MJSAT

Malaysian Journal of Science and Advanced Technology

journal homepage: https://mjsat.com.my/

Exploring the New Sine Unit Gompertz Distribution: Theory, Inference, and Practical Relevance to Financial and Medical Analytics

Ehinomen Emmanuel Ehizojie*[1]

[1]*Department of Statistics, Ahmadu Bello University, Zaria, Nigeria.*

KEYWORDS

Unit Gompertz Distribution
New Sine-G family
Trade share data Estimations
Hazard rates
Relief times data

ARTICLE HISTORY

Received 2 January 2026
Received in revised form 29 May 2026
Accepted 12 June 2026
Available online 26 June 2026

ABSTRACT

This study proposes the New Sine Unit Gompertz distribution (NSUGD), a flexible bounded lifetime model derived from the Unit Gompertz distribution. The probability density function accommodates decreasing, near-symmetric, right-skewed, and left-skewed shapes, while the hazard rate function exhibits increasing, J-shaped, or bathtub forms, enabling effective modelling of unit-interval data across diverse applications. Some statistical properties are derived, including the survival and hazard rate functions, quantile function, moments, median, mean, skewness, and kurtosis. Parameters are estimated via maximum likelihood, maximum product of spacings, and Cramér-von Mises methods; a Monte Carlo simulation study, evaluating average estimates and mean squared error, identifies maximum likelihood as the most efficient approach. Empirical validation on financial and medical datasets demonstrates the NSUGD's superior fit relative to competitors such as the unit Gompertz, Kumaraswamy, and beta distributions, as evidenced by lower AIC and BIC values and non-significant Kolmogorov-Smirnov statistics with higher p-values. Owing to its adaptability, the NSUGD holds promise for diverse fields, including engineering, environmental science, economics, insurance, and other disciplines where unit bounded data are prevalent.

1. INTRODUCTION

Statistical distribution theory forms the cornerstone of modern statistical modelling and inference, offering structured ways to describe and interpret random phenomena through probability distributions. Each distribution provides a mathematical representation of how random variables behave across their range of possible values, thereby aiding in understanding uncertainty, variability, and stochastic behaviour. For bounded data constrained within 0 and 1, such as proportions, success probabilities, growth rates, indices, and percentages; traditional unbounded distributions like the Exponential, Gompertz, or Normal prove inadequate, often necessitating transformations to enforce support constraints while preserving flexibility in skewness, kurtosis, and tail behaviour [1]. Modelling bounded data accurately is crucial inmany scientific disciplines, as it prevents unrealistic predictions and offers a more precise representation of the underlying processes. The Beta and Kumaraswamy distributions address many such needs [2], yet complex empirical patterns in skewness or hazard rates often exceed their scope, spurring advances in generalised bounded models for reliability, finance, and environmental applications [3].

The Unit Gompertz distribution (UGD), introduced by [3], represents a bounded transformation of the classical Gompertz law of mortality, originally formulated to model human mortality rates with an exponential increase over age [4]. Defined on the unit interval (0, 1), this distribution arises from the transformation $X = exp(-Y)$, where Y follows a Gompertz distribution with and as the scale and shape parameters respectively. Its probability density function (PDF) and cumulative distribution function (CDF) take the form

$$g(x) = \alpha\beta x^{-\beta-1} e^{-\alpha(x^{-\beta}-1)} \quad (1)$$

$$G(x) = e^{-\alpha(x^{-\beta}-1)}, \quad (2)$$

yielding a PDF that is increasing, reversed J-shaped and right-skewed; alongside increasing or upside-down bathtub shaped hazard function suited to lifetime data exhibiting such failure rate characteristics. Despite its utility in reliability analysis and survival modelling, the UGD exhibits limitations in flexibility, particularly in accommodating symmetric, left-skewed shapes prevalent in diverse bounded datasets.

*Corresponding author:
E-mail address: Ehinomen Emmanuel Ehizojie < ehizojieee@gmail.com >.
https://doi.org/10.56532/mjsat.v6i2.689
2785-8901/

The New Sine-G family of distributions (NS-G), proposed by [5], uses a sine transformation that introduces non-linear bending in the CDF. The sine function is bounded, smooth, and maps the unit interval onto itself, making it a mathematically natural mechanism for constructing CDFs on (0, 1) without requiring additional parameters. Its CDF and PDF are defined as follows

$$F(x) = sin\left[\frac{\pi}{4}G(x)\{G(x)+1\}\right], \qquad x \in \mathbb{R}, \tag{3}$$

$$f(x) = \frac{\pi}{4}g(x)\{2G(x)+1\}\,cos\left[\frac{\pi}{4}G(x)\{G(x)+1\}\right], \tag{4}$$

respectively, where G(x) and g(x) are the CDF and PDF of a given univariate continuous distribution.

Unlike polynomial-based generators, the trigonometric transformation introduces oscillatory curvature into the baseline CDF, enriching the resulting PDF to exhibit multiple local modes or asymmetric shapes even when the baseline density is monotonic, and as well the hazard shapes at no cost to parsimony [5].

In reliability data, periodic patterns are not literal cycles but rather recurring stress-relaxation processes, seasonal maintenance effects, or oscillatory degradation in mechanical systems; such as vibrations in bearings [7]. Thus, the sine function's inherent periodicity makes it a natural choice for modelling such behaviors. When applied to the UGD, this transformation produces the New Sine Unit Gompertz Distribution (NSUGD). As a result, the NSUGD is better equipped to represent complex distributional shapes compared to the baseline distribution and other well-known conventional bounded models.

In practice, reliability data may exhibit non-monotonic patterns due to cyclic stress, environmental variation, or repeated usage. For instance,

i. fatigue crack growth under cyclic loading where damage accumulates in a step-like periodic manner;
ii. electronic components subjected to daily temperature cycles causing periodic stress;
iii. mechanical systems with rotating parts where failure risk oscillates with rotation angle;
iv. biological responses influenced by treatment cycles; and,
v. alternating phases of remission and relapse in medical conditions;

may display fluctuating failure behaviour [7], [8], and [9]. The sine transformation allows the NSUGD to approximate such patterns more effectively. Hence, the NSUGD stands as a valuable addition to the class of bounded lifetime distributions, promising analytical precision in diverse applications, including actuarial science and biostatistics.

The structure of this paper is organised as follows. Section 2 introduces the formulation of the NSUGD along with its reliability measures and graphical illustrations. Section 3 presents some statistical properties of the NSUGD, including its quantile function, median, moments, skewness, and kurtosis. Section 4 describes the estimation techniques employed for parameter estimation, while Section 5 provides a detailed Monte Carlo simulation study to assess the performance of the estimators. Section 6 demonstrates the practical applicability of the NSUGD using lifetime datasets, and Section 7 offers the concluding remarks and key insights drawn from the study.

2. THE NEW SINE UNIT GOMPERTZ DISTRIBUTION

By substituting Equations (1) and (2) into (3) and (4), the proposed model is derived. Accordingly, the PDF and CDF of a random variable X following the proposed NSUGD are defined as follows

$$f(x) = \frac{\pi}{4}\alpha\beta x^{-\beta-1}e^{-\alpha(x^{-\beta}-1)}\left(2e^{-\alpha(x^{-\beta}-1)}+1\right)cos\left[\frac{\pi}{4}e^{-\alpha(x^{-\beta}-1)}\left(e^{-\alpha(x^{-\beta}-1)}+1\right)\right], \quad x \in (0,1) \tag{5}$$

and

$$F(x) = sin\left[\frac{\pi}{4}e^{-\alpha(x^{-\beta}-1)}\left(e^{-\alpha(x^{-\beta}-1)}+1\right)\right], \tag{6}$$

where $\alpha > 0$ and $\beta > 0$ are scale and shape parameters respectively. The survival function, hazard function (HF) are respectively given by

$$S(x) = 1 - sin\left[\frac{\pi}{4}e^{-\alpha(x^{-\beta}-1)}\left(e^{-\alpha(x^{-\beta}-1)}+1\right)\right], \tag{7}$$

and

$$h(x) = \frac{\pi}{4}\alpha\beta x^{-\beta-1}e^{-\alpha(x^{-\beta}-1)}(2e^{-\alpha(x^{-\beta}-1)}+1) \times \frac{cos\left[\frac{\pi}{4}e^{-\alpha(x^{-\beta}-1)}\left(e^{-\alpha(x^{-\beta}-1)}+1\right)\right]}{1-sin\left[\frac{\pi}{4}e^{-\alpha(x^{-\beta}-1)}\left(e^{-\alpha(x^{-\beta}-1)}+1\right)\right]}. \tag{8}$$

The graphical representation of the PDF, CDF, and HF of the NSUGD for different parameter settings are given in Figure 1. The Figure illustrates the diverse shapes exhibited by the NSUGD under varying parameter combinations. The PDF demonstrates remarkable flexibility, capturing a range of distributional forms including monotonically decreasing, nearly symmetric, right-skewed, and left-skewed unimodal patterns. This adaptability highlights the NSUGD's potential to model diverse datasets within the unit interval. The corresponding CDFs reveal smooth and consistent growth across all parameter settings, confirming the distribution's validity and bounded nature.

