, evaluate different parameter estimation methods, including Bayesian approaches, and develop regression models using XE-G family distributions as response variables. These endeavors aim to further enhance the versatility and effectiveness of the XE-G family as a statistical modeling tool across various applications.

Conflict of interest

I have no conflicts of interest to disclose.

Data availability

The manuscript cites pre-existing data, as no new data is generated or analyzed in this study.

References

Alexander, C., Cordeiro, G. M., Ortega, E. M., & Sarabia, J. M. (2012). Generalized beta-generated distributions. *Computational Statistics & Data Analysis, 56*(6), 1880-1897.

Alizadeh, M., Cordeiro, G. M., Pinho, L. G. B., & Ghosh, I. (2017). The Gompertz-G family of distributions. *Journal of Statistical Theory and Practice, 11*, 179-207.

Alizadeh, M., Emadi, M., Doostparast, M., Cordeiro, G. M., Ortega, E. M., & Pescim, R. R. (2015). A new family of distributions: the Kumaraswamy odd log-logistic, properties and applications. *Hacettepe Journal of Mathematics and Statistics, 44*(6), 1491-1512.

Alizadeh, M., Rasekhi, M., Yousof, H. M., & Hamedani, G. G. (2018). The transmuted Weibull-G family of distributions. *Hacettepe Journal of Mathematics and Statistics, 47*(6), 1671-1689.

Alizadeh, M., Tahir, M. H., Cordeiro, G. M., Mansoor, M., Zubair, M., & Hamedani, G. (2015). The Kumaraswamy marshal-Olkin family of distributions. J*ournal of the Egyptian Mathematical Society, 23*(3), 546-557.

Almetwally, E. M., Alharbi, R., Alnagar, D., & Hafez, E. H. (2021). A new inverted topp-leone distribution: applications to the COVID-19 mortality rate in two different countries. *Axioms, 10*(1), 25.

Alzaatreh, A., & Ghosh, I. (2015). On the Weibull-X family of distributions. *Journal of Statistical Theory and Applications, 14*(2), 169-183.

Bhati, D., Malik, M. A., & Vaman, H. J. (2015). Lindley–Exponential distribution: properties and applications. *Metron, 73*, 335-357.

Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. *Journal of Data Science, 12*(1), 53-68.

Cordeiro, G. M., Afify, A. Z., Yousof, H. M., Pescim, R. R., & Aryal, G. R. (2017). The exponentiated Weibull-H family of distributions: Theory and Applications. *Mediterranean Journal of Mathematics, 14*, 1-22.

Cordeiro, G. M., Alizadeh, M., & Ortega, E. M. (2014). The exponentiated half-logistic family of distributions: Properties and applications. *Journal of Probability and Statistics,* 2014.

Cordeiro, G. M., Ortega, E. M., & da Cunha, D. C. (2013). The exponentiated generalized class of distributions. *Journal of Data Science, 11*(1), 1-27.

Cordeiro, G. M., Ortega, E. M., Popović, B. V., & Pescim, R. R. (2014). The Lomax generator of distributions: Properties, minification process and regression model. *Applied Mathematics and Computation*, *247*, 465-486.

Dey, S., Alzaatreh, A., Zhang, C., & Kumar, D. (2017). A new extension of generalized exponential distribution with application to ozone data. *Ozone: Science & Engineering, 39*(4), 273-285.

Eghwerido, J. T., & Agu, F. I. (2021). The shifted Gompertz-G family of distributions: properties and applications. *Mathematica Slovaca, 71*(5), 1291-1308.

Eghwerido, J. T., Agu, F. I., & Ibidoja, O. J. (2022). The shifted exponential-G family of distributions: Properties and applications. *Journal of Statistics and Management Systems, 25*(1), 43-75.

Eghwerido, J. T., Efe-Eyefia, E., & Zelibe, S. C. (2021). The transmuted alpha power-G family of distributions. *Journal of Statistics and Management Systems, 24*(5), 965-1002.

