

A New Unit Family of Distributions with Applications to Public Health, Environmental, and Financial Data

Okechukwu J. Obulezi¹, Emmanuel E. Oguadimma², Chinyere P. Igboke³, Chinelo U. Chikwelu⁴, Gaber Sallam Salem Abdalla⁵, John T. Mendy^{6,*}, Ahmed S. Abdelzahr⁷, and Mohammed Elgarhy⁸

¹Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, P.O. Box 5025 Awka, Nigeria

²Department of Mathematics, Oregon State University, Corvallis, OR 97331, USA

³Department of Statistics, School of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, 144411, India

⁴Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, P.O. Box 5025 Awka, Nigeria

⁵Department of Insurance and Risk Management, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia

⁶Department of Mathematics, School of Arts and Science, Brikama Campus, University of The Gambia, Brikama, Gambia

⁷Financial Sciences department, The Applied College, Imam Abdulrahman Bin Faisal University, Saudi Arabia

⁸Department of Basic Sciences, Higher Institute of Administrative Sciences, Belbeis, AlSharkia, Egypt

Received: 7 Jun. 2025, Revised: 21 Sep. 2025, Accepted: 13 Oct. 2025

Published online: 1 Nov. 2025

Abstract: Bounded distributions are important in statistical studies especially when investigations involve rates, proportions or experimental outcomes such as success/failure, defective/non-defectives etc. The literature on bounded or unit family of distributions is scanty. Therefore, in this study, we develop a new unit family of distributions called Dhillon-G family of distributions. The Dhillon-G family is parsimonious given that it does not introduce additional parameter. Further, the Kumaraswamy distribution was modified using the Dhillon-G family with the modified distribution offering a better flexibility in application, mathematical tractability and better goodness of fit. We studied the structural behaviour/properties of the new distribution and the studies reveal that the distribution can model both skewed, unimodal, bimodal, bath-tub shape and bounded data. We estimated its parameters using non-Bayesian and Bayesian methods and a comprehensive simulation study is implemented to determine the behaviour of the estimators using maximum likelihood, maximum product of spacing, least squares and weighted least squares at different parameter settings and sample sizes. The results reveal consistent asymptotic behaviour for the parameters in the presence of small and large sample sizes. The Dhillon-Kumaraswamy distribution and its chosen competing distributions which included (Kumaraswamy, Beta, Gumbel, Gamma, Burr XII and Weibull) were fitted to data on mortality of COVID-19 patients in Dominica, Vinyl chloride from clean upgradient ground-water monitoring wells in (g/L), weekly volume of traded Bitcoin in USD, and infant mortality rate per 1000 live-births in selected countries. All four datasets are positively skewed and the Dhillon-Kumaraswamy distribution outperformed all competing models in fitting those datasets and in parameter estimation. These demonstrate that the new distribution is useful in modeling short interval dataset namely $x \in (0, 1)$ as well as asymmetric data.

Keywords: Dhillon distribution, Dhillon-G family, Estimation, Kumaraswamy distribution, Parsimony

1 Introduction

From economics to environmental science, researchers frequently encounter variables that are naturally constrained within a specific range, typically $(0, 1)$. Whether it's calculating a success rate or understanding population proportions, analyzing this double-bounded data is critical. The primary reason for utilizing the unit

interval to launch a new family of distributions is to obtain greater flexibility and control of the ensuing statistical model. This is beneficial as it allows one to map existing distributions, which may have limited supports like the real line or non-negative real line, to a standardized bounded interval. These render them suitable for modeling data that are naturally bounded,

* Corresponding author e-mail: jt.mendy83@gmail.com

e.g., proportions, percentages, or rates. Working on the unit interval, new distributions with a wider variety of shapes can be devised, including complex features like bimodal or J-shaped curves, which were not accessible from the original parent distribution. This is typically accomplished by mapping the cumulative distribution function or a similar function of a parent distribution with the desired known properties to the unit interval, thereby introducing additional shape parameters. Not only does this method enhance the distribution's flexibility, but also links the distribution to an existing theoretical basis, e.g., the Beta distribution, so that the new family of distributions is more interpretable and handy for modeling a variety of real data more accurately.

However, using common statistical models such as the normal or exponential distributions presents a significant challenge. Their unbounded nature can lead to estimates that fall outside the physically possible range. To overcome this, a class of specialized probability models has emerged. The Beta distribution is a seminal example, offering a robust and generic framework for modeling data within the $(0, 1)$ interval, ([10]). This was followed by other notable models, such as the Kumaraswamy ([12]) and Johnson S_B distributions ([11]), logit-normal distribution ([22]), Beta-Cauchy distribution ([23]), Dagum distribution ([24]), generalized Beta distribution of the first kind ([25]), inverse-Weibull distribution ([26]), Mustapha ([16]), unit-Weibull ([27]), unit log-log ([28]), unit-inverse Gaussian ([29]), unit-Gompertz ([30]), power new power function ([31]), unit Teissier ([32]), unit-Birnbaum-Saunders ([33]), Gumbel-Logistic Unit ([34]), unit-Chen ([35]), unit Burr-XII ([36]), unit-Lindley ([37]), each bringing its own strengths to the field. Distributions have helped in understanding human activities and therefore present models that aid decision making, especially in health, environment, engineering, and finance. These include reduced type-I heavy-tailed Weibull ([40]), new family of generalized distributions based on logistic-x transformation ([50]), two-parameter Chris-Jerry distribution ([51]), type-I heavy-tailed exponential distribution ([52]), type-I heavy-tailed Rayleigh distribution ([53]), type-II heavy-tailed family ([54]), kumaraswamy bell-Rayleigh distribution ([55]) and Weibull sine generalized family ([56]), new truncated Lindley-generated family ([57]), bounded sine hyperbolic distribution ([58]), new flexible power-X family ([59]), new Lomax-G family ([60]) and new hyperbolic tangent family ([61]). For more details see [41, 42, 43, 44, 46, 47, 48, 49].

The complexity of emerging data, often characterized by highly skewed, heavy-tailed, multimodal patterns, or $(0, 1)$ interval exceeds the capabilities of existing distributions. This clear gap in the literature drives the search for more flexible and powerful models. In this paper, we introduce a new family of double bounded continuous distributions with the view to providing more flexibility for the description of proportions and related data. The family is parameter parsimonious (no additional

parameter to the baseline model), better goodness of fit to varying datasets and mathematically tractable.

This manuscript is organized into a logical flow, beginning with the Dhillon-G family of distributions (Section 2) and then focusing on a specific member, the Dhillon-Kumaraswamy distribution (Section 3). The article then delves into the distribution's structural properties (Section 4) before moving on to parameter estimation (Section 5) and Monte Carlo simulation (Section 6). These foundational sections lead to the practical applications of the distribution (Section 7). The manuscript concludes with final comments in (Section 8).

2 Dhillon-G family of distributions

[1] proposed a distribution of a non-negative random variable T , that is capable of describing decreasing and inverted bathtub hazard rates. For a special condition when $\theta = \beta = 1$, a reduced Dhillon distribution will be a probability density function (pdf) and cumulative distribution function (cdf) given as

$$r(t) = \frac{1}{(t+1)^2}; \quad t > 0, \quad (1)$$

and

$$R(t) = 1 - e^{-\ln(t+1)} = 1 - \frac{1}{t+1}, \quad (2)$$

respectively. Following the definition provided by [2], let $r(t)$ be the PDF of a random variable $T \in [a, b]$, for $-\infty \leq a < b \leq \infty$ and $W[G(x; \zeta)]$ be a function of the CDF $G(x; \zeta)$ of any random variable X so that $W[FG]$ satisfies the following conditions;

1. $W[G(x; \zeta)] \in [a, b]$.
2. W is differentiable and monotonically nondecreasing.
3. $W[G(x; \zeta)] \rightarrow a$ as $x \rightarrow -\infty$ and $W[G(x; \zeta)] \rightarrow b$ as $x \rightarrow \infty$,

where $[a, b]$ is the domain of the random variable T such that $-\infty \leq a < b \leq \infty$.

Therefore, utilizing the $T - X$ generator, the cdf of the family of distributions is

$$F(x; \zeta) = \int_a^{W[G(x; \zeta)]} r(t) dt = R\{W[G(x; \zeta)]\}. \quad (3)$$

The associated pdf is

$$f(x; \zeta) = \frac{dW[G(x; \zeta)]}{dx} r\{W[G(x; \zeta)]\}; \quad x > a, \quad (4)$$

where ζ is the vector of parameters for any parent distribution with cdf $G(\cdot)$. Next, we provide some choices

Table 1: $W[G(x; \zeta)]$ with associated domains, CDF and PDF of the generated family

Range of x	$W[G(x)]$	$F(x)$	$f(x)$	Remark
$x \in \text{support of } G$	$G(x)$	$1 - \frac{1}{G(x)+1}$	$\frac{g(x)}{(G(x)+1)^2}$	canonical case
$x \in \text{support of } G$	$1 - G(x)$	$1 - \frac{1}{2-G(x)}$	$\frac{g(x)}{(2-G(x))^2}$	tail-reverse or increasing tail weights
$x : G(x) \in (0, 1)$	$-\log(1 - G(x))$	$1 - \frac{1}{1-\log(1-G(x))}$	$\frac{g(x)}{(1-\log(1-G(x)))^2(1-G(x))}$	exponentiated-related families
$x : G(x) \in (0, 1)$	$-\log(G(x))$	$1 - \frac{1}{1-\log(G(x))}$	$\frac{g(x)}{(1-\log(G(x)))^2G(x)}$	reversed hazard frameworks
$x \in \text{support of } G \setminus \{0, 1\}$	$\frac{G(x)}{1-G(x)}$	$G(x)$	$g(x)$	odd transformation
$x \in \text{support of } G \setminus \{0, 1\}$	$\left(\frac{G(x)}{1-G(x)}\right)^\delta$	$1 - \frac{1}{1+\left(\frac{G(x)}{1-G(x)}\right)^\delta}$	$\frac{\delta g(x) G^{\delta-1}(x)}{(1-G(x))^{\delta+1} \left(1 + \left(\frac{G(x)}{1-G(x)}\right)^\delta\right)^2}$	generalized odds transformation
$x \in \text{support of } G$	$G^k(x)$	$1 - \frac{1}{1+G^k(x)}$	$\frac{k G^{k-1}(x) g(x)}{(1+G^k(x))^2}$	shape-controlling transformation power
$x \in \text{support of } G$	$\log(1 + G(x))$	$1 - \frac{1}{1+\log(1+G(x))}$	$\frac{g(x)}{(1+\log(1+G(x)))^2(1+G(x))}$	light-tail modeling
$x \in \text{support of } G$	$\exp(G(x)) - 1$	$1 - \frac{1}{\exp(G(x))}$	$\frac{g(x)}{\exp(G(x))}$	stretches tail moderately
$x \in \text{support of } G$	$\tan\left(\frac{\pi}{2} G(x)\right)$	$1 - \frac{1}{1+\tan\left(\frac{\pi}{2} G(x)\right)}$	$\frac{\pi g(x)}{2(1+\tan\left(\frac{\pi}{2} G(x)\right))^2 \cos^2\left(\frac{\pi}{2} G(x)\right)}$	maps $G(x) \in (0, 1)$ to $t \in (0, \infty)$
$x \in \text{support of } G \setminus \{0\}$	$\frac{1}{G(x)} - 1$	$G(x)$	$g(x)$	inverse function to slow convergence
$x \in \text{support of } G \setminus \{1\}$	$\frac{1}{1-G(x)} - 1$	$G(x)$	$g(x)$	inverse tail probability
$x \in \text{support of } G$	$\arctan(G(x))$	$1 - \frac{1}{1+\arctan(G(x))}$	$\frac{g(x)}{(1+\arctan(G(x)))^2(1+G(x)^2)}$	periodic functions
$x : G(x) \in (0, 1)$	$\log\left(\frac{1}{1-G(x)}\right)$	$1 - \frac{1}{1+\log\left(\frac{1}{1-G(x)}\right)}$	$\frac{g(x)}{\left(1+\log\left(\frac{1}{1-G(x)}\right)\right)^2(1-G(x))}$	log-survival transformation

of $W[G(x; \zeta)]$ with their resulting family CDF and PDF in Table (1).

From Table (1), our interest is on bounded random variable X , basically utilizing $W[G(x)] = \log\left(\frac{1}{1-G(x)}\right)$, in equations (3) and (4), the cdf of the Dhillon-G family is

$$F(x) = 1 - \frac{1}{1 + \log\left(\frac{1}{1-G(x)}\right)}; \quad x \in (0, 1). \quad (5)$$

The corresponding pdf is

$$f(x) = \frac{g(x)}{\left[1 + \log\left(\frac{1}{1-G(x)}\right)\right]^2 (1-G(x))}. \quad (6)$$

The hazard function which is generally given as $h(x) = \frac{f(x)}{1-F(x)}$ is

$$h(x) = \frac{g(x)}{\left[1 + \log\left(\frac{1}{1-G(x)}\right)\right] (1-G(x))}. \quad (7)$$

A significant disadvantage of using the unit interval to propose new families of distributions is the lack of intuitive interpretability of the new parameters. Unlike simple distributions for which the parameters can clearly represent location or scale, the parameters of such newly transformed distributions can often represent a complex combination of shape, skewness, and kurtosis that can be hard for practitioners to understand and deal with. This lack of interpretability is compounded by the fact that such models are computationally intensive; the PDF and CDF functions generally have no closed-form

representation, making parameter estimation by maximum likelihood estimation numerically unstable and computationally intensive. Further, the method can lead to over-parameterization, where model complexity outstrips the data, with the risk of a bad generalization to new data and the model not being reliable for prediction. So, offering more flexibility, this method trades simplicity and ease of calculation for this flexibility.

3 Dhillon-Kumaraswamy distribution

The Kumaraswamy distribution by [12], originally referred to as the double-bounded distribution is one of the earliest truncated distributions for modeling rates, percentages, ratios and related lifetime occurrences which the standard models failed to handle. Due to its structural properties and computational efficiency, it has attracted much attention among researchers. In this paper, the kumaraswamy distribution is used as a baseline, with $W[G(x)] = \log\left(\frac{1}{1-G(x)}\right)$ to formulate what is now known as the Dhillon-Kumaraswamy (DK) distribution. The resulting distribution is also double-bounded at $(0, 1)$. Let X be a continuous random variable in the interval $0 < x < 1$, the pdf and cdf of the Kumaraswamy distribution are respectively

$$g(x; a, b) = abx^{a-1} (1-x)^{b-1}; \quad 0 < x < 1, \quad a, b > 0, \quad (8)$$

and

$$G(x; a, b) = 1 - (1-x^a)^b. \quad (9)$$

Substituting equation (9) into (5), the cdf of the DK distribution is realized as

$$F(x; a, b) = 1 - \frac{1}{1 - b \log(1 - x^a)}; \quad x \in (0, 1), \quad a, b > 0. \quad (10)$$

Similarly, plug in equation (8) and (9) into (6), the corresponding pdf of the DK distribution is obtained as

$$f(x; a, b) = \frac{abx^{a-1}}{(1 - x^a)[1 - b \log(1 - x^a)]^2}; \quad x \in (0, 1), \quad a, b > 0. \quad (11)$$

Remark. The advantage of this DK model is that it does not include any additional parameters however alters the functional form of the baseline distribution (Kumaraswamy). Skewness depends on a : If $a < 1$, peak near 0 — right-skewed. If $a = 1$, shape depends mainly on b . If $a > 1$, peak shifts toward middle or right — possibly left-skewed.