Furthermore, the HF plots display increasing, J-shaped, and bathtub-shaped profiles, reflecting the versatility of the NSUGD in characterising diverse failure behaviours. The bathtub-shaped hazard, evident at $\alpha = 0.5, \beta = 0.1$ (Figure 1), comprises three clinically meaningful phases: an initial elevated-risk period corresponding to early complications or post-treatment relapse; a stable intermediate phase reflecting consistent recovery; and a terminal increasing-risk phase associated with disease progression, treatment wear-out, or age-related deterioration [6]. This tripartite structure enables the NSUGD to simultaneously accommodate early- and late-life failures within a single parsimonious model, offering clinicians a valuable tool for designing risk-stratified follow-up strategies.

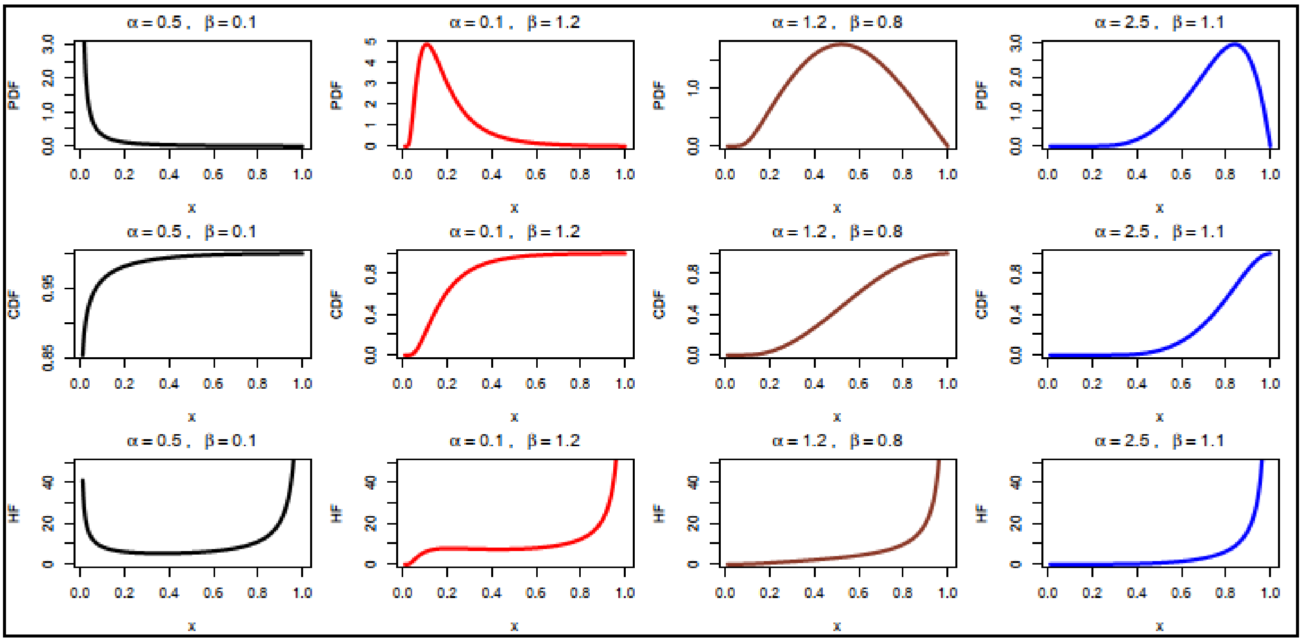

Fig. 1. The PDF, CDF and HF plots of the NSUGD for different values of the parameters.

3. Statistical Features

3.1 Mixture Representation for NSUGD

The closed-form CDF and PDF of the NSUGD involve nested exponential and trigonometric terms that render direct analytical integration intractable. Consequently, derivation of moments and related properties require a series expansion. As shown by [7], the CDF and PDF of the NSUGD admit the following series representations:

$$F(x) = \sum_{k=0}^{+\infty}\sum_{l=0}^{2k+1} a_{k,l}\left[e^{-\alpha(x^{-\beta}-1)}\right]^{l+2k+1}, \quad (9)$$

and

$$f(x) = \alpha\beta\sum_{k=0}^{+\infty}\sum_{l=0}^{2k+1} b_{k,l}x^{-\beta-1}\left[e^{-\alpha(x^{-\beta}-1)}\right]^{l+2k+1}, \quad (10)$$

where

$$a_{k.l} = \frac{(-1)^k}{(2k+1)!}\left(\frac{\pi}{4}\right)^{2k+1}\frac{(2k+1)!}{l!(2k+1-l)!},$$

$$b_{k,l} = (l+2k+1)a_{k,l}$$

3.2 Quantile Function

Let $Q(u)$ be the quantile function (QF) corresponding to $F(x)$, that is, satisfying $F(Q(u)) = u$ for $u \in (0,1)$. Thus, the QF corresponding to $F(x)$ is given by

$$Q(u) = \left[1 - \frac{1}{\alpha}ln\left\{\frac{-1+\sqrt{1+\frac{16}{\pi}arcsin(u)}}{2}\right\}\right]^{-\frac{1}{\beta}}, \quad (11)$$

setting $u = 0.5$, the median of NSUGD is given by

$$x_{0.5} = \left[1 - \frac{1}{\alpha}ln\left\{\frac{-1+\sqrt{1+\frac{16}{\pi}arcsin(0.5)}}{2}\right\}\right]^{-\frac{1}{\beta}}. \quad (12)$$

Let u be a random variable following the uniform distribution on the interval (0, 1). The QF of the NSUGD can therefore be employed to generate random samples of size n from the distribution. This is achieved by substituting the uniform random variates $u_1, u_2, \dots, u_n$ into Equation (11), thereby yielding the corresponding random observations from the NSUGD as

$$x_i = \left[1 - \frac{1}{\alpha}ln\left\{\frac{-1+\sqrt{1+\frac{16}{\pi}arcsin(u_i)}}{2}\right\}\right]^{-\frac{1}{\beta}}. \quad (13)$$

Table 1 presents the quartile values corresponding to selected parameter combinations of the proposed model, while their graphical representations are depicted in Figure 2. It is evident that the quantiles increase with rising values of both parameters, indicating the influence of α and β on the model's location and spread.