Eghwerido, J. T., Ikwuoche, J. D., & Adubisi, O. D. (2020). Inverse odd Weibull generated family of distribution. *Pakistan Journal of Statistics and Operation Research,* 617-633.

Eghwerido, J. T., Oguntunde, P. E., & Agu, F. I. (2021). The alpha power Marshall-Olkin-G distribution: properties, and applications. *Sankhya A*, 1-26.

Eliwa, M. S., & El-Morshedy, M. (2019). Bivariate Gumbel-G family of distributions: statistical properties, Bayesian and non-Bayesian estimation with application. *Annals of Data Science, 6*(1), 39-60.

Eugene, N., Lee, C., & Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods, 31*(4), 497-512.

Ghitany, M. E., Al-Mutairi, D. K., Balakrishnan, N., & Al-Enezi, L. J. (2013). Power Lindley distribution and associated inference. *Computational Statistics & Data Analysis, 64*, 20-33.

Gross, A. J., & Clark, V. (1975). Survival distributions: reliability applications in the biomedical sciences.

Gupta, R. D., & Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences, 43*(1), 117-130.

Klakattawi, H., Alsulami, D., Elaal, M. A., Dey, S., & Baharith, L. (2022). A new generalized family of distributions based on combining Marshal-Olkin transformation with TX family. *PloS one, 17*(2), e0263673.

Merovci, F., Alizadeh, M., Yousof, H. M., & Hamedani, G. G. (2017). The exponentiated transmuted-G family of distributions: theory and applications. *Communications in Statistics-Theory and Methods, 46*(21), 10800-10822.

Nadarajah, S., Cordeiro, G. M., & Ortega, E. M. (2012). General results for the Kumaraswamy-G distribution. *Journal of Statistical Computation and Simulation, 82*(7), 951-979.

Nadarajah, S., & Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematica, 92*, 97-111.

Nassar, M., Alzaatreh, A., Mead, M., & Abo-Kasem, O. (2017). Alpha power Weibull distribution: Properties and applications. *Communications in Statistics-Theory and Methods, 46*(20), 10236-10252.

Reyad, H., Alizadeh, M., Jamal, F., & Othman, S. (2018). The Topp Leone odd Lindley-G family of distributions: Properties and applications. *Journal of Statistics and Management Systems, 21*(7), 1273-1297.

Ristić, M. M., & Balakrishnan, N. (2012). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation, 82*(8), 1191-1206.

Tahir, M. H., Cordeiro, G. M., Alizadeh, M., Mansoor, M., Zubair, M., & Hamedani, G. G. (2015). The odd generalized exponential family of distributions with applications. *Journal of Statistical Distributions and Applications, 2*, 1-28.

Tahir, M. H., Cordeiro, G. M., Alzaatreh, A., Mansoor, M., & Zubair, M. (2016). The logistic-X family of distributions and its applications. *Communications in Statistics-Theory and Methods, 45*(24), 7326-7349.

Tahir, M. H., Zubair, M., Mansoor, M., Cordeiro, G. M., ALİZADEHK, M., & Hamedani, G. (2016). A new Weibull-G family of distributions. *Hacettepe Journal of Mathematics and Statistics, 45*(2), 629-647.

Torabi, H., & Hedesh, N. M. (2012). The gamma-uniform distribution and its applications. *Kybernetika, 48*(1), 16-30.

, H., & Montazeri, N. H. (2014). The logistic-uniform distribution and its applications. *Communications in Statistics-Simulation and Computation, 43*(10), 2551-2569.

Yousof, H. M., Afify, A. Z., Alizadeh, M., Butt, N. S., & Hamedani, G. (2015). The transmuted exponentiated generalized-G family of distributions. *Pakistan Journal of Statistics and Operation Research*.

Zografos, K., & Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology, 6*(4), 344-362.

Acknowledgments

Not applicable

Authors contributions

Dr. Shahid Mohammad was responsible for all contributions to the article.

Funding

Not applicable

Competing interests

The author declares no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education. The journal¡¯s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

Copyrights