The hazard rate function or force of mortality is expressed as

$$h(x; a, b) = \frac{abx^{a-1}}{(1 - x^a)[1 - b \log(1 - x^a)]}, \quad x \in (0, 1), \quad a, b > 0. \quad (12)$$

3.1 Unimodality of the DK distribution

To ascertain the unimodality conditions, we must analyze the first derivative of the pdf with respect to x , $f'(x)$, and identify its critical points by setting $f'(x) = 0$. Using logarithmic differentiation, we consider $L(x) = \ln f(x; a, b)$:

$$L(x) = \ln(ab) + (a-1) \ln x - \ln(1 - x^a) - 2 \ln[1 - b \ln(1 - x^a)]$$

Differentiating $L(x)$ with respect to x yields:

$$\frac{f'(x)}{f(x)} = \frac{a-1}{x} + \frac{ax^{a-1}}{1-x^a} - \frac{2abx^{a-1}}{(1-x^a)[1-b \ln(1-x^a)]}$$

Setting $\frac{f'(x)}{f(x)} = 0$ leads to the following equation for critical points:

$$(a-1)(1-x^a)[1-b \ln(1-x^a)] + ax^a[1-b \ln(1-x^a)] - 2abx^a = 0 \quad (13)$$

This is a complex implicit equation involving x^a and $\ln(1-x^a)$. Obtaining an explicit analytical condition for unimodality (i.e., a simple closed-form expression for a and b defining the unimodal region) from this equation is generally not feasible. Understanding the behavior of the

pdf at the boundaries of its support, $x \rightarrow 0^+$ and $x \rightarrow 1^-$, provides insight into its shape:

$$\begin{cases} f(x) \rightarrow \infty & \text{as } x \rightarrow 0^+ \text{ if } a < 1 \\ f(x) \rightarrow b & \text{as } x \rightarrow 0^+ \text{ if } a = 1 \\ f(x) \rightarrow 0 & \text{as } x \rightarrow 0^+ \text{ if } a > 1 \\ f(x) \rightarrow \infty & \text{as } x \rightarrow 1^- \end{cases}$$

The tendency of $f(x)$ to ∞ as $x \rightarrow 1^-$ (for all $a, b > 0$) implies that the DK distribution is unlikely to exhibit a classic bell-shaped unimodality with tails approaching zero at both ends. Instead, it typically presents as J-shaped or L-shaped, with a potential mode at or very near $x = 1$.

For the special case when $a_0 = 1$, the derivative equation simplifies to:

$$1 - \frac{2b}{1 - b \ln(1-x)} = 0$$

This yields a critical point at $x = 1 - \exp\left(\frac{1-2b}{b}\right)$. For this mode to be valid within $(0, 1)$, we require $b > 0.5$. Due to the inherent complexity of Eq. (13) and the boundary behavior of the pdf, the unimodality of the DK distribution for general a, b pairs often necessitates numerical analysis. This involves numerically solving $f'(x) = 0$ for x and examining the second derivative $f''(x)$ to determine the nature of the critical points (maximum, minimum, or inflection). Graphically, we present the area plot of the pdf for various parameter combinations in Figure 1.

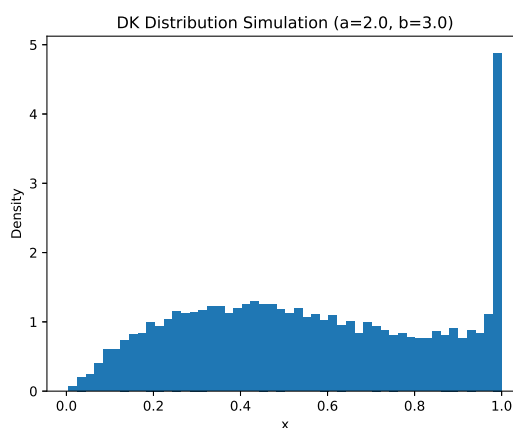


Fig. 1: Area plot of the DK pdfs for selected parameter pairs (a, b) .

Figure 2 plots are polar plots that play a crucial role in understanding the behavior of the DK distribution. The plots provide a graphical representation of the pairs of two parameters that determine if the distribution is

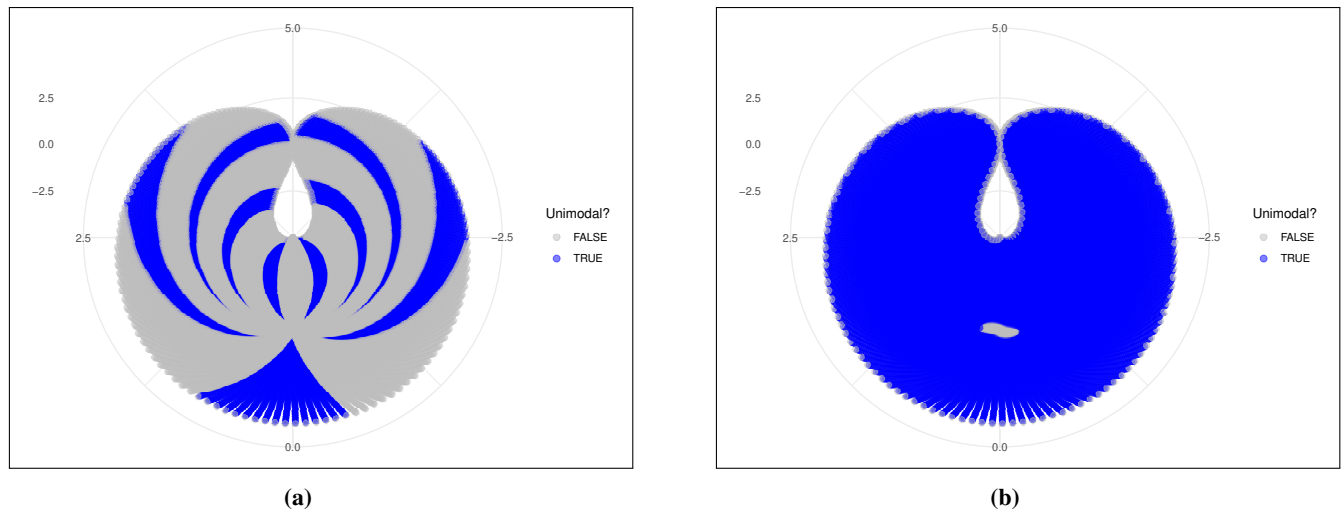


Fig. 2: Polar plots of parameter dependences yielding a unimodal DK distribution, with some fixed values of $a_0 \in (0, 1)$ and two different angular intervals: (a) $b \in [0, 8\pi]$; (b) $b \in [0, 80\pi]$.

unimodal with one peak or multimodal with many peaks. The blue regions represent parameter combinations that give a unimodal distribution, a characteristic that many statistical uses prefer, and the gray regions represent parameter values that give a multimodal shape. These types of plots are important because they lead researchers to select ranges for parameters carefully so that the behavior of their model aligns with the characteristics of the data that they are modeling. They also provide us with information regarding the sensitivity of the modality to changes in parameters of the distribution, revealing complex mathematical dependencies that would be difficult to determine from equations alone. The difference between the two plots, driven by the change in the range of the angular parameter, demonstrates how even a seemingly slight change can have a profound impact on the overall shape of the distribution.

The plots in Figure 3 illustrate the properties of the DK distribution. Figure (a) shows the PDF for various parameter values of a and b . The curves exhibit a wide range of shapes, including J-shapes, reverse J-shapes, and skewed unimodal forms. This versatility demonstrates the distribution's ability to model different types of bounded data. Figure (b) displays the corresponding hazard functions. These functions show increasing, decreasing, and bathtub-shaped patterns, highlighting the flexibility of the DK distribution in modeling phenomena with varying hazard rates over time or a specific interval.

4 Structural Properties

In this section, we discuss some properties of the DK model which include quantile function, moment, order

statistics, moment generating function, and mean residual life function.

4.1 Quantile Function

Let $F(x; a, b)$ be the cdf defined on $x \in (0, 1)$. Then the quantile function F^{-1} is a mapping: $F^{-1} : (0, 1) \rightarrow (0, 1)$ such that for each $u \in (0, 1)$, the quantile function is given by:

$$F^{-1}(u; a, b) = \left(1 - \exp\left(-\frac{u}{b(1-u)}\right) \right)^{1/a}, \quad (14)$$

$$a > 0, b > 0.$$

4.2 Random Variate Generation (Inverse Transform)

Let F denote the DK cdf and $Q(u) = F^{-1}(u)$ its quantile (14). For $U \sim \text{Unif}(0, 1)$, define

$$X = Q(U) = \left(1 - \exp\left(-\frac{U}{b(1-U)}\right) \right)^{1/a}.$$

Proposition 1. X has cdf F , hence the map $U \mapsto Q(U)$ yields i.i.d. DK samples.

Proof. Since F is continuous and strictly increasing on $(0, 1)$, for any $x \in (0, 1)$,

$$\mathbb{P}(X \leq x) = \mathbb{P}(Q(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x).$$

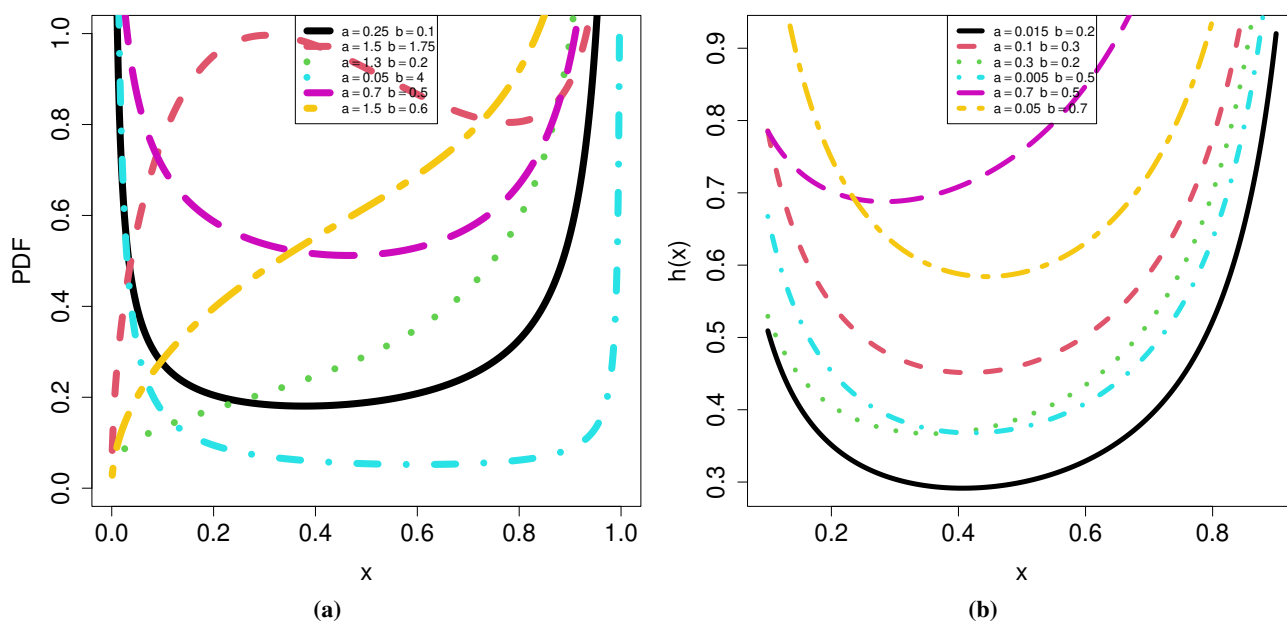


Fig. 3: Plots of (a) pdf (b) hazard function for DK distribution

Remark. To avoid catastrophic cancellation near $U \approx 1$, compute $t = U/(b(1 - U))$ and use $1 - e^{-t} = -\text{expm1}(-t)$. This handles endpoints by clipping U to $(\varepsilon, 1 - \varepsilon)$ with $\varepsilon \in [10^{-16}, 10^{-12}]$ in double precision.

Input: $a > 0, b > 0$, sample size n

for $i = 1$ **to** n **do**

 Draw $U_i \sim \text{Unif}(0, 1)$;

 Set $t_i \leftarrow U_i/(b(1 - U_i))$;

 Set $X_i \leftarrow (-\text{expm1}(-t_i))^{1/a}$;

end

Output: $X_1, \dots, X_n$

Algorithm 1: Inverse-Transform Sampling for DK(a, b)

$$\begin{aligned}\mu'_r &= E(X^r) = \int_0^1 x^r f(x) dx \\ &= ab \int_0^1 \frac{x^{r+a-1}}{(1-x^a)[1-b\log(1-x^a)]^2} dx.\end{aligned}$$

Using integration by parts

$$u = x^r; \quad dv = \frac{abx^{a-1}}{(1-x^a)[1-b\log(1-x^a)]^2} dx.$$

The final form of the r th moment is given as

$$\begin{aligned}\mu'_r &= \frac{r}{ab} \sum_{j=0}^{\infty} (-1)^j \binom{r/a-1}{j} e^{(j+1)/b} E_1\left(\frac{j+1}{b}\right), \\ E_1(z) &= \int_z^{\infty} \frac{e^{-t}}{t} dt.\end{aligned}\tag{15}$$

This equation expresses the r th moment of the DK distribution as an infinite series involving the exponential integral special function $E_1(z)$. The first moment, which corresponds to the mean, is derived by replacing r with 1 in the Eq. (15).

$$\mu'_1 = \mu = \frac{1}{ab} \sum_{j=0}^{\infty} (-1)^j \binom{1/a-1}{j} e^{(j+1)/b} E_1\left(\frac{j+1}{b}\right).$$

The second, third, and fourth crude moments can be derived by substituting $r = 2, r = 3$, and $r = 4$ into

4.3 Moment

The r th crude moment, denoted as μ'_r or $E(X^r)$, of a continuous random variable X is defined by the integral of x^r multiplied by the pdf $f(x)$ over the support of the distribution. First, we write the formal definition for the r th moment using the DK distribution (10) alongside its supported interval $(0, 1)$.

Eq. (15).

$$\mu'_2 = \frac{2}{ab} \sum_{j=0}^{\infty} (-1)^j \binom{2/a-1}{j} e^{(j+1)/b} E_1 \left(\frac{j+1}{b} \right),$$

$$\mu'_3 = \frac{3}{ab} \sum_{j=0}^{\infty} (-1)^j \binom{3/a-1}{j} e^{(j+1)/b} E_1 \left(\frac{j+1}{b} \right),$$

and

$$\mu'_4 = \frac{4}{ab} \sum_{j=0}^{\infty} (-1)^j \binom{4/a-1}{j} e^{(j+1)/b} E_1 \left(\frac{j+1}{b} \right).$$

We now demonstrate the numerical computation and simulation of raw moments for the DK distribution, leveraging the explicit quantile function derived in Eq. (14). This allows us to simulate DK random variables efficiently (see Algorithm 1), and empirically study the behavior of moments as functions of the order r and parameters a, b . Finally, we compare the series-based analytical formula in Eq. (15) to a quantile-based numerical integral.

Remark. We use the closed form quantile (15) to compute

$$\mu'_r = \mathbb{E}[X^r] = \int_0^1 F^{-1}(u; a, b)^r du \quad (16)$$

via the trapezoidal rule; this is equivalent to $\int_0^r x^r f(x) dx$ by the change of variables $u = F(x)$.

We now establish the dependence of μ'_r on the parameters a and b . Recalling Eq. (14), set

$$s(u, b) := 1 - \exp\left(-\frac{u}{b(1-u)}\right) \in (0, 1),$$

$$Q(u; a, b) = F^{-1}(u; a, b) = s(u, b)^{1/a}.$$

Then for a fixed b and $u \in (0, 1)$,

$$\frac{\partial Q}{\partial a} = -\frac{1}{a^2} \ln s(u, b) \cdot s(u, b)^{1/a} > 0,$$

since $\ln s(u, b) < 0$. Therefore $Q(u; a, b)$ is pointwise increasing in a , and by monotone dominance we have from Eq. (16) that $\mu'_r(a, b)$ is strictly increasing in a , for every $r > 0$.

Next, differentiating $s(u, b)$ w.r.t b yields

$$\frac{\partial s}{\partial b} = -\exp\left(-\frac{u}{b(1-u)}\right) \frac{u}{b^2(1-u)} < 0.$$

Since Q is increasing in s , we have $\partial Q / \partial b < 0$ pointwise, hence $\mu'_r(a, b)$ is strictly decreasing in b .