Table 1. The Quantiles for some parameter values of the NSUGD (α, β).

u	(0.5, 0.1)	(0.1, 1.2)	(1.2, 0.8)	(2.5, 1.1)
0.1	0.00000	0.07419	0.27530	0.56688
0.2	0.00000	0.09655	0.35417	0.64452
0.3	0.00000	0.11733	0.41865	0.69877
0.4	0.00002	0.13892	0.47736	0.74260
0.5	0.00008	0.16295	0.53402	0.78078
0.6	0.00029	0.19141	0.59121	0.81587
0.7	0.00107	0.22768	0.65157	0.84971
0.8	0.00430	0.27921	0.71910	0.88425
0.9	0.02271	0.36971	0.80314	0.92309

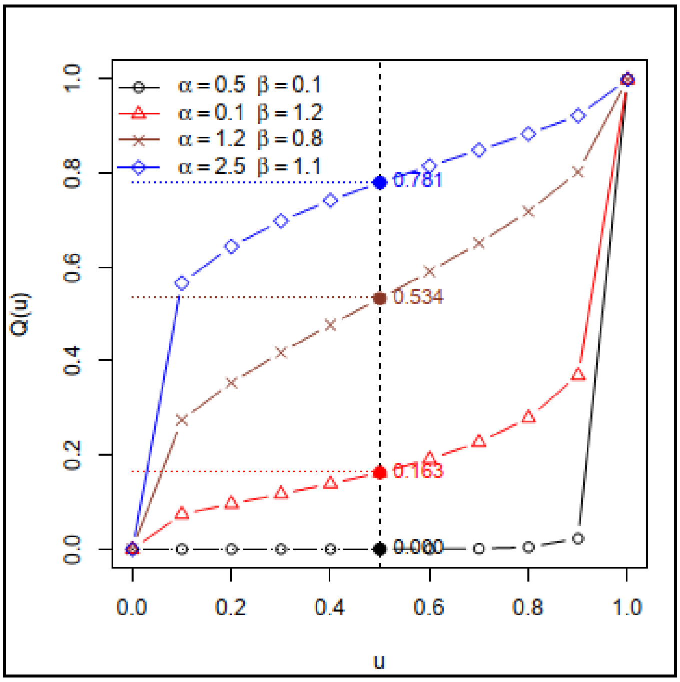

Fig. 2. The QF plots with the median of the NSUGD across various parameter values.

3.3 Moments

From the PDF series expansion given in Equation (10), the rth moments of the NSUGD is obtained as follows

$$\mu_r' = \mathrm{E}[x^r],$$

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

$$= \sum_{k=0}^{+\infty}\sum_{l=0}^{2k+1} b_{k,l}\int_0^1 x^{r-\beta-1}\left[e^{-\alpha(x^{-\beta}-1)}\right]^{l+2k+1} dx,$$

set $m = l + 2k + 1, m \in \mathbb{Z}_{>0}$, but

$$l = \int_0^1 x^{r-\beta-1}\left[e^{-\alpha(x^{-\beta}-1)}\right]^{l+2k+1} dx,$$

$$= e^{\alpha m}\int_0^1 x^{r-\beta-1}\, e^{-\alpha m x^{-\beta}} dx,$$

let $t = \alpha m x^{-\beta} \rightarrow x = \left(\frac{\alpha m}{t}\right)^{\frac{1}{\beta}}$ and $dx = -\frac{1}{\beta}(\alpha m)^{\frac{1}{\beta}} t^{\frac{-1}{\beta}-1} dt$.

The limits $x: 0 \rightarrow 1$, map to $t: \infty \rightarrow \alpha m$. After simplifying the powers of t and αm the integral becomes

$$I = \frac{1}{\beta}(\alpha m)^{\frac{r}{\beta}}\int_{\alpha m}^{\infty} t^{\left(1-\frac{r}{\beta}\right)-1} e^{-t} dt = \frac{1}{\beta}(\alpha m)^{\frac{r}{\beta}}\Gamma\left(1-\frac{r}{\beta}, \alpha m\right),$$

Therefore, the rth moment of NSUGD has the form

$$\mu_r' = \frac{\alpha^{\frac{r}{\beta}}}{\beta}\sum_{k=0}^{+\infty}\sum_{l=0}^{2k+1} b_{k,l} m^{\frac{r}{\beta}}\Gamma\left(1-\frac{r}{\beta}, \alpha m\right), \quad (14)$$

where

$$m = l + 2k + 1,$$

$$a_{k.l} = \frac{(-1)^k}{(2k+1)!}\left(\frac{\pi}{4}\right)^{2k+1}\frac{(2k+1)!}{l!\,(2k+1-l)!}$$

$$b_{k,l} = (l + 2k + 1)a_{k,l}.$$

Let $\mu_1', \mu_2', \mu_3', \mu_4'$ be the 1st, 2nd, 3rd, 4th moments. Using Equation (14), the mean, variance, skewness, kurtosis in terms of rth moments of the NSUGD can be computed as

Mean: $\mu = \mathrm{E}[x] = \mu_1'$,

Variance: $\sigma^2 = \mu_2' - (\mu_1')^2$,

Skewness: $S_k = \dfrac{\mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3}{\sigma^3}$,

Kurtosis: $K_u = \dfrac{\mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4}{\sigma^4}$.

The numerical values of these measures are presented in Table 2. The results reveal that as both parameters increase, the mean and variance decline, whereas skewness and kurtosis rise substantially. This pattern suggests that the distribution becomes increasingly right-skewed and heavy-tailed with higher parameter values, making it well suited for modelling asymmetric and leptokurtic data.

Table 2. Mean, variance, skewness and kurtosis of the NSUGD (α, β).

α	β	Mean	Variance	Skewness	Kurtosis
0.2	4.1	0.0968	0.0469	1.5805	2.3684
0.2	4.5	0.0924	0.0481	1.7986	3.4203
0.2	5.0	0.0872	0.0488	2.0271	4.5297
0.4	4.1	0.1483	0.0946	1.6399	3.8065
0.4	4.5	0.1381	0.0917	1.7764	4.2592
0.4	5.0	0.1272	0.0879	1.9360	4.8387
0.8	4.1	0.1234	0.0927	2.0935	5.4657
0.8	4.5	0.1138	0.0877	2.2440	6.1104
0.8	5.0	0.1038	0.0820	2.4204	6.9254
1.2	4.1	0.0791	0.0652	2.9443	9.7492
1.2	4.5	0.0727	0.0610	3.1224	10.8231
1.2	5.0	0.0661	0.0564	3.3326	12.1727

4. Parameter Estimation

In this section, we discuss the estimation of NSUGD parameters using Maximum Likelihood (ML), Maximum Product Spacing (MPS), and Cramér-von Mises (CVM) estimation methods. Furthermore, the performance of such estimators will be examined through simulation study.

4.1 The Maximum Likelihood Estimation

Suppose that a random sample of size n is drawn from a population having a NSUGD given by Equation (5) with unknown parameter vector $\Theta = (\alpha, \beta)'$. Then, the likelihood function for is given by

$$\ell(\Theta) = n\,ln\left(\frac{\pi}{4}\right) + n\,ln\,\alpha + n\,ln\,\beta - (\beta+1)\sum_{i=1}^{n} ln\,x_i - \sum_{i=1}^{n}\alpha\left(x_i^{-\beta}-1\right) + \sum_{i=1}^{n} ln(2A_i+1) + \sum_{i=1}^{n} ln\left(cos\left[\frac{\pi}{4}A_i(A_i+1)\right]\right), \tag{15}$$

where $A_i = e^{-\alpha(x_i^{-\beta}-1)}$.

From Equation (15), the partial derivatives with respect to α and β are given by

$$\frac{\partial\ell}{\partial\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n}\left(x_i^{-\beta}-1\right) - \sum_{i=1}^{n}\frac{2\left(x_i^{-\beta}-1\right)A_i}{2A_i+1} + \sum_{i=1}^{n}\frac{\pi}{4}\left(x_i^{-\beta}-1\right)A_i(2A_i+1)\,tan\left(\frac{\pi}{4}A_i(A_i+1)\right), \tag{16}$$

$$\frac{\partial\ell}{\partial\beta} = \frac{n}{\beta} - \sum_{i=1}^{n} ln\,x_i - \sum_{i=1}^{n}\alpha x_i^{-\beta}\,ln\,x_i + \sum_{i=1}^{n}\alpha x_i^{-\beta}\,ln\,x_i + \sum_{i=1}^{n}\frac{2\alpha x_i^{-\beta}\,ln\,x_i\,A_i}{2A_i+1} - \sum_{i=1}^{n}\frac{\pi}{4}\alpha x_i^{-\beta}\,ln\,x_i\,A_i(2A_i+1)\,tan\left(\frac{\pi}{4}A_i(A_i+1)\right), \tag{17}$$

respectively, where $A_i = e^{-\alpha(x_i^{-\beta}-1)}$. The ML estimates of (α,β), say $(\hat{\alpha}_{ML},\hat{\beta}_{ML})$, are obtained by setting $\frac{\partial\ell}{\partial\alpha} = \frac{\partial\ell}{\partial\beta} = 0$ and solving Equations (16) and (17) numerically using statistical software programs such as R.

4.2 *The Maximum Product Spacing Estimation*

Let $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ be the order statistics of a random sample of size n from the NSUGD. The MPS estimator can be determined by minimizing

$$M = \sum_{i=1}^{n+1} ln[D_i], \qquad D_i = F\left(x_{(i)}\right) - F\left(x_{(i-1)}\right),$$

with respect to the unknown parameters. Thus, the MPS estimates of the NSUGD are obtained by solving the following nonlinear equations

$$M(\alpha,\beta) = \sum_{i=1}^{n+1}\frac{1}{D_i}\left[\Delta_s\left(x_{(i)}\right) - \Delta_s\left(x_{(i-1)}\right)\right] = 0, \tag{18}$$

where $\Delta_s\left(x_{(i)}\right), s = 1,2$, are defined as

$$\Delta_1\left(x_{(i)}\right) = \frac{\partial F\left(x_{(i)}\right)}{\partial\alpha} = cos\left[\frac{\pi}{4}u_i(u_i+1)\right]\frac{\pi}{4}(2u_i+1)\left(-u_i\left(x_{(i)}{}^{-\beta}-1\right)\right), \tag{19}$$

$$\Delta_1\left(x_{(i-1)}\right) = \frac{\partial F\left(x_{(i-1)}\right)}{\partial\alpha} = cos\left[\frac{\pi}{4}u_{i-1}(u_{i-1}+1)\right]\frac{\pi}{4}(2u_{i-1}+1)\left(-u_{i-1}\left(x_{(i-1)}{}^{-\beta}-1\right)\right), \tag{20}$$

$$\Delta_2\left(x_{(i)}\right) = \frac{\partial F\left(x_{(i)}\right)}{\partial\beta} = cos\left[\frac{\pi}{4}u_i(u_i+1)\right]\frac{\pi}{4}(2u_i+1)\left(\alpha u_i x_{(i)}{}^{-\beta}\,ln\,x_{(i)}\right), \tag{21}$$

$$\Delta_2\left(x_{(i-1)}\right) = \frac{\partial F\left(x_{(i-1)}\right)}{\partial\beta} = cos\left[\frac{\pi}{4}u_{i-1}(u_{i-1}+1)\right]\frac{\pi}{4}(2u_{i-1}+1)\left(\alpha u_{i-1}x_{(i-1)}{}^{-\beta}\,ln\,x_{(i-1)}\right), \tag{22}$$

given that $u_i = e^{-\alpha\left(x_{(i)}{}^{-\beta}-1\right)}$, and $u_{i-1} = e^{-\alpha\left(x_{(i-1)}{}^{-\beta}-1\right)}$.

4.3 *The Cramér-von Mises Estimation*

The CVM estimator can be determined by minimizing

$$C = \sum_{i=1}^{n}\left[F\left(x_{(i)}\right) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n}.$$

Thus, the MPS estimates of the NSUGD are obtained by solving the following nonlinear equations

$$2\sum_{i=1}^{n}\left[sin\left\{\frac{\pi}{4}u_i(1+u_i)\right\} - \frac{2i-1}{2n}\right]\Delta_s\left(x_{(i)}\right) = 0, \tag{23}$$

where $u_i = e^{-\alpha\left(x_{(i)}{}^{-\beta}-1\right)}$, and $\Delta_s\left(x_{(i)}\right), s = 1,2$, are defined Equations (19) and (21), respectively.

5. SIMULATION

Simulation experiments were conducted to evaluate the performance of the ML, MPS, and CVM estimation methods under varying sample sizes and parameter values for the proposed NSUGD. Random samples from the NSUGD were generated using its quantile function. The simulation considered sample sizes of *n* = 50, 100, 150, 200, 250, 300, 350, 400, 450, and 500, with each configuration replicated 1000

Table 3. Average estimates and MSEs of ML, MPS and CVM for the NSUGD at ($\alpha = 0.8, \beta = 0.5$).

	ML		**MPS**		**CVM**	
n	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)
50	0.7883(0.1170)	0.5276(0.0472)	0.9935(0.3260)	0.4795(0.0189)	0.9690(0.9704)	0.5384(0.0421)
100	0.7929(0.0586)	0.5079(0.0275)	0.8825(0.0949)	0.4883(0.0080)	0.8876(0.2088)	0.5122(0.0197)
150	0.7703(0.0373)	0.5150(0.0202)	0.8377(0.0498)	0.4982(0.0054)	0.8331(0.0998)	0.5156(0.0129)
200	0.7742(0.0296)	0.5132(0.0138)	0.8242(0.0390)	0.4998(0.0044)	0.8174(0.0655)	0.5139(0.0100)
250	0.7888(0.0231)	0.5030(0.0159)	0.8289(0.0305)	0.4973(0.0034)	0.8367(0.0566)	0.5033(0.0080)
300	0.7776(0.0182)	0.5143(0.0050)	0.8168(0.0214)	0.4996(0.0027)	0.8149(0.0427)	0.5085(0.0065)
350	0.7846(0.0172)	0.5026(0.0132)	0.8116(0.0193)	0.4998(0.0024)	0.8180(0.0400)	0.5056(0.0060)
400	0.7827(0.0142)	0.5083(0.0061)	0.8052(0.0153)	0.5020(0.0021)	0.8205(0.0315)	0.5018(0.0049)
450	0.7784(0.0140)	0.5066(0.0096)	0.8058(0.0151)	0.5013(0.0019)	0.8123(0.0268)	0.5036(0.0042)
500	0.7845(0.0112)	0.5052(0.0075)	0.8082(0.0122)	0.5002(0.0016)	0.8177(0.0237)	0.5013(0.0037)

Table 4. Average estimates and MSEs of ML, MPS and CVM for the NSUGD at ($\alpha = 1.2, \beta = 0.8$).

	ML		**MPS**		**CVM**	
n	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)
50	1.3399(1.4835)	0.8794(0.0931)	2.2321(15.3266)	0.6985(0.0957)	2.1421(18.5338)	0.8889(0.2026)
100	1.2477(0.2463)	0.8382(0.0425)	1.5378(0.6284)	0.7301(0.0463)	1.4400(1.2660)	0.8280(0.0837)
150	1.2156(0.1320)	0.8298(0.0270)	1.4044(0.2462)	0.7485(0.0278)	1.3208(0.5529)	0.8324(0.0583)
200	1.2254(0.1178)	0.8197(0.0220)	1.3667(0.1864)	0.7567(0.0225)	1.3040(0.3138)	0.8129(0.0408)
250	1.2123(0.0747)	0.8160(0.0156)	1.3304(0.1159)	0.7622(0.0165)	1.2658(0.1851)	0.8127(0.0314)
300	1.2033(0.0621)	0.8153(0.0125)	1.2998(0.0868)	0.7698(0.0129)	1.2422(0.1460)	0.8163(0.0272)
350	1.2049(0.0544)	0.8128(0.0112)	1.2935(0.0752)	0.7711(0.0116)	1.2533(0.1391)	0.8080(0.0242)
400	1.1941(0.0465)	0.8167(0.0101)	1.2698(0.0601)	0.7802(0.0100)	1.2129(0.0942)	0.8205(0.0200)
450	1.1873(0.0422)	0.8181(0.0091)	1.2560(0.0510)	0.7843(0.0087)	1.2045(0.0843)	0.8213(0.0181)
500	1.2056(0.0370)	0.8078(0.0078)	1.2704(0.0478)	0.7768(0.0082)	1.2349(0.0796)	0.8047(0.0165)

Table 5. Average estimates and MSEs of ML, MPS and CVM for the NSUGD at ($\alpha = 0.5, \beta = 1.2$).