Figure 4 is visual comparison of moments μ'_r computed by the quantile-integral method in Eq. (16) and the series formula in Eq. (15) for the DK distribution with $a = 2, b = 3$ (a). Moments μ'_r of the DK distribution with $a = 50, b = 3$, computed using the series-integral method

in Eq. (15) (b). The plot in panel 4a compares two different methods of calculating the raw moments of the DK distribution for specified values of the parameters ($a = 2$ and $b = 3$). The blue line with circles indicates the moments calculated via the quantile function, a generally powerful and easy method of calculating moments, especially for complex distributions. The orange line with squares represents the moments calculated using the series expansion method, involving the approximation of the moment integral by an infinite series. That both lines are very close to and decline in a similar fashion as the moment order (r) rises provides a significant observation. This near perfect alignment validates the theoretical accuracy of both the quantile and series forms of the moments of the DK distribution. It proves the mathematical expressions derived for the moments to be correct and consistent, which is of utmost importance to the credibility of the new distribution. Panel 4b shows a graph of the raw moments of the DK distribution as a function of the order of the moment (r), for another choice of parameters ($a = 50$ and $b = 3$). The curve increases. This is exactly the expectation for any distribution for which, say, fourth-order moments exist: as the order of the moment increases, so does the value of the moment. The plot provides a clear graphical representation of the behavior of the moments and can be used to interpret the characteristics of the distribution for higher-order statistics (e.g., higher-order kurtosis), which usually are of interest in more specialized applications like risk analysis.

The plots in Figure 5 illustrate the behavior of the key moments of the DK distribution as a function of its parameters a and b . Panel (a) shows that the mean is an increasing function of both parameters, indicating that higher values of a and b shift the central tendency of the distribution to the right. The variance, as depicted in panel (b), also increases with a and b but exhibits a more complex, non-monotonic behavior with a distinct peak, suggesting that the dispersion of the distribution is highly sensitive to the interplay between the two parameters. Similarly, the plots for skewness and kurtosis, in panels (c) and (d) respectively, reveal that both are increasing functions of a and b , which implies that the distribution becomes more positively skewed and leptokurtic (heavy-tailed) as the parameter values increase.

4.4 Distribution of Order Statistics

Order statistics are essential for addressing complex challenges in statistical analysis. Suppose we have a random sample $X_1, X_2, \dots, X_n$ taken from a population. The r^{th} order statistic, denoted as $X_{(r)}$ (where $r = 1, 2, \dots, n$), is obtained by arranging the sample values in increasing order. These statistics prove especially valuable when analyzing data following the

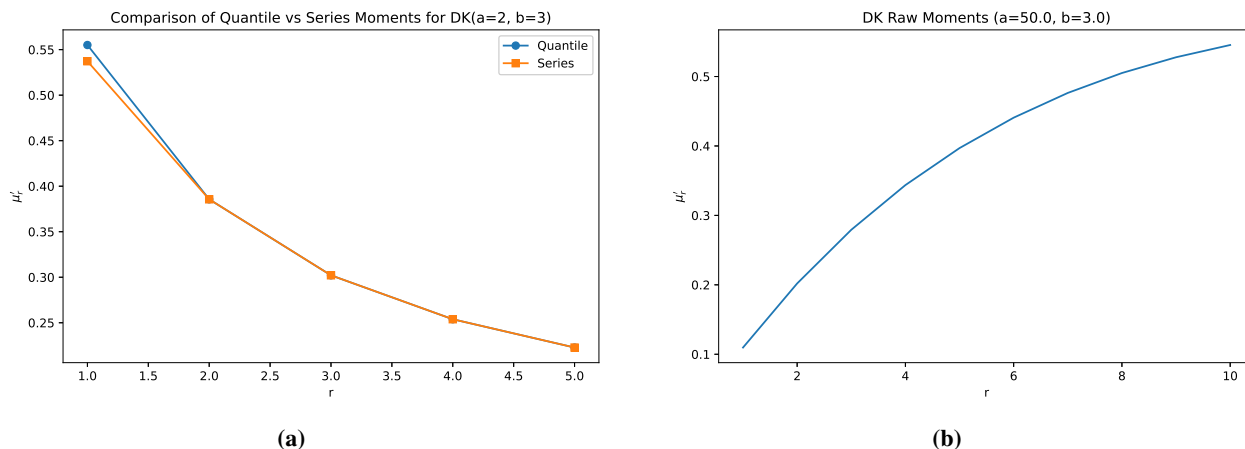


Fig. 4: Comparison of moments μ'_r

DK distribution.

$$y_{r:n}(x, a, b) = \binom{n}{r} f(x) \{F(x)\}^{r-1} [1 - F(x)]^{n-r}.$$

$$y_{r:n}(x, a, b) = \binom{n}{r} \left(\frac{abx^{a-1}}{(1-x^a)[1-b\log(1-x^a)]^2} \right) \times \left\{ 1 - \frac{1}{1-b\log(1-x^a)} \right\}^{r-1} \left[\frac{1}{1-b\log(1-x^a)} \right]^{n-r}.$$

The PDF for the largest order statistic is derived by setting $r = n$, yielding:

$$y_{n:n}(x, a, b) = \left(\frac{abx^{a-1}}{(1-x^a)[1-b\log(1-x^a)]^2} \right) \times \left\{ 1 - \frac{1}{1-b\log(1-x^a)} \right\}^{n-1}.$$

For the case where $r = 1$, the PDF of the first-order statistic is given by:

$$y_{1:n}(x, a, b) = \left(\frac{nabx^{a-1}}{(1-x^a)[1-b\log(1-x^a)]^2} \right) \times \left[\frac{1}{1-b\log(1-x^a)} \right]^{n-1}.$$

4.5 Moment Generating Function (MGF)

The moment generating function (MGF), where it exists is defined as

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_0^1 e^{tx} \frac{abx^{a-1}}{(1-x^a)[1-b\log(1-x^a)]^2} dx.$$

By change of variable transformation, the integral becomes:

$$M_X(t) = \int_0^{\infty} e^{t(1-e^{-y})^{1/a}} \frac{b}{(1+by)^2} dy. \quad (17)$$

The definition of the MGF as a power series is:

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E[X^k],$$

where $E[X^k]$ is the k -th moment of the distribution. The k -th moment is given by:

$$E[X^k] = \int_0^1 x^k f(x) dx = \int_0^1 x^k \frac{abx^{a-1}}{(1-x^a)[1-b\log(1-x^a)]^2} dx = \int_0^{\infty} (1-e^{-y})^{k/a} \frac{b}{(1+by)^2} dy.$$

$$M_X(t)$$

$$= 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \left[\frac{k}{ab} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(k/a)}{j! \Gamma(k/a-j)} e^{(j+1)/b} E_1\left(\frac{j+1}{b}\right) \right]$$

$$E_1(x) = \int_x^{\infty} \frac{e^{-t}}{t} dt. \quad (18)$$

The MGF $M_X(t)$ for the randomly distributed DK model satisfies $M_X(0) = 1$ and $M'_X(0) = \mu'_1 = \mathbb{E}[X]$. Since $0 < X < 1$ a.s., $M_X(t)$ is strictly increasing and convex in t ; as $t \rightarrow +\infty$ it grows exponentially (dominated by mass near $x = 1$), while as $t \rightarrow -\infty$ it decays to 0. Near $t = 0$ we have the usual expansion $M_X(t) = 1 + \mu'_1 t + \frac{1}{2} \mu'_2 t^2 + \dots$, so the curve is almost linear with slope equal to the mean. Parameter effects are intuitive: smaller a concentrates mass near 0, reducing μ'_1

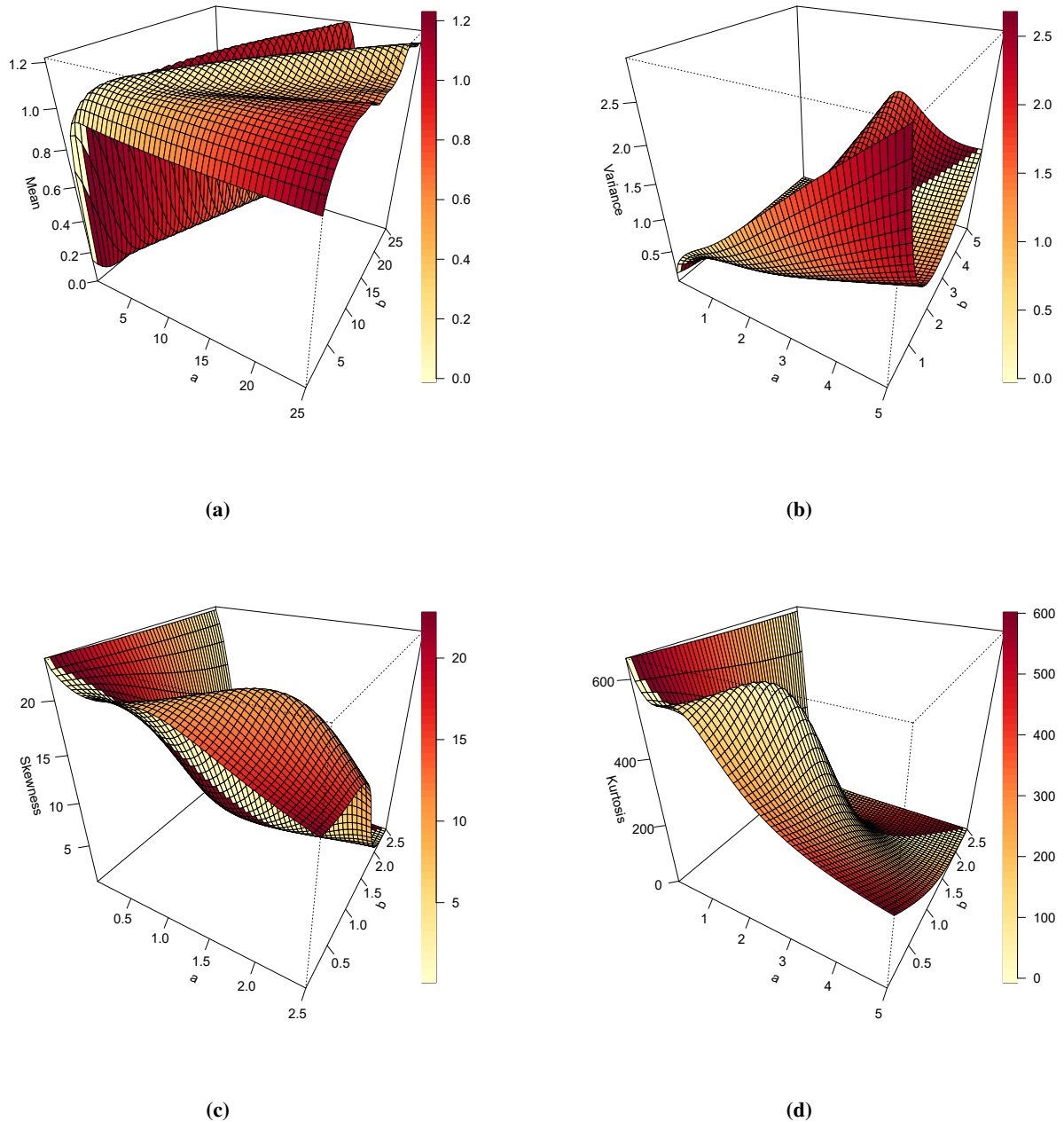


Fig. 5: Plots of (a) Mean, (b) Variance, (c) Skewness and (d) Kurtosis of DK distribution

and flattening the rise for $t > 0$; larger b concentrates mass near 1, increasing μ'_1 , steepening the growth for $t > 0$, and slowing the decay for $t < 0$. Because DK has an endpoint singularity at $x = 1$, the positive- t growth is especially pronounced. We now validate computation by overlaying the single-integral in Eq. (17) and series-based in Eq. (18) evaluations of $M_X(t)$, then we plot $M_X(t)$ for several b values (holding a fixed) to illustrate the

influence of b (Figure 6). On the same grid $\{t_\ell\}_{\ell=1}^L$ we report pointwise errors defined by the absolute error $\mathcal{E}_A(t_\ell) = |M_{\text{series}}(t_\ell) - M_{\text{integral}}(t_\ell)|$ and the relative error $\mathcal{E}_R(t_\ell) = |(M_{\text{series}}(t_\ell) - M_{\text{integral}}(t_\ell))/M_{\text{series}}(t_\ell)|$ (depicted in Figure 7).

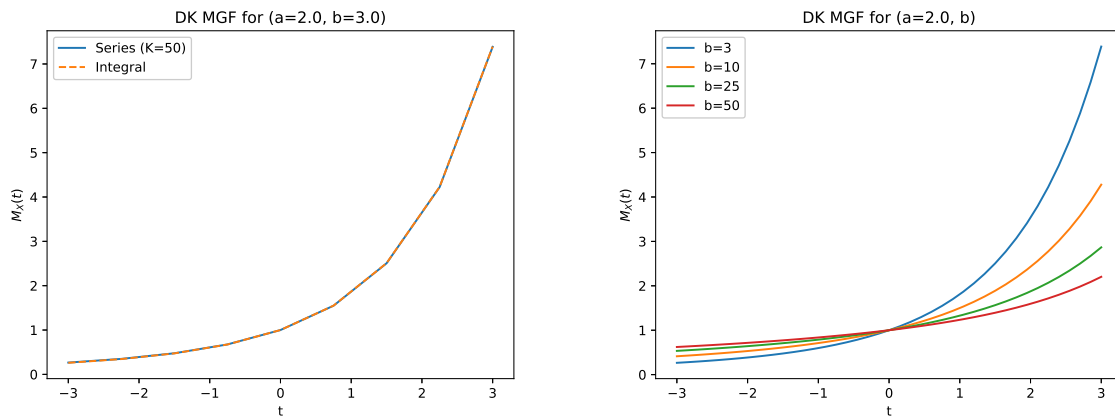


Fig. 6: MGF, $M_X(t)$, for fixed a and different values of b

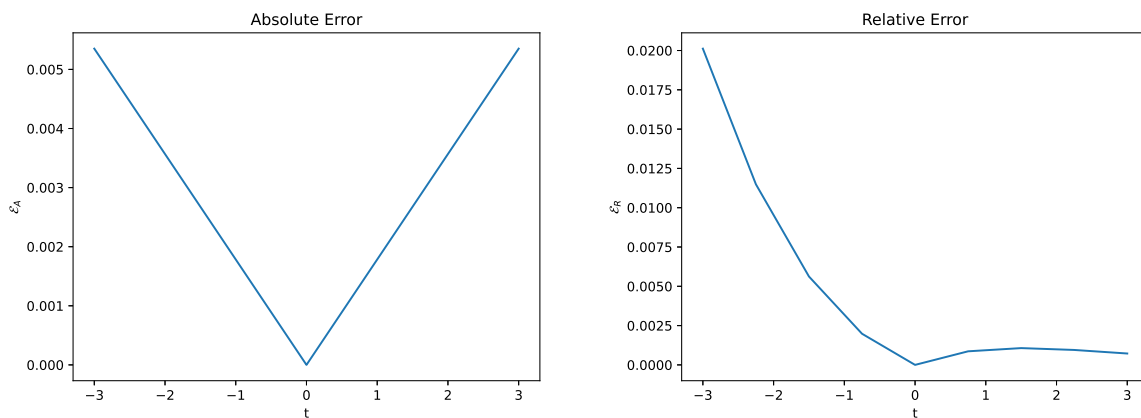


Fig. 7: Absolute (left) and Relative Error (right)

4.6 Mean Residual Life Function

The Mean Residual Life (MRL) function, denoted by $m(t)$, represents the expected additional lifetime of an item given that it has already survived up to time t . Mathematically, it is defined as:

$$m(t) = E[X - t | X > t] = \frac{1}{S(t)} \int_t^\infty S(x) dx,$$

where $S(x)$ is the survival function $S(t) = 1 - F(x; a, b)$ in Eq. 10.