	ML		**MPS**		**CVM**	
n	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)
50	0.5672(0.0498)	1.2116(0.0801)	0.7429(0.1503)	1.0332(0.1026)	0.5713(0.1482)	1.2532(0.1545)
100	0.5472(0.0267)	1.2004(0.0432)	0.6438(0.0569)	1.0910(0.0536)	0.5328(0.0550)	1.2311(0.0748)
150	0.5446(0.0172)	1.1849(0.0277)	0.6149(0.0331)	1.1018(0.0354)	0.5212(0.0311)	1.2203(0.0518)
200	0.5313(0.0122)	1.1938(0.0215)	0.5858(0.0213)	1.1275(0.0259)	0.5084(0.0199)	1.2200(0.0350)
250	0.5277(0.0092)	1.1915(0.0166)	0.5720(0.0156)	1.1372(0.0206)	0.5056(0.0146)	1.2179(0.0278)
300	0.5321(0.0082)	1.1822(0.0137)	0.5714(0.0137)	1.1352(0.0175)	0.5106(0.0126)	1.2074(0.0228)
350	0.5258(0.0065)	1.1877(0.0117)	0.5575(0.0103)	1.1489(0.0142)	0.5061(0.0103)	1.2087(0.0195)
400	0.5223(0.0061)	1.1920(0.0108)	0.5522(0.0093)	1.1545(0.0128)	0.5038(0.0093)	1.2120(0.0174)
450	0.5238(0.0051)	1.1882(0.0093)	0.5513(0.0077)	1.1534(0.0110)	0.5075(0.0088)	1.2056(0.0165)
500	0.5217(0.0046)	1.1899(0.0082)	0.5462(0.0069)	1.1593(0.0098)	0.5036(0.0067)	1.2077(0.0130)

Table 6. Average estimates and MSEs of ML, MPS and CVM for the NSUGD at (α=0.3, β=0.3).

	ML		**MPS**		**CVM**	
n	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)	$\hat{\alpha}$, MSE($\hat{\alpha}$)	$\hat{\beta}$, MSE($\hat{\beta}$)
50	0.3086(0.0159)	0.2957(0.0027)	0.3912(0.0380)	0.2643(0.0039)	0.3261(0.0355)	0.3134(0.0066)
100	0.2966(0.0073)	0.2989(0.0014)	0.3359(0.0107)	0.2808(0.0018)	0.3087(0.0126)	0.3072(0.0029)
150	0.3021(0.0054)	0.2955(0.0010)	0.3307(0.0074)	0.2828(0.0013)	0.3069(0.0086)	0.3048(0.0020)
200	0.3028(0.0039)	0.2959(0.0008)	0.3254(0.0050)	0.2852(0.0010)	0.3074(0.0061)	0.3026(0.0014)
250	0.3000(0.0028)	0.2964(0.0006)	0.3173(0.0033)	0.2882(0.0007)	0.3035(0.0044)	0.3026(0.0011)
300	0.2971(0.0020)	0.2975(0.0005)	0.3140(0.0025)	0.2894(0.0005)	0.2996(0.0034)	0.3036(0.0009)
350	0.3002(0.0018)	0.2967(0.0004)	0.3152(0.0023)	0.2897(0.0005)	0.3036(0.0031)	0.3014(0.0008)
400	0.2966(0.0018)	0.2985(0.0004)	0.3090(0.0021)	0.2926(0.0004)	0.2987(0.0027)	0.3039(0.0007)
450	0.2993(0.0015)	0.2971(0.0003)	0.3099(0.0017)	0.2919(0.0004)	0.3014(0.0023)	0.3016(0.0006)
500	0.2997(0.0015)	0.2973(0.0003)	0.3104(0.0016)	0.2921(0.0004)	0.3024(0.0024)	0.3013(0.0006)

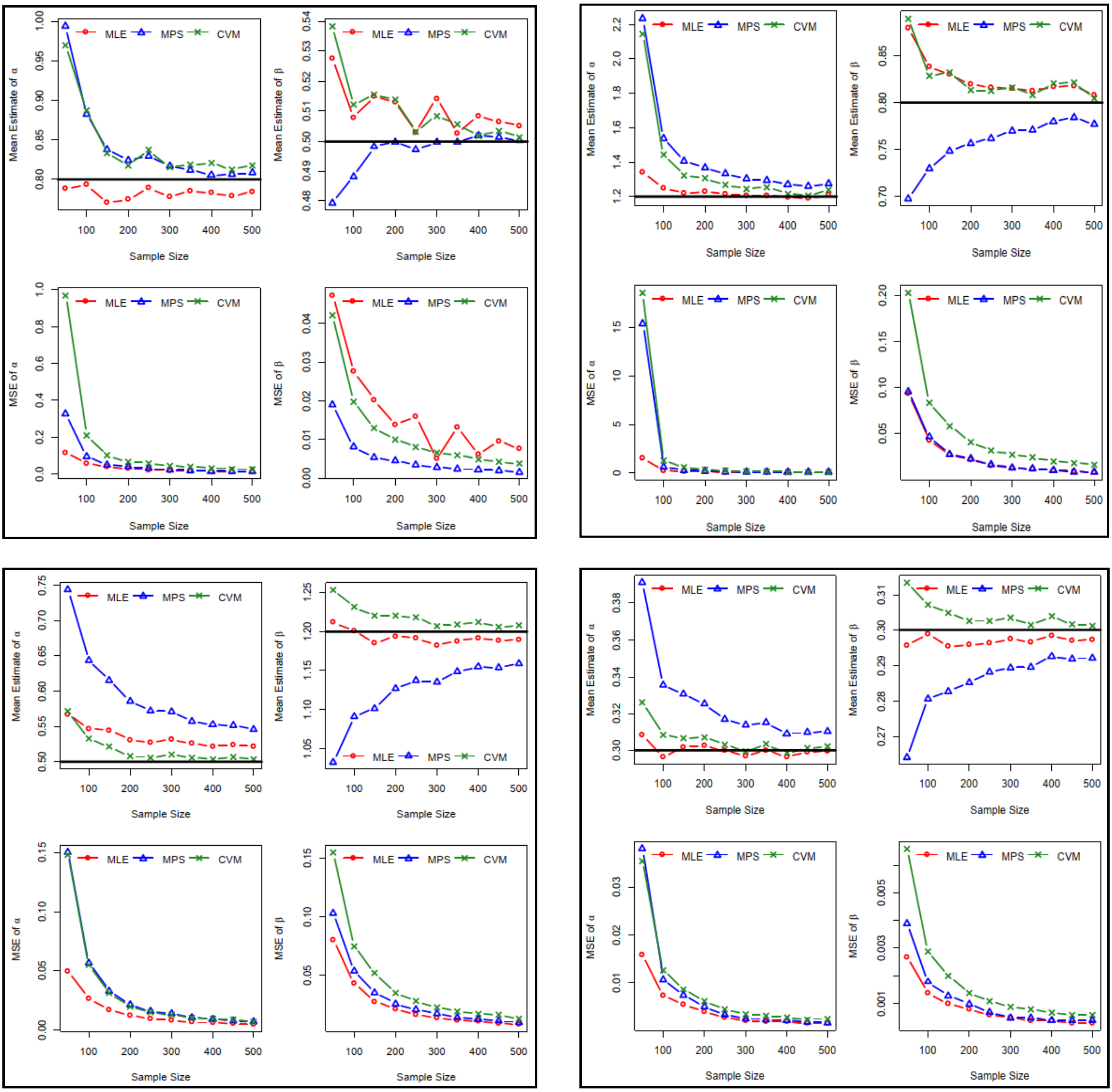

Fig. 3. Graphical representation of mean estimate and MSE values in Table 3 (top-left), Table 4 (top-right), Table 5 (bottom-left), and Table 6 (bottom-right) respectively.

times. The R Statistical software was used to compute the mean estimates and mean squared errors (MSE), which are given as

$$\mathrm{Mean}(\Theta) = \frac{1}{1000}\sum_{i=1}^{1000}\hat{\Theta}_i,$$

$$\mathrm{MSE}(\Theta) = \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{\Theta}_i - \Theta\right)^2,$$

where $\hat{\Theta} = (\hat{\alpha}, \hat{\beta})$ is an estimate of $\Theta = (\alpha, \beta)$.

The mean estimates and their corresponding MSEs for each parameter are presented in Tables 3 - 6, while Figure 3 provides the graphical representations that reflect the numerical outcomes of the simulations. At the smallest n = 50, the MSEs are noticeably higher for all three estimation methods. This is a common feature of finite-sample estimation in two-parameter models and reflects greater sampling variability in small samples [10]. As the *n* increases, the errors decline steadily, supporting the consistency of all estimators.