$\therefore$ the final form of the MRL function is

$$m(t) = \frac{1 - b \log(1 - t^a)}{ab} \sum_{n=0}^{\infty} (-1)^n \binom{(1-a)/a}{n} e^{(n+1)/b} \times E_1 \left(\frac{n+1}{b} [1 - b \log(1 - t^a)] \right). \quad (19)$$

For the DK law, the mean residual life $m(t)$ in Eq. (19) lives on $t \in [0, 1)$, starts at $m(0) = \mathbb{E}[X]$, and vanishes as $t \rightarrow 1^-$. Using $m'(t) = h(t)m(t) - 1$ with the DK hazard, we get the universal right-endpoint behavior $m(t) \sim 1 - t$ (slope -1) and the bound $0 \leq m(t) \leq 1 - t$. In practice the curve is decreasing for most parameter settings (DMRL), though when $a < 1$ a small initial uptick can occur before the descent to zero. Parameter effects are monotone: increasing a shifts mass to the right and raises $m(t)$ pointwise, whereas increasing b shifts mass to the left and lowers $m(t)$ across t .

In Figure 8, for the Top-left, varying $b \in \{3, 10, 25, 50\}$ at $a = 2.0$; $m(t)$ decreases pointwise as b increases. Top-right: varying a at fixed $b = 3.0$; $m(t)$ increases pointwise with a . Bottom-right (zoom near $t = 1$): all curves meet at 0 and are tangent to the reference line $1 - t$, confirming a common slope -1 . Bottom-left: cases with $a < 1$ (e.g., $a \in \{0.4, 0.6, 0.8, 0.95\}$ at $b = 3.0$) display the predicted initial

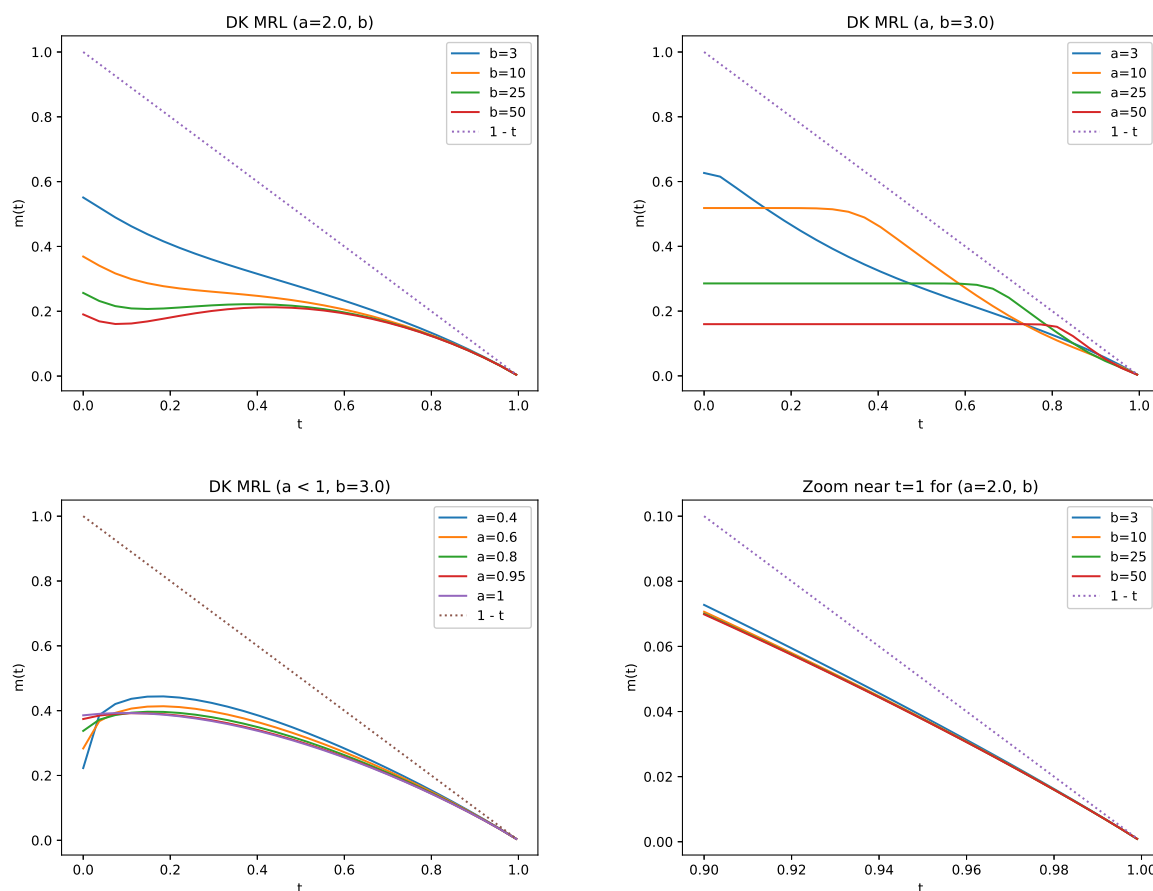


Fig. 8: Mean residual life $m(t)$ for the DK distribution.

uptick before decreasing, while remaining below $1 - t$ for all t .

5 Parameter Estimation

In section, we derive the estimators of the parameters of the DK distribution using both non-bayesian and bayesian techniques.

5.1 Maximum Likelihood Estimation (MLE)

The method of MLE, first proposed by [62,63,64,?], consider a random sample of size n , denoted by $x_1, x_2, \dots, x_n$, where each observation is independently and identically distributed (i.i.d.) following an DK distribution with unspecified parameters a and b . The

log-likelihood function can be derived as

$$\ell(a, b) = \log [L(a, b)] = n \log a + n \log b + (a - 1) \sum_{i=1}^n \log x_i - \sum_{i=1}^n \log(1 - x_i^a) - 2 \sum_{i=1}^n \log[1 - b \log(1 - x_i^a)].$$

Take partial derivative with respect to a yields;

$$\frac{\partial \ell}{\partial a} = \frac{n}{a} + \sum_{i=1}^n \log x_i + \sum_{i=1}^n \frac{x_i^a \log x_i}{1 - x_i^a} - 2b \sum_{i=1}^n \frac{x_i^a \log x_i}{(1 - x_i^a)[1 - b \log(1 - x_i^a)]}.$$

Equating this expression to zero yields the MLE condition for the parameter a :

$$\frac{n}{a} + \sum_{i=1}^n \log x_i + \sum_{i=1}^n \frac{x_i^a \log x_i}{1 - x_i^a} - 2b \sum_{i=1}^n \frac{x_i^a \log x_i}{(1 - x_i^a)[1 - b \log(1 - x_i^a)]} = 0. \quad (20)$$

Similarly, take partial derivative with respect to b gives;

$$\frac{\partial \ell}{\partial b} = \frac{n}{b} - 2 \sum_{i=1}^n \frac{-\log(1 - x_i^a)}{1 - b \log(1 - x_i^a)}$$

$$\frac{\partial \ell}{\partial b} = \frac{n}{b} + 2 \sum_{i=1}^n \frac{\log(1 - x_i^a)}{1 - b \log(1 - x_i^a)}.$$

Equating this expression to zero yields the MLE condition for the parameter b :

$$\frac{n}{b} + 2 \sum_{i=1}^n \frac{\log(1 - x_i^a)}{1 - b \log(1 - x_i^a)} = 0. \quad (21)$$

The MLEs of the parameters a and b are obtained by numerically solving the coupled equations specified in Eqs. (20) and (21).

5.2 Least Squares Estimation (LSE)

The method of LSE, first proposed by [4], offers an approach for determining parameters of the Beta distribution. Building upon the foundational concepts presented in their work, we develop an estimation technique that minimizes the sum of squared differences between empirical observations and theoretical predictions. This approach, which focuses on optimizing the fit of the Beta distribution model, continues to be a crucial tool for deriving accurate parameter estimates.

$$E[F(x_{k:n}|a, b)] = \frac{k}{n+1} \quad \text{and}$$

$$V[F(x_{k:n}|a, b)] = \frac{k(n-k+1)}{(n+1)^2(n+2)}.$$

The estimates of the parameters a and b are derived through the least squares method, yielding the estimators $\hat{a}_{LSE}$ and $\hat{b}_{LSE}$. These estimators are computed by minimizing the objective function $L(a, b)$ with respect to a and b .

$$L(a, b) = \arg \min_{(a, b)} \sum_{k=1}^n \left[F(x_{k:n}|a, b) - \frac{k}{n+1} \right]^2. \quad (22)$$

The estimates are derived by solving the following set of non-linear equations:

$$\sum_{k=1}^n \left[F(x_{k:n}|a, b) - \frac{k}{n+1} \right]^2 \Delta_1(x_{k:n}|a, b) = 0 \quad (23)$$

$$\sum_{k=1}^n \left[F(x_{k:n}|a, b) - \frac{k}{n+1} \right]^2 \Delta_2(x_{k:n}|a, b) = 0. \quad (24)$$

Where

$$\Delta_1(x_{k:n}|a, b) = \frac{bx^a \log(x)}{(1 - x^a)(1 - b \log(1 - x^a))^2}. \quad (25)$$

$$\Delta_2(x_{k:n}|a, b) = -\frac{\log(1 - x^a)}{(1 - b \log(1 - x^a))^2}. \quad (26)$$

The expressions given in Eqs. (25) and (26) are obtained by computing the first-order partial derivatives of the DK distribution's CDF, shown in Eq. (10), with respect to the parameters a and b , respectively.

5.3 Weighted Least Squares Estimation (WLSE)

The parameters a and b of the DK distribution are estimated using the weighted least squares approach. The resulting estimators, denoted by $\hat{a}_{WLSE}$ and $\hat{b}_{WLSE}$, are derived by minimizing the objective function $W(a, b)$ with respect to these parameters.

$$W(a, b) = \arg \min_{(a, b)} \sum_{k=1}^n w_k \left[F(x_{k:n}|a, b) - \frac{k}{n+1} \right]^2, \quad (27)$$

where

$$w_k = \frac{(n+1)^2(n+2)}{k(n-k+1)}.$$

$$\sum_{k=1}^n w_k \left[G(x_{k:n}|a, b) - \frac{k}{n+1} \right]^2 \Delta_1(x_{k:n}|a, b) = 0. \quad (28)$$

$$\sum_{k=1}^n w_k \left[G(x_{k:n}|a, b) - \frac{k}{n+1} \right]^2 \Delta_2(x_{k:n}|a, b) = 0. \quad (29)$$

The estimators $\Delta_1(x_{k:n}|a, b)$ and $\Delta_2(x_{k:n}|a, b)$ are as given in Eqs. (25) and (26), respectively.

5.4 Maximum Product of Spacing Estimation (MPS)

The MPS method was proposed by [5] as a different approach compared to traditional maximum likelihood estimation. This technique provides an approximation of the Kullback-Leibler information criterion while employing a unique methodology independent of maximum likelihood principles. When working with ordered data in ascending sequence, this method proves particularly suitable for implementation.

$$I_s(X|a, b) = \left[\prod_{k=1}^{n+1} D_j(x_{k:n}|a, b) \right]^{\frac{1}{n+1}}, \quad (30)$$

Where $D_j(x_{k:n}|a, b) = F(x_k|a, b) - F(x_{k-1}|a, b)$; $k = 1, 2, 3, \dots, n$. In a similar way, it is possible to opt for maximizing the function.

$$N(a, b) = \frac{1}{n+1} \sum_{k=1}^{n+1} \ln(D_j(x_{k:n}|a, b)). \quad (31)$$

To obtain the parameter estimates, we compute the partial derivatives of the function $N(a, b)$ with respect to a and b . By setting these derivatives to zero, $\frac{\partial N(a, b)}{\partial a} = 0$ and $\frac{\partial N(a, b)}{\partial b} = 0$, and solving the resulting nonlinear system of equations, we can determine the optimal values for the parameters.

5.5 Bayesian Inference

The objective of this study is to obtain Bayesian estimators (BEs) for the unknown parameters of the DK distribution. Within the Bayesian inference framework, various loss functions such as quadratic (squared error), LINEX (linear-exponential), and generalized entropy loss functions will be utilized for parameter estimation. Assuming independence between the parameters, we adopt gamma-distributed priors for both a and b , which are subsequently integrated into the joint prior distribution of the DK model.

$$\left. \begin{aligned} \phi_1(a) &\propto a^{\theta_1-1} e^{-(1-\theta_1)a}, & a > 0, \theta_1 > 0, (1-\theta_1)_1 > 0 \\ \phi_2(b) &\propto b^{\theta_2-1} e^{-(1-\theta_2)b}, & b > 0, \theta_2 > 0, (1-\theta_2)_2 > 0. \end{aligned} \right\} \quad (32)$$

The specification of the hyperparameters θ_k and $(1-\theta)_k$ for $k = 1, 2$ is informed by prior beliefs about the unknown parameters. The joint prior distribution for the parameter vector $\pi = (a, b)$ is consequently defined as:

$$\left. \begin{aligned} \phi_1(\pi) &= \phi_1(a)\phi_2(b) \\ \phi_1(\pi) &\propto a^{\theta_1-1} b^{\theta_2-1} e^{-((1-\theta_1)a + (1-\theta_2)b)}. \end{aligned} \right\} \quad (33)$$

The posterior distribution, after observing the data $X = (x_1, x_2, \dots, x_n)$, is proportional to:

$$\phi(\pi | X) = \frac{\phi(\pi) \mathbf{1}(\pi)}{\int_{\pi} \phi(\pi) \mathbf{1}(\pi) d\pi}.$$

Consequently, the form of the posterior distribution is given by:

$$\begin{aligned} \phi(\pi | X) &\propto \frac{b^{2n+\theta_2-1} a^{\theta_1-1}}{(ab+2)^2} \\ &\times e^{-b \sum_{k=1}^n (x_k - (1-\theta_1)a - (1-\theta_2)b)} \prod_{k=1}^n (a + bx^2). \end{aligned} \quad (34)$$

Bayesian parameter estimates are derived under various loss functions. For the squared error loss (SEL), the Bayes estimator is given by:

$$\hat{\phi}_{\text{BE-SEL}} = \mathbb{E}[\phi | X] = \int \phi \pi(\phi | X) d\phi. \quad (35)$$

Alternative loss functions, such as LINEX and generalized entropy loss (GEL), address asymmetric estimation scenarios. The Bayes estimator under LINEX loss is defined as:

$$\hat{\phi}_{\text{LINEX}} = -\frac{1}{\eta} \log \left(\int e^{-\eta \phi} \pi(\phi | X) d\phi \right), \quad (36)$$

where $\eta \neq 0$ reflects the asymmetry in estimation. For GEL, the Bayes estimator becomes:

$$\hat{\phi}_{\text{GEL}} = \left(\int \phi^{-l} \pi(\phi | X) d\phi \right)^{-1/l}, \quad (37)$$

where $l \neq 0$ is the asymmetry parameter.

Since these estimators often lack closed-form solutions, numerical methods like Markov chain Monte Carlo (MCMC) are employed. The MCMC procedure for approximating Bayesian estimates involves the following steps:

1. Initialize the parameters $\phi^{(0)}$ and set the number of iterations M .
2. Generate samples $\phi^{(j)}$ from the posterior distribution $\pi(\phi | X)$ using algorithms like the Metropolis-Hastings or Gibbs sampler.
3. Discard the initial τ_D samples as the burn-in period to ensure convergence.
4. Use the remaining $M - \tau_D$ samples to compute Bayesian estimates as follows:

$$\begin{aligned} \hat{\phi}_{\text{SEL}} &= \frac{1}{M - \tau_D} \sum_{j=\tau_D+1}^M \phi^{(j)}, \\ \hat{\phi}_{\text{LINEX}} &= -\frac{1}{\eta} \log \left(\frac{1}{M - \tau_D} \sum_{j=\tau_D+1}^M e^{-\eta \phi^{(j)}} \right), \\ \hat{\phi}_{\text{GEL}} &= \left(\frac{1}{M - \tau_D} \sum_{j=\tau_D+1}^M (\phi^{(j)})^{-l} \right)^{-1/l}. \end{aligned}$$

This algorithm enables the computation of Bayesian estimates under SEL, LINEX, and GEL loss functions, providing robust parameter estimates tailored to specific applications. For further reading, refer to [38] and [39].