Furthermore, under standard regularity conditions satisfied by the NSUGD, the ML estimators are asymptotically unbiased, consistent, and efficient in the sense of attaining the Cramér-Rao lower bound: the minimum achievable variance among all consistent estimators [11]. MPS estimators, while consistent, are based on geometric mean spacings and do not generally attain this bound, whereas CVM estimators, being minimum-distance methods, sacrifice efficiency when the parametric model is correctly specified [12]. These theoretical rankings are confirmed empirically: across all simulation scenarios in Tables 3 - 6, ML records the smallest MSEs for both parameters at nearly every sample size, converging most rapidly to the true values as *n* increases. Consequently, it emerges as the most effective approach.

6. REAL APPLICATIONS

To demonstrate the flexibility of the proposed model, its performance is compared against several established unit-bounded distributions, namely the Unit Gompertz (UGD) by [3], Unit Weibull (UWD) by [13], Beta by [14], and Kumaraswamy by [15]. Their respective PDFs are given as

$$f_{UGD}(x;\alpha,\beta) = \alpha\beta x^{-\beta-1} e^{-\alpha(x^{-\beta}-1)},$$

$$f_{UWD}(x;\alpha,\beta) = \frac{\alpha\beta}{x}(-\ln x)^{\beta-1} \exp\left[-\alpha(-\ln x)^{\beta}\right],$$

$$f_{\text{Beta}}(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1},$$

$$f_{\text{Kum}}(x;\alpha,\beta) = \alpha\beta x^{\alpha-1}(1-x^{\alpha})^{\beta-1}.$$

Model selection for real-life datasets is based on a comprehensive set of analytical criteria, including the log-likelihood $(\hat{\ell})$, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). Furthermore, goodness-of-fit (g-o-f) statistic, which is, Kolmogorov-Smirnov (KS) with its associated *p*-value [KS (*p*-value)] are also considered.

Dataset 1: Trade Share Data - The first dataset comprises of 61 observations on trade share data sourced from [16] and previously used by [17] represent the ratio of a country's total trade (exports + imports) to its GDP; a standard measure of economic openness that is bounded strictly within (0, 1), such data naturally satisfy the assumptions of unit-bounded distributions. The measurements of the dataset are as follows:

0.140501976, 0.156622976, 0.157703221, 0.160405084, 0.160815045, 0.22145839, 0.299405932, 0.31307286, 0.324612707, 0.324745566, 0.329479247, 0.330021679, 0.337879002, 0.339706242, 0.352317631, 0.358856708, 0.393250912, 0.41760394, 0.425837249, 0.43557933, 0.442142904, 0.444374621, 0.450546652, 0.4557693, 0.46834656, 0.473254889, 0.484600782, 0.488949597, 0.509590268, 0.517664552, 0.527773321, 0.534684658, 0.543337107, 0.544243515, 0.550812602, 0.552722335, 0.56064254, 0.56074965, 0.567130983, 0.575274825, 0.582814276, 0.603035331, 0.605031252, 0.613616884, 0.626079738, 0.639484167, 0.646913528, 0.651203632, 0.681555152, 0.699432909, 0.704819918, 0.729232311, 0.742971599, 0.745497823, 0.779847085, 0.798375845, 0.814710021, 0.822956383, 0.830238342, 0.834204197, 0.979355395.

The descriptive statistics (mean $= 0.5142$, median $= 0.5278$, mode $= 0.1405$, skewness $= 0.0060$, kurtosis $= 2.5526$) reveal a near-symmetric, platykurtic distribution. Such unit-bounded proportions are common in econometric studies, where accurate modelling of central tendency and tails is essential. The NSUGD is therefore well placed to describe this type of financial data.

Dataset 2: Relief Times Data for Arthritic Patients - The relief times dataset originates from clinical trials measuring the duration until patients experience relief after treatment. This dataset records hours-to-pain-relief for 50 arthritic patients following analgesic administration, as originally reported by [18] and applied by [19]. The data are as follows:

0.70, 0.84, 0.58, 0.50, 0.55, 0.82, 0.59, 0.71, 0.72, 0.61, 0.62, 0.49, 0.54, 0.36, 0.36, 0.71, 0.35, 0.64, 0.85, 0.55, 0.59, 0.29, 0.75, 0.46, 0.46, 0.60, 0.60, 0.36, 0.52, 0.68, 0.80, 0.55, 0.84, 0.34, 0.34, 0.70, 0.49, 0.56, 0.71, 0.61, 0.57, 0.73, 0.75, 0.44, 0.44, 0.81, 0.80, 0.87, 0.29, 0.50.

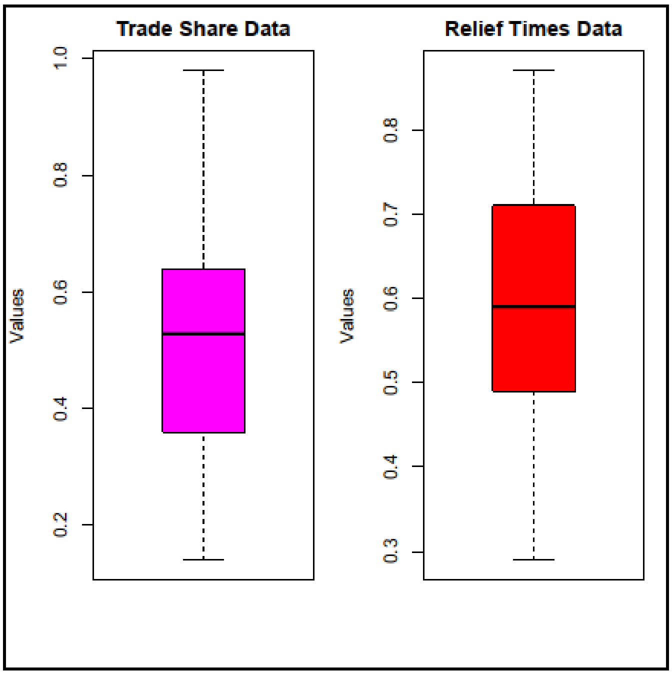

Fig. 4. Box plots for the datasets.

The descriptive analysis shows that mean $= 0.5908$, median $= 0.59$, mode $= 0.55$, skewness $= -0.0976$, and platykurtic kurtosis $= 2.0899$. The observed left-skewness reflects the tendency for many patients to respond relatively quickly, while a smaller number experience delayed relief due to differences in physiological response and disease severity [20]. This pattern is common in medical data and supports the use of flexible bounded models such as the NSUGD.

Figure 4 visually represents the two datasets through box plots. The plots reveal that the first and second data are right-skewed and left-skewed respectively. Also, the data are both platykurtic indicating the absence of outliers.

Although the trade share dataset exhibits a platykurtic behaviour, many financial proportion datasets such as loss ratios, recovery rates, and market concentration indices frequently exhibit heavier tails, reflecting occasional extreme events than the beta or Kumaraswamy distributions can capture [21]. The NSUGD addresses this through the joint influence of α and β: increasing both parameters concentrates mass asymmetrically and elongates the right tail, producing higher skewness and kurtosis, as demonstrated in Table 2. This flexibility ensures the model remains applicable across a wide

Table 7. The MLEs, SEs (in parentheses), selection criteria and g-o-f statistic for the trade share data.

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\ell}$	**AIC**	**BIC**	**KS (*p*-value)**
NSUGD	1.1716 (0.5662)	0.7597 (0.2512)	**-14.4712**	**-24.9423**	**-20.7206**	**0.0622 (0.9608)**
UGD	0.6163 (0.2665)	1.0922 (0.2475)	-10.8759	-17.7518	-13.5300	0.1098 (0.4232)
UWD	1.3397 (0.1726)	1.7342 (0.1695)	-14.2436	-24.4872	-20.2654	0.0681 (0.9215)
Beta	2.7950 (0.4882)	2.6001 (0.4512)	-13.9559	-23.9118	-19.6900	0.0626 (0.9586)
Kumaraswamy	2.3237 (0.3052)	2.7574 (0.5544)	-13.6249	-23.2498	-19.0281	0.0683 (0.9196)

Table 8. The MLEs, SEs (in parentheses), selection criteria and g-o-f statistic for relief times data.

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\ell}$	**AIC**	**BIC**	**KS (*p*-value)**
NSUGD	0.5672 (0.2304)	1.6204 (0.3773)	**-22.2123**	**-40.4245**	**-36.6005**	**0.0826 (0.8850)**
UGD	0.2743 (0.1254)	2.2722 (0.4333)	-19.2387	-34.4774	-30.6534	0.0971 (0.7340)
UWD	2.4840 (0.2261)	2.0541 (0.2261)	-22.0385	-40.0770	-36.2530	0.0854 (0.8590)
Kumaraswamy	3.7479 (0.5155)	4.2126 (1.0391)	-21.9991	-39.9982	-36.1741	0.0874 (0.8394)

range of empirical tail behaviours in unit-bounded financial data.

The superior fit achieved by this model, as given in Tables 7 and 8, is therefore entirely consistent with the empirical evidence revealed in Figures 5 and 6. As presented in Tables 7 and 8, the analytical measures, together with the ML estimates, their associated standard errors (SEs), selection criteria, and g-o-f test statistic are reported for all models considered for both datasets.

For the trade share data, the NSUGD ($AIC = -24.9423, BIC = -20.7206$) outperforms the UWD ($AIC = -24.4872, BIC = -20.2654$) by $0.455\, AIC$ units, and substantially surpasses the UGD by $7.19\, AIC$ units. The NSUGD also records the highest KS *p*-value (0.9608) versus

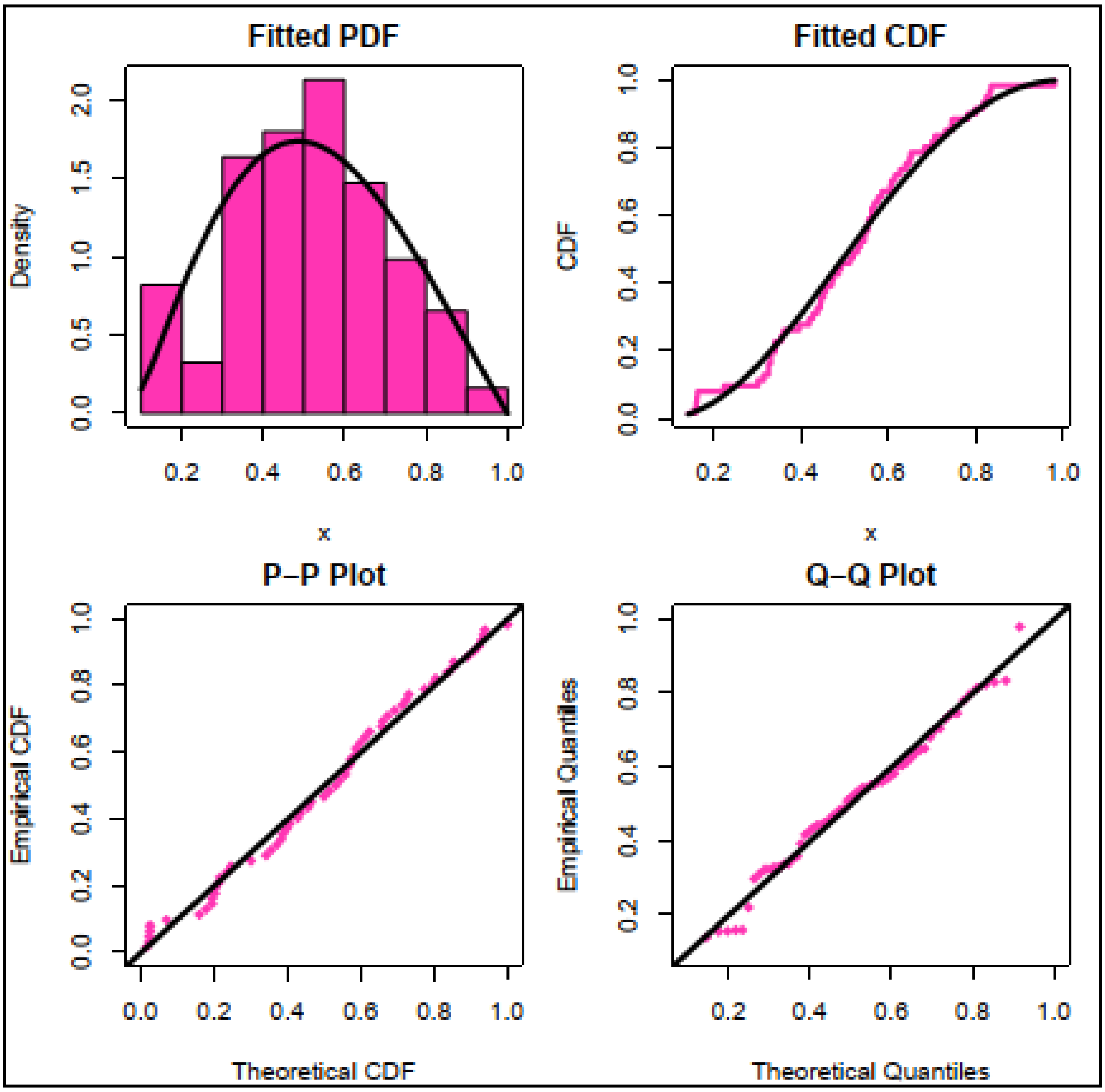

Fig. 5. The histogram with the fitted PDF, fitted CDF, P-P and Q-Q plots of Trade Share Data.