6 Monte Carlo Simulation

The aim of this simulation study is to compare the performance of non-Bayesian (MLE, MPS, LS and WLS) and Bayesian estimation methods for the parameters of the DK distribution. The study employs a 1,000-replication design for each sample size considered ($n = 15, 25, 50, 75$). Estimator accuracy and reliability are assessed using the averaged bias and root mean square error (RMSE) across replications. For the Bayesian context, the study considered squared error loss function (SEL) and the asymmetry in loss functions is explored using:

iLINEX function: $\eta = -0.5$ (LINEX1) and $\eta = 0.5$ (LINEX2).

iiGEL loss function: $l = -0.5$ (GEL1) and $l = 0.5$ (GEL2).

The study also examines different scenarios by varying the initial parameter values, which includes;

Case Ia = 0.5 and $b = 1.75$

Case IIa = 1.75 and $b = 1.95$

Case IIIa = 1.5 and $b = 1.55$

Case IVa = 1.25 and $b = 1.70$

The simulation results presented in Tables 2, 3, 4, and 5 consistently show that the performance of all estimators improves as the sample size (n) increases. Across all four cases, both the bias and the RMSE for both parameters ($\hat{a}$ and $\hat{b}$) decrease significantly with larger n . Comparing the non-Bayesian methods, the MPS method generally provides the lowest bias and RMSE for both estimators across all sample sizes. For the non-Bayesian estimators, the LS and WLS methods often show higher variability, particularly for the $\hat{b}$ estimator at smaller sample sizes. Among the Bayesian estimators, which are presented under various loss functions (SEL, LINEX, and GEL), their performance is generally inferior to the non-Bayesian methods in terms of bias and RMSE for both parameters, especially at smaller sample sizes. For example, in Table 2, the Bayesian estimators for $\hat{a}$ and $\hat{b}$ have much larger biases and RMSEs than the non-Bayesian estimators, a trend that is consistently observed in Tables 3, 4, and 5. The Bayesian estimators with the GEL2 loss function consistently show the worst performance with very high bias and RMSE values.

The plots in Figure 9 illustrate the performance of different non-Bayesian estimation methods for parameters $\hat{a}$ and $\hat{b}$ across four distinct simulation cases. A consistent trend observed across all panels (a), (b), (c), and (d) is that both the Bias and the RMSE decrease as the sample size (n) increases, which is a desirable property for any estimator. For the parameter $\hat{a}$, the MPS and MLE methods generally show the lowest bias and RMSE, particularly as the sample size grows. In contrast, the LS and WLS methods tend to exhibit higher bias and RMSE for $\hat{a}$. For the parameter $\hat{b}$, the MPS estimator consistently outperforms the others by having the lowest bias and RMSE across all cases and sample sizes. The performance of the MLE and LS/WLS estimators for $\hat{b}$ is more variable, with their biases and RMSEs often being considerably higher than those of the MPS estimator, especially at smaller sample sizes.

The plots in Figure 10 illustrate the performance of various Bayesian estimators under different loss functions across four simulation cases. A clear trend across all panels (a), (b), (c), and (d) is that both the bias and the Root Mean Square Error (RMSE) for parameters $\hat{a}$ and $\hat{b}$ decrease as the sample size (n) increases, confirming the asymptotic properties of these estimators. For the

parameter $\hat{a}$, the SEL, LINEX1, LINEX2, and GEL1 estimators show very similar performance, with their bias and RMSE values being closely clustered together. The GEL2 estimator consistently performs worse than the others for $\hat{a}$, exhibiting a significantly higher bias and RMSE across all cases and sample sizes. For the parameter $\hat{b}$, the SEL, LINEX1, LINEX2, and GEL1 estimators also have similar performance, though the differences between them are slightly more pronounced compared to $\hat{a}$. The GEL2 estimator for $\hat{b}$ again shows the worst performance, with markedly higher bias and RMSE values.

7 Applications

The Dhillon-G distribution family is well equipped to model bounded data in the unit interval with complex shapes like bimodality and bathtub curves, which extends its application far beyond the examples shown in this paper. In reliability and engineering, it can be used to estimate the percentage of system operating time or percentage of defective items in a batch manufactured, its ability to fit a bathtub curve being particularly useful in failure analysis. In biology and medicine, the model can be used to examine the disease prevalence, effectiveness rate of treatment, or percentage of gene expression among members of a population. Its usage in economics and finance is in the model's capacity to accommodate bounded variables such as asset return or unemployment, which have bimodal or skewed distributions. Secondly, in social science and ecology, the model is a good statistical tool for analysis of data with survey data being given as proportions or tracking the population proportion of the species in question with room for several different analytical ends.

The first data (Data I) is the Dominica Weekly death rate due to COVID-19 from 22/3/2020 to 20/12/2020 assessed from <https://data.who.int/dashboards/covid19/data?n=c> and presented in Table 6.

The second data (Data II) is the Vinyl chloride data from clean upgradient ground-water monitoring wells in (g/L) first studied by [7] and later investigated by others including [8] [52]. The data is presented in Table 7 below.

The third data (Data III) is the weekly volume of traded Bitcoin-USD ('000,000,000) from 17/09/2023 to 15/9/2024, assessed from <https://finance.yahoo.com/quote/BTC-USD/history/?frequency=1d> and reported in Table 8.

The fourth data (Data IV) is the infant mortality rate per 1000 live-births for a few chosen nations in 2021, as reported by <https://data.worldbank.org/indicator/SP.DYN.IMRT.IN>, (accessed on August 9, 2025). This data has been studied by [51] and it is presented in Table 9.

Table 2: Simulation for Case I

Type	Method	Estimator	$n = 15$		$n = 25$		$n = 50$		$n = 75$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Non-Bayesian	MLE	$\hat{a}$	0.08539	0.03543	0.04332	0.01252	0.01254	0.00230	0.00517	0.00070
		$\hat{b}$	0.79883	4.50797	0.38646	1.45428	0.09763	0.16832	0.03202	0.03489
	MPS	$\hat{a}$	0.01019	0.01556	0.00977	0.00625	0.00552	0.00128	0.00313	0.00040
		$\hat{b}$	0.27028	1.38653	0.17359	0.64233	0.05822	0.09812	0.02135	0.02221
	LS	$\hat{a}$	0.00094	0.05244	0.00086	0.02837	0.00228	0.01233	0.00033	0.00796
		$\hat{b}$	0.38828	5.02942	0.31708	13.41343	0.11083	0.52009	0.06652	0.29475
	WLS	$\hat{a}$	0.01227	0.05208	0.01073	0.02475	0.01300	0.01072	0.00902	0.00690
		$\hat{b}$	0.50928	9.92605	0.32621	7.19041	0.15906	0.50660	0.10843	0.28568
Bayesian	SEL	$\hat{a}$	0.25501	0.11042	0.23612	0.08056	0.20963	0.05425	0.19703	0.04633
		$\hat{b}$	2.02396	12.51036	1.70880	5.95677	1.27155	2.76905	1.18662	2.29889
	LINEX1	$\hat{a}$	0.25574	0.11093	0.23677	0.08094	0.21016	0.05451	0.19748	0.04653
		$\hat{b}$	2.08418	13.66924	1.73368	6.14611	1.28726	2.83796	1.19886	2.34560
	LINEX2	$\hat{a}$	0.25427	0.10991	0.23548	0.08018	0.20911	0.05400	0.19659	0.04613
		$\hat{b}$	1.97015	11.61668	1.68337	5.76414	1.25584	2.70131	1.17419	2.25144
	GEL1	$\hat{a}$	0.25407	0.10988	0.23527	0.08012	0.20891	0.05393	0.19640	0.04607
		$\hat{b}$	2.01608	12.41940	1.70320	5.92530	1.26710	2.75300	1.18286	2.28714
	GEL2	$\hat{a}$	2.85939	11.74827	2.86761	10.75902	2.50817	7.42885	2.42530	6.76427
		$\hat{b}$	2.00035	12.24002	1.69195	5.86205	1.25817	2.72092	1.17530	2.26351

Table 3: Simulation for Case II

Type	Method	Estimator	$n = 15$		$n = 25$		$n = 50$		$n = 75$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Non-Bayesian	MLE	$\hat{a}$	0.24754	0.44077	0.12071	0.14032	0.03657	0.03413	0.02767	0.01704
		$\hat{b}$	1.02620	18.67731	0.45648	1.46033	0.13246	0.22683	0.07606	0.11155
	MPS	$\hat{a}$	0.02593	0.23602	0.01446	0.08860	0.00004	0.02500	0.01394	0.01127
		$\hat{b}$	0.30689	4.11485	0.17670	0.63538	0.06540	0.12892	0.05037	0.07118
	LS	$\hat{a}$	0.02903	0.61595	0.00949	0.28822	0.01281	0.13974	0.00156	0.09992
		$\hat{b}$	0.60871	30.99072	0.23479	1.83133	0.12371	0.66018	0.06093	0.40186
	WLS	$\hat{a}$	0.01029	0.53885	0.03411	0.26847	0.02068	0.12425	0.03323	0.08720
		$\hat{b}$	0.59157	19.78864	0.29978	1.85343	0.16987	0.62285	0.10834	0.38348
Bayesian	SEL	$\hat{a}$	0.76220	1.01732	0.71875	0.76299	0.66666	0.55457	0.62948	0.48036
		$\hat{b}$	2.30912	16.71031	1.92376	7.67530	1.43977	3.55824	1.33550	2.99522
	LINEX1	$\hat{a}$	0.77077	1.03516	0.72620	0.77657	0.67274	0.56375	0.63480	0.48784
		$\hat{b}$	2.39928	18.84873	1.95784	7.98171	1.45978	3.65913	1.35122	3.06557
	LINEX2	$\hat{a}$	0.75373	0.99993	0.71139	0.74973	0.66064	0.54563	0.62422	0.47308
		$\hat{b}$	2.23185	15.18872	1.88903	7.36670	1.41982	3.45972	1.31948	2.92352
	GEL1	$\hat{a}$	0.75893	1.01174	0.71585	0.75837	0.66420	0.55111	0.62729	0.47749
		$\hat{b}$	2.29929	16.57082	1.91708	7.63137	1.43475	3.53757	1.33124	2.97975
	GEL2	$\hat{a}$	2.47968	17.24881	2.10369	8.34461	1.62470	4.10618	1.52268	3.51768
		$\hat{b}$	2.27968	16.29694	1.90369	7.54313	1.42470	3.49630	1.32268	2.94861

Table 10 provides a summary of the basic descriptive statistics for four datasets. A key observation is the variation in sample size (n), which ranges from 27 to 53. The central tendency, as measured by the mean and median, is similar for Data II, III, and IV, all falling around 0.1 to 0.2, while Data I has a considerably smaller mean and median, suggesting its values are concentrated at the lower end of the scale. The dispersion of the data, indicated by the variance, standard deviation, and range, is notably low for Data I compared to the other datasets. The presence of outliers is common across all datasets, although the number and magnitude of these outliers vary.

A consistent finding is that all four datasets exhibit positive skewness, with values greater than zero, indicating a tail on the right side of the distribution. Similarly, the kurtosis values, all greater than 3, suggest that each distribution is leptokurtic, meaning they have heavier tails and a sharper peak than a normal distribution.

The results in Table 11 summarize the performance of several distributions ([12], Beta [10], Gumbel [15], Gamma, Burr XII [13] and [14]) fitted to four different datasets, based on information criteria and goodness-of-fit tests. A consistent pattern emerges where the newly

Table 4: Simulation for Case III

Type	Method	Estimator	$n = 15$		$n = 25$		$n = 50$		$n = 75$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Non-Bayesian	MLE	$\hat{a}$	0.23049	0.35180	0.11755	0.13087	0.03194	0.02219	0.01489	0.00645
		$\hat{b}$	0.72284	4.07025	0.35299	1.03608	0.08598	0.13230	0.04319	0.04902
	MPS	$\hat{a}$	0.00322	0.18308	0.01073	0.07785	0.00915	0.01431	0.00605	0.00390
		$\hat{b}$	0.25695	1.29985	0.16087	0.46941	0.05017	0.07867	0.02939	0.03182
	LS	$\hat{a}$	0.01070	0.46698	0.00899	0.25878	0.00358	0.12065	0.00528	0.07450
		$\hat{b}$	0.41194	8.46459	0.25404	1.60182	0.08730	0.43912	0.06173	0.21818
	WLS	$\hat{a}$	0.04619	0.41554	0.04940	0.23254	0.02977	0.10456	0.03423	0.06471
		$\hat{b}$	0.41051	4.97288	0.29771	1.42628	0.12508	0.41255	0.10574	0.22477
Bayesian	SEL	$\hat{a}$	0.77967	1.02222	0.74236	0.76232	0.66281	0.50072	0.64855	0.47371
		$\hat{b}$	1.98717	11.31740	1.72694	5.49966	1.32062	2.43042	1.25119	2.35232
	LINEX1	$\hat{a}$	0.78640	1.03702	0.74838	0.77410	0.66812	0.50850	0.65252	0.47931
		$\hat{b}$	2.03761	12.14097	1.74641	5.63570	1.33337	2.47750	1.26171	2.39103
	LINEX2	$\hat{a}$	0.77301	1.00782	0.73640	0.75080	0.65758	0.49315	0.64461	0.46824
		$\hat{b}$	1.94026	10.63161	1.70720	5.36153	1.30808	2.38481	1.24055	2.31333
	GEL1	$\hat{a}$	0.77687	1.01725	0.73983	0.75797	0.66044	0.49737	0.64673	0.47126
		$\hat{b}$	1.97969	11.23966	1.72221	5.47478	1.31664	2.41751	1.24775	2.34165
	GEL2	$\hat{a}$	2.01472	11.28472	1.76275	5.59862	1.35871	2.52527	1.29084	2.44674
		$\hat{b}$	1.96472	11.08575	1.71275	5.42485	1.30871	2.39190	1.24084	2.32016

Table 5: Simulation for Case IV

Type	Method	Estimator	$n = 15$		$n = 25$		$n = 50$		$n = 75$	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Non-Bayesian	MLE	$\hat{a}$	0.18962	0.23923	0.09552	0.08463	0.02465	0.01673	0.01342	0.00595
		$\hat{b}$	0.78850	4.67055	0.37047	1.13767	0.10433	0.16848	0.04770	0.05765
	MPS	$\hat{a}$	0.01704	0.12614	0.00241	0.04967	0.00122	0.01240	0.00496	0.00393
		$\hat{b}$	0.23678	1.41910	0.15170	0.50849	0.05103	0.10032	0.03018	0.03674
	LS	$\hat{a}$	0.01464	0.36134	0.00951	0.18852	0.00547	0.07736	0.00916	0.05268
		$\hat{b}$	0.41692	4.26303	0.22948	1.68815	0.08651	0.46534	0.06368	0.35310
	WLS	$\hat{a}$	0.04559	0.33982	0.02050	0.17353	0.02185	0.06715	0.03112	0.04625
		$\hat{b}$	0.47397	4.44364	0.26931	1.50162	0.13633	0.44526	0.10210	0.33008
Bayesian	SEL	$\hat{a}$	0.59274	0.62438	0.55598	0.45995	0.52376	0.32511	0.50449	0.28572
		$\hat{b}$	2.07026	12.78619	1.73611	5.89750	1.40135	2.95545	1.20286	2.20698
	LINEX1	$\hat{a}$	0.59724	0.63194	0.55983	0.46575	0.52699	0.32883	0.50716	0.28870
		$\hat{b}$	2.13177	13.92612	1.75949	6.07042	1.41712	3.02250	1.21376	2.24600
	LINEX2	$\hat{a}$	0.58830	0.61700	0.55217	0.45426	0.52056	0.32146	0.50184	0.28279
		$\hat{b}$	2.01494	11.89443	1.71255	5.72387	1.38584	2.89085	1.19181	2.16748
	GEL1	$\hat{a}$	0.59043	0.62126	0.55398	0.45734	0.52198	0.32315	0.50301	0.28413
		$\hat{b}$	2.06210	12.69266	1.73074	5.86845	1.39685	2.93946	1.19933	2.19682
	GEL2	$\hat{a}$	2.49582	14.55190	2.16999	7.56074	1.83787	4.35932	1.64224	3.45189
		$\hat{b}$	2.04582	12.50816	1.71999	5.81025	1.38787	2.90774	1.19224	2.17637

proposed DK distribution provides the best fit across all four datasets. This is evidenced by its lowest values for the AIC, CAIC, BIC, and HQIC information criteria, which penalize models with more parameters while rewarding a better log-likelihood (LL). For example, in Data I, the DK distribution has a higher log-likelihood than Gumbel, and better information criteria values than all other distributions. Furthermore, the goodness-of-fit tests, specifically the Anderson-Darling (A) and Cramér-von Mises (W) statistics, are consistently at their minimum for the DK distribution in all four cases. Most notably, the Kolmogorov-Smirnov (KS) test for the DK

distribution consistently yields the highest p-values across all datasets (0.9984 for Data I, 0.9873 for Data II, 0.7851 for Data III, and 0.9837 for Data IV), confirming that it provides the most plausible fit to the observed data. The other distributions, such as Kumaraswamy, Beta, and Gamma, generally perform similarly to each other but are outperformed by the DK distribution.

Table 12 provides the Maximum Likelihood Estimates (MLEs) and their corresponding standard errors for the parameters of various distributions fitted to four datasets. The standard error, which indicates the precision of the estimates, generally appears to be relatively small for the

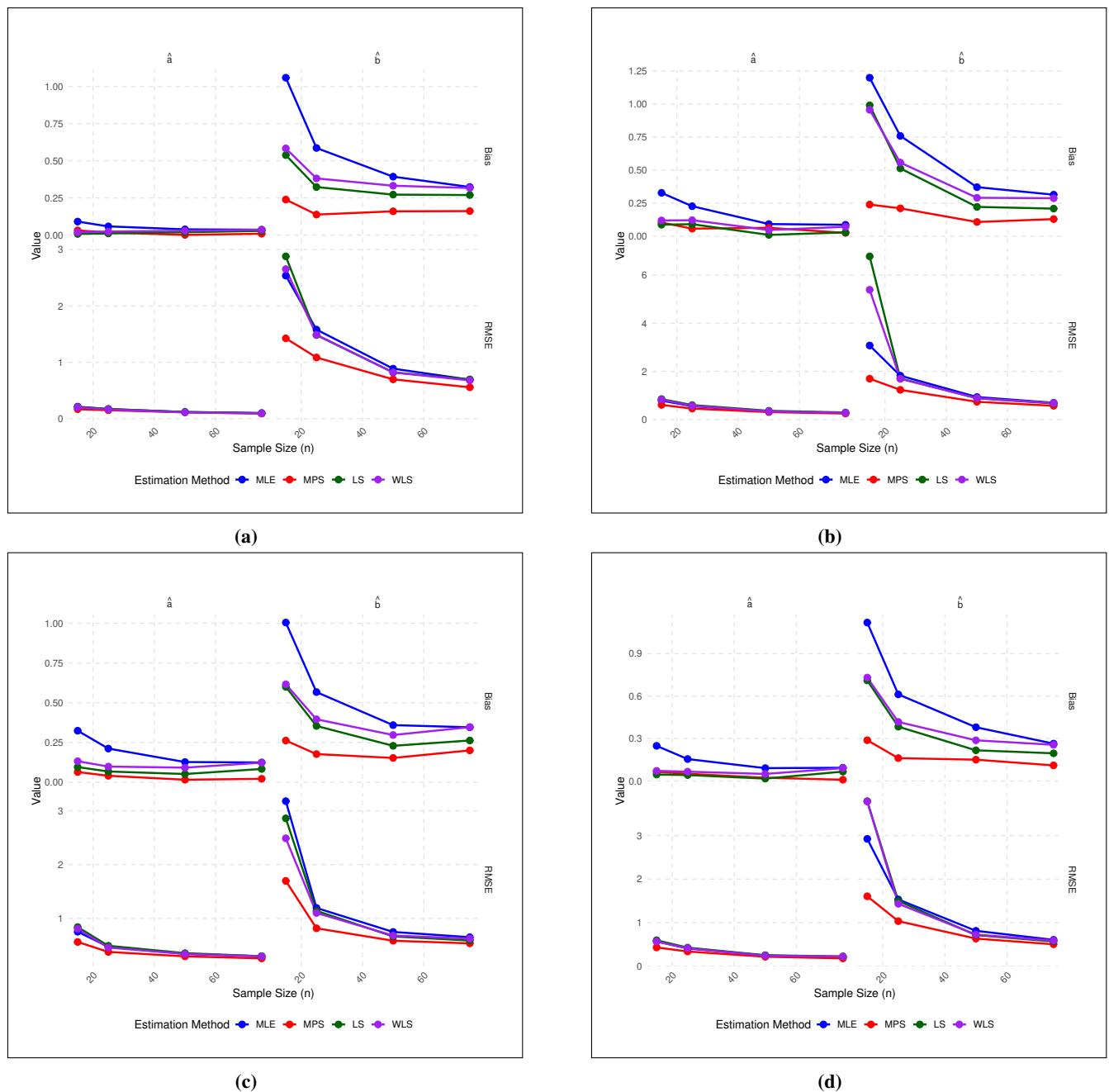
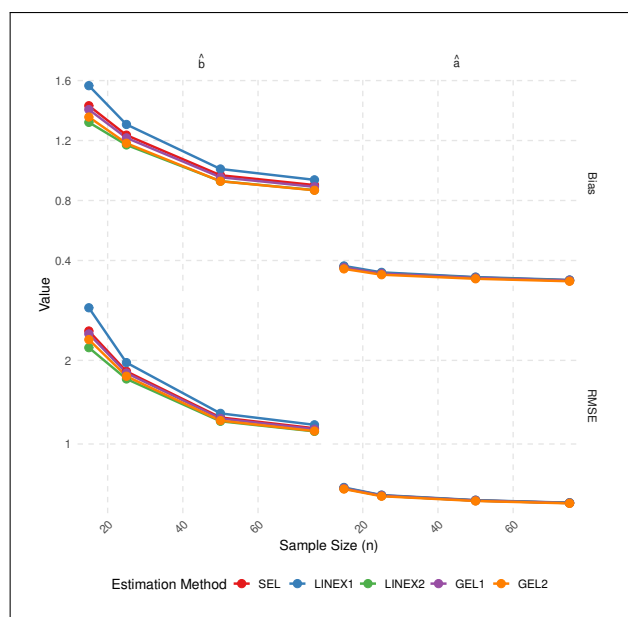


Fig. 9: Plots of (a) Case I, (b) II, (c) III and (d) IV for Non-Bayesian Simulation Results

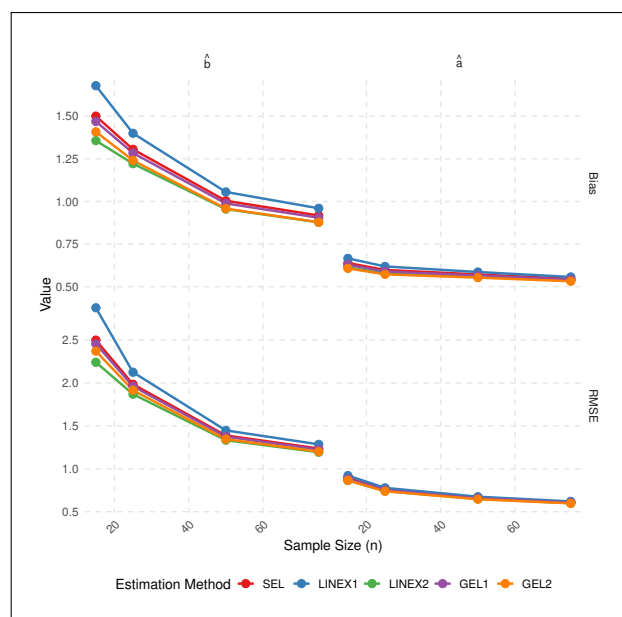
DK distribution, especially in comparison to some of the other models. For instance, in Data I, the standard error for $\hat{b}$ of the DK distribution is 1.2620, which is considerably lower than the large standard errors observed for the BurrXII distribution at 201.5043, Beta at 24.4471, and Gamma at 25.3974. This trend suggests that the parameters of the DK distribution are estimated with a higher degree of precision and stability. The Kumaraswamy distribution also shows relatively small

standard errors, indicating good precision. The Gumbel distribution's parameter estimates are small in magnitude, which is consistent with its very low standard errors. Overall, the consistently low standard errors for the DK distribution across all datasets further support its superior performance and reliability as a suitable model for the analyzed data.

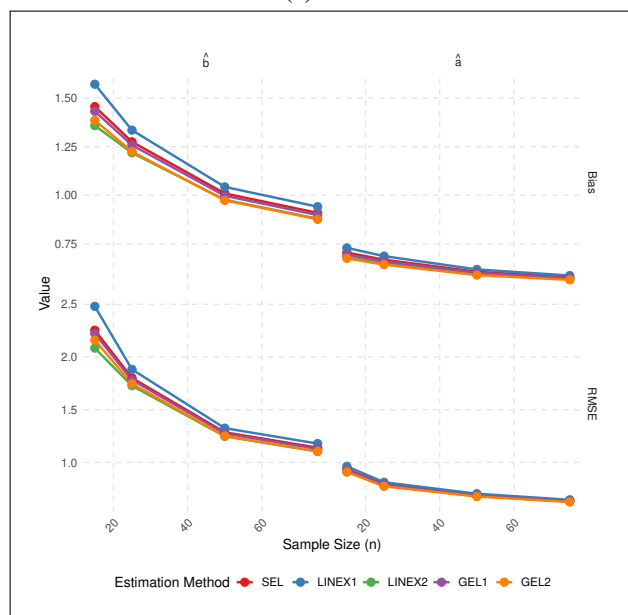
The following inferences are drawn from Figures 11-19.



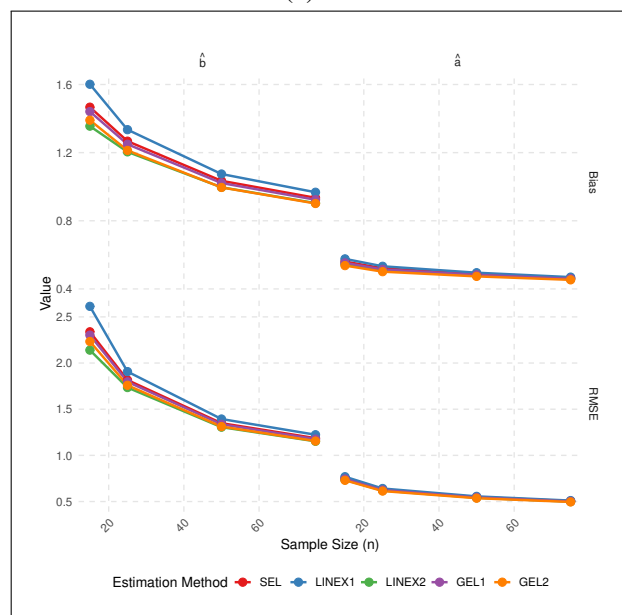
(a)



(b)



(c)



(d)

Fig. 10: Plots of (a) Case I, (b) II, (c) III and (d) IV for Bayesian Simulation Results

iFigure 11 (Box Plot & Violin Plot): Most notably, the violin plot for Data-III reveals a bimodal distribution that the box plot completely obscures.

iiFigure 12 (TTT Plots): These plots effectively differentiate underlying failure rates, showing Data-I with a constant rate, Data-II with an increasing rate, Data-III with a bathtub-shaped rate, and Data-IV with an upside-down bathtub-shaped rate.

iiiFigure 13 (Histograms with Density Plots): All four datasets consistently exhibit right-skewness, with most data concentrated at lower values and a longer tail extending to higher values.

ivFigure 14 (ecdf vs. Theoretical cdf Plots): For all datasets, the empirical cdf closely follows the DK cdf, strongly indicating a good fit of the DK model.

vFigure 15 (Empirical vs. Theoretical Survival Function Plots): The empirical survival function

Table 6: Dominica Weekly death rate due to COVID-19

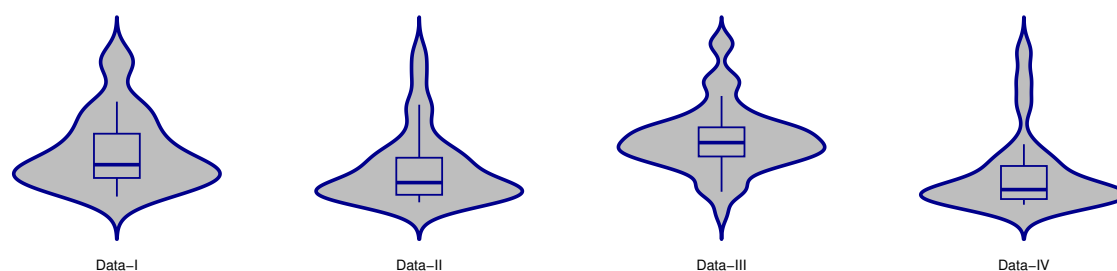
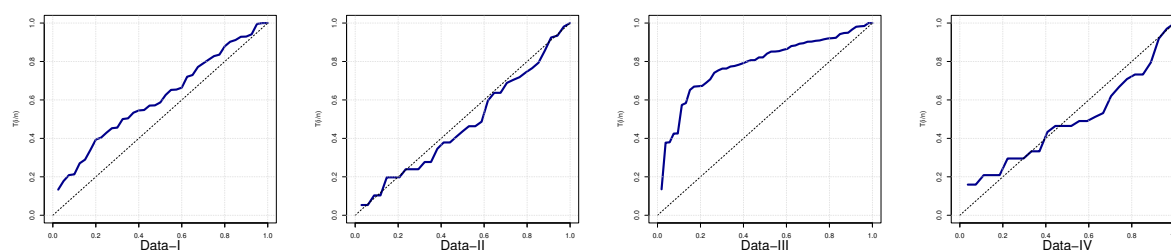
0.029850746	0.035363458	0.05292172	0.051236749	0.053061224	0.035197989
0.032082324	0.025607639	0.01929982	0.012975779	0.016090105	0.016615654
0.012140954	0.024329382	0.01301384	0.012219227	0.013564214	0.010826889
0.008958089	0.010932598	0.016161891	0.02201862	0.023125997	0.037598736
0.033069307	0.02618165	0.019748264	0.015044519	0.011788481	0.009540117
0.009618688	0.008082768	0.008333333	0.006851922	0.005307263	0.005720572
0.004095843	0.003398641	0.002529511	0.004015331		

Table 7: Vinyl chloride data

5.1	1.2	1.3	0.6	0.5	2.4	0.5	1.1	8.0	0.8	0.4	0.6	0.9	0.4	2.0	0.5	5.3
3.2	2.7	2.9	2.5	2.3	1.0	0.2	0.1	0.1	1.8	0.9	2.0	4.0	6.8	1.2	0.4	0.2

Table 8: Weekly volume of traded bitcoin-usd

0.256559779	0.822519052	0.730122407	0.821837757	0.726736545	1.134023831	1.65718773
1.155200566	1.347674948	1.435353611	1.30439894	1.34360108	1.885036149	1.68493482
1.520841715	1.578071262	2.02311774	2.599718197	1.391358248	1.553146029	1.351841874
1.578263954	2.135600415	1.607691456	3.228956661	4.057096933	4.059577506	3.280971217
2.163246258	2.294272601	2.882436754	2.578895559	1.760310882	2.177170118	1.616178571
1.91456349	2.25779718	1.709494099	1.875695007	1.759732716	1.634490154	1.710300939
2.131075314	1.8740702	2.221618762	2.27265185	2.510023388	3.171587346	1.937313931
2.006702741	2.085891553	2.070528442	1.832794312			

**Fig. 11:** Boxplot superimposed on Violin plot for the Datasets**Fig. 12:** TTT plots for the Datasets

consistently aligns very closely with the DK model's curve, confirming the model's suitability for describing survival patterns.

viFigure 16 (P-P Plots): The close alignment of empirical probabilities to the diagonal reference line

across all datasets strongly indicates a good fit of the DK distribution to the data.