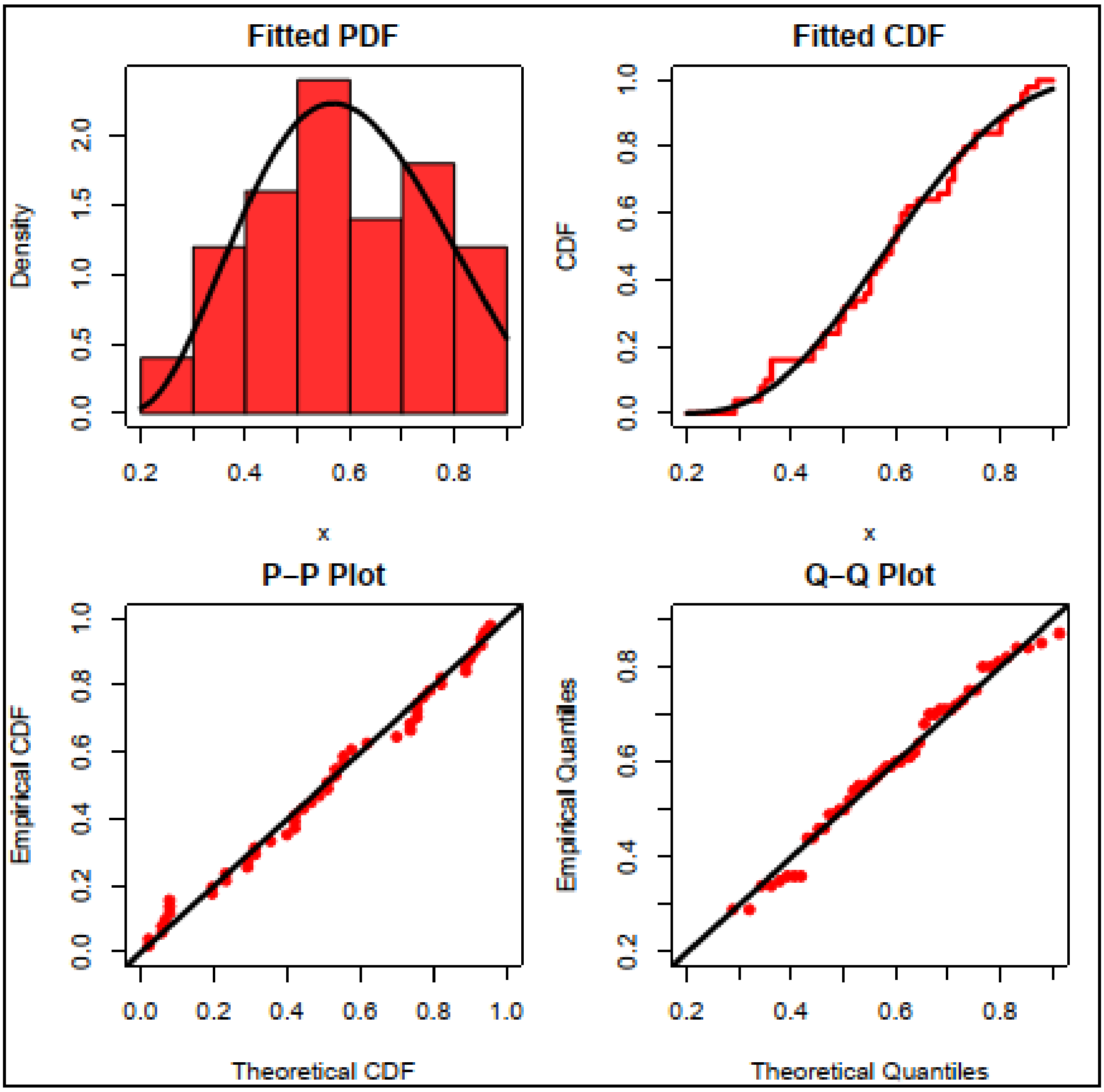

Fig. 6. The histogram with the fitted PDF, fitted CDF, P-P and Q-Q plots of Relief Times Data.

the UWD (0.9215). For the relief times data, the NSUGD ($AIC = -40.4245, BIC = -36.6005$) again leads the UWD ($AIC = -40.0770, BIC = -36.2530$) by 0.347 AIC units and the UGD by 5.95 AIC units, while achieving the highest KS *p*-value (0.8850) across all competing models.

These consistent improvements across both datasets confirm the NSUGD as the superior fitting distribution. In addition, the fitted PDF, CDF, Probability-Probability (P-P) and the Quantile-Quantile (Q-Q) plots, illustrated in Figures 5 and 6, further confirm the adequacy of the NSUGD in modelling the data.

7. CONCLUSION

This study introduced a new bounded lifetime distribution, termed the New Sine Unit Gompertz Distribution (NSUGD), as a modification of the Unit Gompertz distribution. The statistical and reliability characteristics of the NSUGD were extensively explored. Its probability density function exhibits considerable flexibility, accommodating decreasing, nearly symmetric, right-skewed, and left-skewed shapes, while the hazard rate function can display increasing, J-shaped, or bathtub patterns. Such versatility renders the model suitable for analysing diverse types of data defined on the unit interval. Several key properties were derived, including the survival and hazard functions, quartile function, moments, median, mean, skewness, and kurtosis. The model parameters were estimated using three classical methods: Maximum Likelihood, Maximum Product of Spacings, and Cramér-von Mises estimations. A Monte Carlo simulation study was performed to assess the performance of these estimators based on their mean estimates and mean squared errors, revealing that the Maximum Likelihood method provided the most efficient estimates. The practical usefulness of the NSUGD was demonstrated through its application to financial and medical datasets, where it exhibited superior fitting performance compared to several competing unit-bounded models. Owing to its adaptability, the NSUGD holds promise for diverse fields, including engineering, economics, environmental science, insurance, and other disciplines where unit bounded data are prevalent.

REFERENCES

[1] M. C. Jones, “The complementary beta distribution,” J. Statistical Planning and Inference, vol. 104, no. 2, pp. 329-337, 2002. doi: https://doi.org/10.1016/S0378-3758(01)00260-9

[2] G. M. Cordeiro and M. de Castro, “A new family of generalized distributions,” J. Statistical Computation and Simulation, vol. 81, no. 7, pp. 833-898, 2011. doi: https://doi.org/10.1080/00949655.2010.5030745

[3] J. Mazucheli, A. F. B. Menezes, and S. Dey, “Unit-Gompertz distribution with applications,” Statistica, vol. 79, no. 1, pp. 25–43, 2019. doi: https://doi.org/10.6092/issn.1973-2201/8497

[4] B. Gompertz, “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,” Philos. Trans. R. Soc. London, vol. 115, pp. 513-583, 1825. doi: https://doi.org/10.1098/rstl.1825.0026

[5] Z. Mahmood, C. Chesneau, and M. H. Tahir, "A new sine-G family of distributions: Properties and applications," Bull. Comput. Appl. Math., vol. 7, no. 1, pp. 53–81, 2019.

[6] C. D. Lai, and M. Xie, Stochastic ageing and dependence for reliability. Springer, 2006, ISBN: 978-0-387-29742-2.

[7] W. Q. Meeker and L. A. Escobar, Statistical Methods for Reliability Data. New York: Wiley, 1998, ISBN: 978-0471143284.

[8] J. E. Muller, P. H. Stone, Z. G. Turi, J. D. Rutherford, C. A. Czeisler, C. Parker, W. K. Poole, E. Passamani, R. Roberts, T. Robertson, and B. E. Sobel, "Circadian variation in the frequency of onset of acute myocardial infarction," New England J. Medicine, vol. 313, no. 21, pp. 1315–1322, 1985. doi: https://doi.org/10.1056/NEJM198511213132103

[9] H. Pham and H. Wang, "Imperfect maintenance," European J. Operational Research, vol. 94, no. 3, pp. 425–438, 1996. doi: https://doi.org/10.1016/S0377-2217(96)00099-9

[10] E. L. Lehmann and G. Casella, Theory of Point Estimation, 2nd ed. New York: Springer, 1998, ISBN: 978-0387985022.

[11] G. Casella and R. L. Berger, Statistical Inference, 2nd ed. Pacific Grove, CA: Duxbury Press, 2002, ISBN: 978-0534243128.

[12] R. C. H. Cheng and N. A. K. Amin, "Estimating parameters in continuous univariate distributions with a shifted origin," J. Royal Statistical Society: Series B, vol. 45, no. 3, pp. 394–403, 1983. doi: https://doi.org/10.1111/j.2517-6161.1983.tb01268.x

[13] J. Mazucheli, A. F. B. Menezes, and M. E. Ghitany, “The unit-Weibull distribution and associated inference,” J. Applied Probability and Statistics, vol. 13, no. 2, pp. 1–22, 2018.

[14] T. Bayes, “An Essay Towards Solving a Problem in the Doctrine of Chances,” Philos. Trans. R. Soc. London, vol. 53, pp. 370-418, 1763. doi: https://doi.org/10.1098/rstl.1763.0053

[15] P. Kumaraswamy, “A generalized probability density function for double-bounded random processes,” J. Hydrology, vol. 46, no. 1-2, pp. 79-88, 1980. doi: https://doi.org/10.1016/0022-1694(80)90036-0

[16] J. H. Stock and M. W. Watson, Introduction to Econometrics, 2nd ed. Boston, MA: Addison-Wesley, 2007, ISBN: 978-0321278621.

[17] J. C. Ehiwario, J. N. Igabari, and P. E. Ezimadu, “The alpha power Topp-Leone distribution: Properties, simulations and applications,” J. Applied Mathematics and Physics, vol. 11, no. 1, pp. 316–331, 2023. doi: https://doi.org/10.4236/jamp.2023.111018

[18] D. R. Wingo, "Maximum likelihood methods for fitting the Burr XII distribution to life test data," Biometrical Journal, vol. 25, no. 1, pp. 77–84, 1983. doi: https://doi.org/10.1002/bimj.19830250109

[19] F. G. Akgül, "Comparison of the estimation methods for the parameters of exponentiated reduced Kies distribution," Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, vol. 22, no. 3, pp. 1209–1216, 2018. doi: https://doi.org/10.19113/sdufenbed.498870

[20] A. Bruno, S. Tacconelli, and P. Patrignani, “Variability in the response to non-steroidal anti-inflammatory drugs: mechanisms and perspectives,” Basic & Clinical Pharmacology & Toxicology, vol. 114, no. 1, pp. 56-63, 2014. doi: https://doi.org/10.1111/bcpt.12117

[21] R. Cont, "Empirical properties of asset returns: Stylized facts and statistical issues," Quantitative Finance, vol. 1, no. 2, pp. 223–236, 2001. doi: https://doi.org/10.1080/713665670