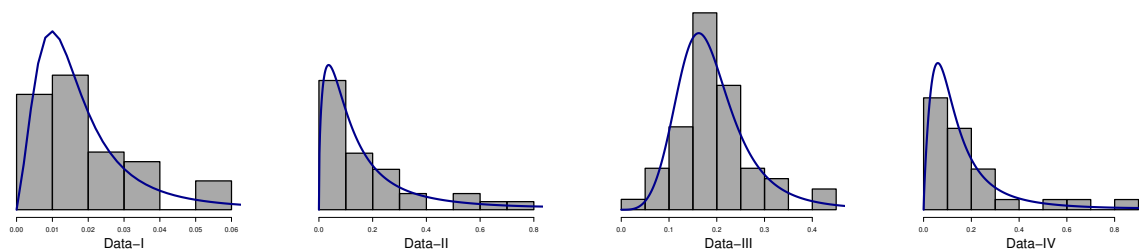
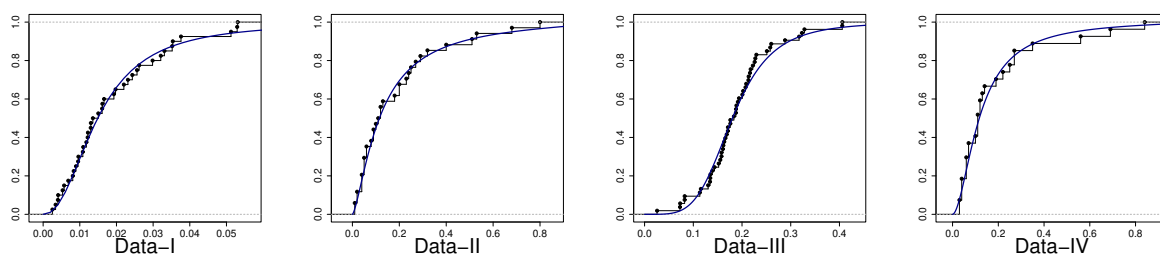
viiFigure 17 (Q-Q Plots): While generally showing a good fit, Data-IV exhibits noticeable deviations at both tails, suggesting the DK model may not fully

Table 9: Infant mortality rate per 1000 live-births

56	10	22	3	69	6	7	11	4	4	19	13	7	27
12	3	4	11	84	27	25	6	35	14	11	12	6	

Table 10: Summary of basic statistics

Statistic	Data I	Data II	Data III	Data IV
n	40	34	53	27
Q_1	0.00939461	0.05	0.1520842	0.06
Q_3	0.02575114	0.2475	0.217717	0.235
IQR	0.01635653	0.1975	0.06563284	0.175
Outlier	0.05292172, 0.05123675, 0.05306122	0.8, 0.68	0.02565598, 0.3228957, 0.4057097, 0.4059578, 0.3280971, 0.3171587	0.56, 0.69, 0.84
Mean	0.018963	0.1879412	0.1900383	0.1881481
Median	0.01430437	0.115	0.1832794	0.11
Variance	0.0001862517	0.03812594	0.005644836	0.04206182
Std. dev.	0.01364741	0.1952586	0.07513212	0.2050898
Range	0.050531709	0.79	0.38030182	0.81
Skewness	1.077715	1.603688	0.7768817	1.953977
Kurtosis	3.416908	5.005408	4.267464	6.050707

**Fig. 13:** Density plot superimposed on Histogram for the Datasets**Fig. 14:** Empirical versus theoretical CDF plots for the Datasets

capture its extreme values as accurately as for other datasets.

Figures 18 and 19 (Likelihood Profile Plots): For both parameters ($\hat{a}$ and $\hat{b}$), the plots show well-defined unimodal curves, indicating that the maximum

likelihood estimates are unique and globally identifiable, suggesting the stability of the parameters.

Table 11: Model performance indicators and goodness of fit tests

Data	Distribution	LL	AIC	CAIC	BIC	HQIC	W	A	KS	p-value
Data I	DK	122.48	-240.9605	-240.6361	-237.5827	-239.7392	0.0291	0.2410	0.0575	0.9984
	Kumaraswamy	122.97	-241.9391	-241.6148	-238.5614	-240.7178	0.0546	0.3521	0.0958	0.8224
	Beta	123.49	-242.9882	-242.6638	-239.6104	-241.7669	0.0358	0.2399	0.0858	0.9059
	Gumbel	121.26	-238.5247	-238.2004	-235.1470	-237.3034	0.0752	0.4787	0.1110	0.6664
	Gamma	123.51	-243.0156	-242.6913	-239.6378	-241.7943	0.0346	0.2328	0.0843	0.9157
	BurrXII	122.98	-241.9562	-241.6319	-238.5785	-240.7349	0.0540	0.3482	0.0951	0.8286
	Weibull	122.97	-241.9478	-241.6235	-238.5701	-240.7266	0.0543	0.3501	0.0955	0.8257
Data II	DK	23.3	-42.6009	-42.2138	-39.5482	-41.5598	0.0316	0.2131	0.0772	0.9873
	Kumaraswamy	22.00	-39.9997	-39.6126	-36.9470	-38.9587	0.0794	0.5235	0.1208	0.7037
	Beta	21.80	-39.6043	-39.2172	-36.5516	-38.5632	0.0848	0.5580	0.1317	0.5974
	Gumbel	16.06	-28.1201	-27.7330	-25.0673	-27.0790	0.1633	1.0409	0.1569	0.3724
	Gamma	22.87	-41.7495	-41.3624	-38.6967	-40.7084	0.0459	0.2975	0.0973	0.9041
	BurrXII	22.94	-41.8900	-41.5029	-38.8372	-40.8489	0.0399	0.2570	0.0843	0.9690
	Weibull	22.84	-41.6766	-41.2895	-38.6238	-40.6355	0.0463	0.3000	0.0918	0.9366
Data III	DK	63.47	-122.9477	-122.7077	-119.0071	-121.4323	0.1040	0.6588	0.0870	0.7851
	Kumaraswamy	62.81	-121.6122	-121.3722	-117.6716	-120.0968	0.1609	0.9613	0.1294	0.3099
	Beta	63.20	-122.3951	-122.1551	-118.4545	-120.8798	0.1381	0.8216	0.1097	0.5110
	Gumbel	63.30	-122.5983	-122.3583	-118.6577	-121.0829	0.1210	0.7350	0.1062	0.5527
	Gamma	62.83	-121.6676	-121.4276	-117.7271	-120.1523	0.1416	0.8399	0.1068	0.5457
	BurrXII	63.07	-122.1426	-121.9026	-118.2020	-120.6272	0.1501	0.8996	0.1248	0.3520
	Weibull	62.95	-121.9010	-121.6610	-117.9604	-120.3857	0.1550	0.9276	0.1269	0.3325
Data IV	DK	21.02	-38.0379	-37.5379	-35.4462	-37.2673	0.0549	0.3725	0.0887	0.9837
	Kumaraswamy	16.33	-28.6582	-28.1582	-26.0665	-27.8876	0.2081	1.3004	0.1991	0.2349
	Beta	16.15	-28.2971	-27.7971	-25.7055	-27.5265	0.2145	1.3361	0.2136	0.1703
	Gumbel	13.76	-23.5159	-23.0159	-20.9242	-22.7452	0.2218	1.3703	0.2015	0.2228
	Gamma	18.58	-33.1615	-32.6615	-30.5698	-32.3908	0.1239	0.8064	0.1726	0.3974
	BurrXII	18.77	-33.5420	-33.0420	-30.9504	-32.7714	0.1110	0.7247	0.1501	0.5775
	Weibull	18.23	-32.4575	-31.9575	-29.8658	-31.6869	0.1306	0.8475	0.1615	0.4817

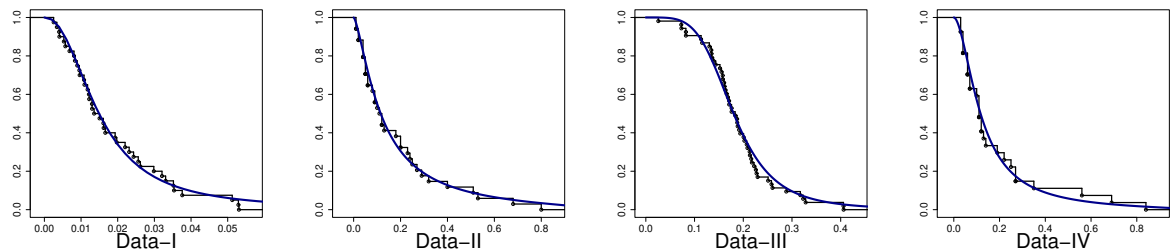
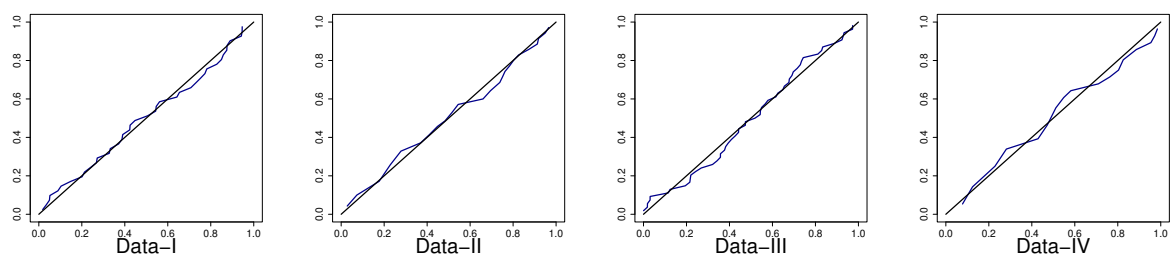
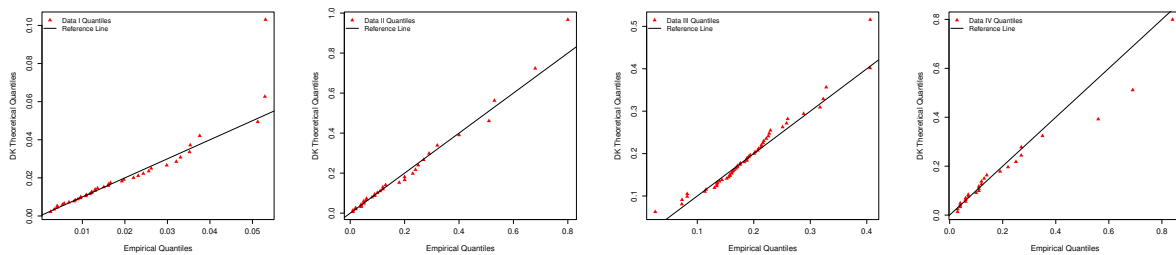
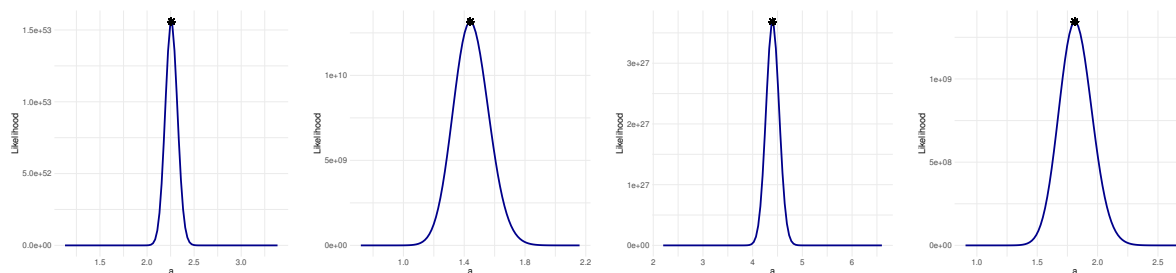
**Fig. 15:** Empirical versus theoretical survival function plots for the Datasets**Fig. 16:** P-P plots for the Datasets

Table 12: MLEs and standard errors for the parameters

Data	Distribution	$\hat{a}$ (standard error)	$\hat{b}$ (standard error)
Data I	DK	2.2573(0.2949)	9.4967(1.2620)
	Kumaraswamy	1.4807(0.1809)	5.7113(0.6640)
	Beta	2.0099(0.4175)	103.9624(24.4471)
	Gumbel	0.0130(0.0016)	0.0095(0.0012)
	Gamma	2.0449(0.4252)	107.8358(25.3974)
	BurrXII	1.4854(0.1782)	309.4328(201.5043)
	Weibull	0.0211(0.0024)	1.4830(0.7971)
Data II	DK	1.4394(0.2290)	3.0880(0.5884)
	Kumaraswamy	0.8366(0.1407)	1.1464(0.2708)
	Beta	0.82363(0.1742)	3.3498(0.8692)
	Gumbel	0.1092(0.0209)	0.1170(0.0174)
	Gamma	1.0627(0.2281)	5.6543(1.5357)
	BurrXII	1.0932(0.1325)	6.9315(1.5800)
	Weibull	0.1888(0.0339)	1.0102(0.1327)
Data III	DK	4.3998(0.5220)	7.5392(0.9176)
	Kumaraswamy	2.6413(0.2771)	4.0658(0.4098)
	Beta	4.8456(0.9114)	20.6689(4.0457)
	Gumbel	0.1553(0.0095)	0.0656(0.0066)
	Gamma	5.8668(1.1088)	30.8720(6.0916)
	BurrXII	2.6900(0.2634)	64.7313(24.7615)
	Weibull	0.2133(0.0116)	2.6660(0.2697)
Data IV	DK	1.8118(0.3110)	3.8780(0.7728)
	Kumaraswamy	0.8748(0.1596)	1.1614(0.2957)
	Beta	0.8941(0.2131)	3.4259(0.9879)
	Gumbel	0.1110(0.0216)	0.1086(0.0186)
	Gamma	1.2802(0.3133)	6.8041(2.0278)
	BurrXII	1.1685(0.1503)	7.5892(1.9310)
	Weibull	0.1943(0.0370)	1.0739(0.1504)

**Fig. 17:** Q-Q plots for the Datasets**Fig. 18:** Likelihood Profile for $\hat{a}$ for Data I, II, III and IV respectively

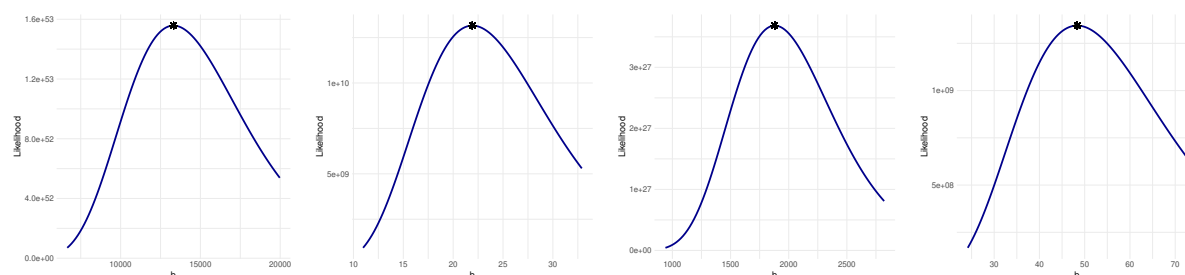


Fig. 19: Likelihood Profile for $\hat{b}$ for Data I, II, II and IV respectively

8 Final Comments and Future Work

This study introduced a new unit family of distributions called the Dhillon-G family that uses a reduced Dhillon distribution as a transformer in the T-X generator. Different generalizers were suggested in Table 1 for future studies. The Kumaraswamy distribution was further used as a parent distribution to determine the usefulness of the new family in enhancing bounded distributions, leading to a new distribution referred to DK. The structural properties of the Dk distribution were studied with evidence of DK distribution capable of modeling both unimodal and bimodal data. The parameters of the DK model were estimated using MLE, MPS, LS and WLS methods as well as Bayesian estimation under SEL, LINEX and GEL functions. A comprehensive simulation studies were carried out at different parameter settings and sample sizes, and results reveal consistent asymptotic behavior for small and large sample sizes. Four real data sets were deployed to demonstrate the applicability.

Future research can pursue this work along many different directions. A promising direction is to compare the performance of other generalizers found in Table 1. By employing a different parent distribution, such as the Beta or Topp-Leone distribution, from the Dhillon-G family may yield new, highly flexible models targeted for specific kinds of data. Furthermore, an analysis of the stress-strength reliability of the DK distribution would be useful, particularly for engineering and quality control applications where the stress and strength of components are significant. Another future research area is to extend the current model to a multivariate or bivariate model. Construction of a bivariate Dhillon-G family would allow researchers to represent the relationship between two or more bounded variables that are correlated with each other, which is crucial in finance and medicine. The DK distribution parameters may even serve as regressors in a regression model so that the model can receive covariate information. This would render the DK distribution no longer an independent model but a general regression tool.

Lastly, although in this research we worked with complete data, subsequent studies can benefit from

generalizing the DK distribution to other censoring schemes, such as Type I or progressive censoring. This is very important in the case of reliability and life time data when complete observations are not possible. Implementing these ideas will not only add depth to the theoretical structure of the Dhillon-G family but also significantly increase its applicability in many fields of science.

The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

- [1] B. S. Dhillon, "Statistical functions to represent various types of hazard rates," *Microelectronics Reliability*, vol. 20, no. 5, pp. 581–584, 1980. DOI: [https://doi.org/10.1016/0026-2714\(80\)90386-8](https://doi.org/10.1016/0026-2714(80)90386-8)
- [2] A. Alzaatreh, C. Lee, and F. Famoye, "A new method for generating families of continuous distributions," *Metron*, vol. 71, no. 1, pp. 63–79, 2013. URL: <https://doi.org/10.1007/s40300-013-0007-y>
- [3] J. W. Strutt, "On the resultant of a large number of vibrations of the same pitch and of arbitrary phase," *Philosophical Magazine*, vol. 10, no. 60, pp. 73–78, 1880. DOI: <https://doi.org/10.1080/14786448008626893>
- [4] J. J. Swain, S. Venkatraman, and J. R. Wilson, "Least-squares estimation of distribution functions in Johnson's translation system," *Journal of Statistical Computation and Simulation*, vol. 29, no. 4, pp. 271–297, 1988. URL: <https://doi.org/10.1080/00949658808811068>
- [5] R. Cheng and N. Amin, "Maximum product of spacings estimation with application to the lognormal distribution (Mathematical Report 79-1)," *Cardiff: University of Wales IST*, 1979.
- [6] E. Q. Chinedu, Q. C. Chukwudum, N. Alsadat, O. J. Obulezi, E. M. Almetwally, and A. H. Tolba, "New lifetime distribution with applications to single acceptance sampling plan and scenarios of increasing hazard rates," *Symmetry*, vol. 15, no. 10, pp. 1881, 2023. DOI: <https://doi.org/10.3390/sym15101881>
- [7] D. K. Bhaumik, K. Kapur, and R. D. Gibbons, "Testing parameters of a gamma distribution for small samples," *Technometrics*, vol. 51, no. 3, pp. 326–334, 2009. DOI: <https://doi.org/10.1198/tech.2009.07038>

- [8] A. H. Tolba, C. K. Onyekwere, A. R. El-Saeed, N. Alsadat, H. Alohal, and O. J. Obulezi, "A New Distribution for Modeling Data with Increasing Hazard Rate: A Case of COVID-19 Pandemic and Vinyl Chloride Data," *Sustainability*, vol. 15, no. 17, pp. 12782, 2023. DOI: <https://doi.org/10.3390/su151712782>
- [9] B. C. Nwankwo, H. O. Obiora-Ilouno, F. A. Almulhim, M. S. Mustafa, and O. J. Obulezi, "Group acceptance sampling plans for type-I heavy-tailed exponential distribution based on truncated life tests," *AIP Advances*, vol. 14, no. 3, 2024. DOI: <https://doi.org/10.1063/5.0194258>
- [10] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions, volume 2*, vol. 2, John Wiley & sons, 1995.
- [11] N. L. Johnson, "Systems of frequency curves generated by methods of translation," *Biometrika*, vol. 36, no. 1/2, pp. 149–176, 1949. DOI: <https://doi.org/10.2307/2332539>
- [12] P. Kumaraswamy, "A generalized probability density function for double-bounded random processes," *Journal of hydrology*, vol. 46, no. 1-2, pp. 79–88, 1980. DOI: [https://doi.org/10.1016/0022-1694\(80\)90036-0](https://doi.org/10.1016/0022-1694(80)90036-0)
- [13] I. W. Burr, "Cumulative frequency functions," *The Annals of mathematical statistics*, vol. 13, no. 2, pp. 215–232, 1942. DOI: <https://www.jstor.org/stable/2235756>
- [14] W. Weibull, "A statistical theory of strength of materials," *IVB-Handl.*, 1939. URL: <https://cir.nii.ac.jp/crid/1573668925268278656>
- [15] E. J. Gumbel, "The return period of flood flows," *The annals of mathematical statistics*, vol. 12, no. 2, pp. 163–190, 1941.
- [16] M. Muhammad, "A new two parameter distribution with a finite support," *Applied Mathematical Sciences*, vol. 10, pp. 1–12, 2016.
- [17] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4, Springer, 2006.
- [18] F. J. Fabozzi, S. M. Focardi, and P. N. Kolm, *Financial modeling of the equity market: from CAPM to cointegration*, John Wiley & Sons, 2006.
- [19] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, vol. 2, Springer, 2009.
- [20] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. Siegelbaum, A. J. Hudspeth, S. Mack, and others, *Principles of neural science*, vol. 4, McGraw-hill New York, 2000.
- [21] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965. DOI: [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- [22] J. Atchison and S. M. Shen, "Logistic-normal distributions: Some properties and uses," *Biometrika*, vol. 67, no. 2, pp. 261–272, 1980. DOI: <https://doi.org/10.1093/biomet/67.2.261>
- [23] E. Alshawarbeh, F. Famoye, and C. Lee, "Beta-Cauchy distribution: some properties and applications," *Journal of Statistical Theory and Applications*, vol. 12, no. 4, pp. 378–391, 2013. DOI: <https://doi.org/10.2991/jsta.2013.12.4.5>
- [24] C. Dagum, "A new model of personal income distribution: specification and estimation," in *Modeling income distributions and Lorenz curves*, Springer, pp. 3–25, 2008. DOI: https://doi.org/10.1007/978-0-387-72796-7_1
- [25] J. B. McDonald and Y. J. Xu, "A generalization of the beta distribution with applications," *Journal of Econometrics*, vol. 66, no. 1-2, pp. 133–152, 1995. DOI: [https://doi.org/10.1016/0304-4076\(94\)01612-4](https://doi.org/10.1016/0304-4076(94)01612-4)
- [26] M. Bourguignon, R. B. Silva, and G. M. Cordeiro, "The Weibull-G family of probability distributions," *Journal of data science*, vol. 12, no. 1, pp. 53–68, 2014.
- [27] J. Mazucheli, A. F. B. Menezes, and M. E. Ghitany, "The unit-Weibull distribution and associated inference," *J. Appl. Probab. Stat*, vol. 13, no. 2, pp. 1–22, 2018.
- [28] M. Ç. Korkmaz and Z. S. Korkmaz, "The unit log-log distribution: A new unit distribution with alternative quantile regression modeling and educational measurements applications," *Journal of Applied Statistics*, vol. 50, no. 4, pp. 889–908, 2023. DOI: <https://doi.org/10.1080/02664763.2021.2001442>
- [29] M. E. Ghitany, J. Mazucheli, A. F. B. Menezes, and F. Alqallaf, "The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval," *Communications in Statistics-Theory and methods*, vol. 48, no. 14, pp. 3423–3438, 2019. DOI: <https://doi.org/10.1080/03610926.2018.1476717>
- [30] J. Mazucheli, A. F. 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>
- [31] K. Karakaya, C. S. Rajitha, Ş. Sağlam, Y. A. Tashkandy, M. E. Bakr, A. H. Muse, A. Kumar, E. Hussam, and A. M. Gemeay, "A new unit distribution: properties, estimation, and regression analysis," *Scientific Reports*, vol. 14, no. 1, pp. 7214, 2024. DOI: <https://doi.org/10.1038/s41598-024-57390-7>
- [32] A. Krishna, R. Maya, C. Chesneau, and M. R. Irshad, "The unit Teissier distribution and its applications," *Mathematical and Computational Applications*, vol. 27, no. 1, pp. 12, 2022. DOI: <https://doi.org/10.3390/mca27010012>
- [33] J. Mazucheli, A. F. B. Menezes, and S. Dey, "The unit-Birnbaum-Saunders distribution with applications," *Chilean Journal of Statistics*, vol. 9, no. 1, pp. 47–57, 2018.
- [34] V. S. Stojanović, M. Jovanović, B. Pažun, Z. Langović, and Ž. Grujić, "Gumbel-Logistic Unit Distribution with Application in Telecommunications Data Modeling," *Symmetry*, vol. 16, no. 11, pp. 1513, 2024. DOI: <https://doi.org/10.3390/sym16111513>
- [35] M. C. Korkmaz, E. Altun, C. Chesneau, and H. M. Yousof, "On the unit-Chen distribution with associated quantile regression and applications," *Mathematica Slovaca*, vol. 72, no. 3, pp. 765–786, 2022. DOI: <https://doi.org/10.1515/ms-2022-0052>
- [36] M. Ç. Korkmaz and C. Chesneau, "On the unit Burr-XII distribution with the quantile regression modeling and applications," *Computational and Applied Mathematics*, vol. 40, no. 1, pp. 29, 2021. DOI: <https://doi.org/10.1007/s40314-021-01418-5>
- [37] J. Mazucheli, A. F. B. Menezes, and S. Chakraborty, "On the one parameter unit-Lindley distribution and its associated regression model for proportion data," *Journal of Applied Statistics*, vol. 46, no. 4, pp. 700–714, 2019. DOI: <https://doi.org/10.1080/02664763.2018.1511774>
- [38] S. Brooks, "Markov chain Monte Carlo method and its application," *Journal of the royal statistical society: series D (the Statistician)*, vol. 47, no. 1, pp. 69–100, 1998. URL: <https://doi.org/10.1111/1467-9884.00117>
- [39] D. Van Ravenzwaaij, P. Cassey, and S. D. Brown, "A simple introduction to Markov Chain Monte-Carlo sampling,"

- Psychonomic bulletin & review*, vol. 25, no. 1, pp. 143–154, 2018. URL: <https://cir.nii.ac.jp/crid/1572824500518986496>
- [40] Prakash, A., Maurya, R. K., Alsadat, N., & Obulezi, O. J. (2024). Parameter estimation for reduced Type-I Heavy-Tailed Weibull distribution under progressive Type-II censoring scheme. *Alexandria Engineering Journal*, 109, 935–949. doi: [10.1016/j.aej.2024.09.029](https://doi.org/10.1016/j.aej.2024.09.029)
- [41] Noori, N. A., Khaleel, M. A., Khalaf, S. A., & Dutta, S. (2025). Analytical modeling of expansion for odd Lomax generalized exponential distribution in framework of neutrosophic logic: A theoretical and applied on neutrosophic data. *Innovation in Statistics and Probability*, 1(1), 47–59.
- [42] Husain, Q. N., Qaddoori, A. S., Noori, N. A., Abdullah, K. N., Suleiman, A. A., & Balogun, O. S. (2025). New expansion of Chen distribution according to the neutrosophic logic using the Gompertz family. *Innovation in Statistics and Probability*, 1(1), 60–75.
- [43] El Gazar, A. M., Ramadan, D. A., ElGarhy, M., and El-Desouky, B. S. "Estimation of parameters for inverse power Ailamujia and truncated inverse power Ailamujia distributions based on progressive type-II censoring scheme." *Innovation in Statistics and Probability*, 1(1):76–87, 2025.
- [44] Orji, G. O., Etaga, H. O., Almetwally, E. M., Igbokwe, C. P., Aguwa, O. C., and Obulezi, O. J. "A new odd reparameterized exponential transformed-X family of distributions with applications to public health data." *Innovation in Statistics and Probability*, 1(1):88–118, 2025.
- [45] Sapkota, L. P., Kumar, V., Tekle, G., Alrweili, H., Mustafa, M. S., & Yusuf, M. (2025). Fitting real data sets by a new version of Gompertz distribution. *Modern Journal of Statistics*, 1(1), 25–48.
- [46] M. A., M., Al-Kandari, N. M., & Mohammad, R. Z. (2025). Fundamental properties of the characteristic function using the compound Poisson distribution as the sum of the gamma model. *Modern Journal of Statistics*, 1(1), 49–57.
- [47] Noori, N. A., Abdullah, K. N., & Khaleel, M. A. (2025). Development and applications of a new hybrid Weibull-inverse Weibull distribution. *Modern Journal of Statistics*, 1(1), 80–103.
- [48] Ragab, I., & Elgarhy, M. (2025). Type II half logistic Ailamujia distribution with numerical illustrations to medical data. *Computational Journal of Mathematical and Statistical Sciences*, (), -. doi: [10.21608/cjmss.2025.346849.1095](https://doi.org/10.21608/cjmss.2025.346849.1095)
- [49] Elgarhy, M., Abdalla, G., Hassan, A., & Almetwally, E. (2025). Bayesian and Non-Bayesian Analysis of the Novel Unit Inverse Exponentiated Lomax Distribution Using Progressive Censoring Schemes with Optimal Scheme and Data Application. *Computational Journal of Mathematical and Statistical Sciences*, (), -. doi: [10.21608/cjmss.2025.374277.1151](https://doi.org/10.21608/cjmss.2025.374277.1151)
- [50] Obulezi, O. J., Obiora-Iloano, H. O., Osuji, G. A., Kayid, M., & Balogun, O. S. (2025). A new family of generalized distributions based on logistic-x transformation: sub-model, properties and useful applications. *Research in Statistics*, 3(1), 2477232. doi: [10.1080/27684520.2025.2477232](https://doi.org/10.1080/27684520.2025.2477232)
- [51] Chinedu, E. Q., Chukwudum, Q. C., Alsadat, N., Obulezi, O. J., Almetwally, E. M., & Tolba, A. H. (2023). New lifetime distribution with applications to single acceptance sampling plan and scenarios of increasing hazard rates. *Symmetry*, 15(10), 1881. doi: [10.3390/sym15101881](https://doi.org/10.3390/sym15101881)
- [52] Nwankwo, B. C., Obiora-Iloano, H. O., Almulhim, F. A., SidAhmed Mustafa, M., & Obulezi, O. J. (2024). Group acceptance sampling plans for type-I heavy-tailed exponential distribution based on truncated life tests. *AIP Advances*, 14(3). doi: [10.1063/5.0194258](https://doi.org/10.1063/5.0194258)
- [53] Nwankwo, M. P., Alsadat, N., Kumar, A., Bahloul, M. M., & Obulezi, O. J. (2024). Group acceptance sampling plan based on truncated life tests for Type-I heavy-tailed Rayleigh distribution. *Heliyon*, 10(19), e38150. doi: [10.1016/j.heliyon.2024.e38150](https://doi.org/10.1016/j.heliyon.2024.e38150)
- [54] El-Saeed, A. R., Obulezi, O. J., & Abd El-Raouf, M. M. (2025). Type II heavy tailed family with applications to engineering, radiation biology and aviation data. *Journal of Radiation Research and Applied Sciences*, 18(3), 101547. doi: [10.1016/j.jrras.2025.101547](https://doi.org/10.1016/j.jrras.2025.101547)
- [55] Nadir, S., Aslam, M., Anyiam, K. E., Alshawarbeh, E., & Obulezi, O. J. (2025). Group acceptance sampling plan based on truncated life tests for the Kumaraswamy Bell-Rayleigh distribution. *Scientific African*, 27, e02537. doi: [10.1016/j.sciaf.2025.e02537](https://doi.org/10.1016/j.sciaf.2025.e02537)
- [56] Obulezi, O. J., Obiora-Iloano, H. O., Osuji, G. A., Kayid, M., & Balogun, O. S. (2025). Weibull Sine Generalized Distribution Family: Fundamental Properties, Sub-model, Simulations, with Biomedical Applications. *Electronic Journal of Applied Statistical Analysis*, 18(01), 183–212. doi: [10.1285/i20705948v18n1p183](https://doi.org/10.1285/i20705948v18n1p183)
- [57] Hussein, M., Rodrigues, G. M., Ortega, E. M. M., Vila, R., & Elsayed, H. (2023). A New Truncated Lindley-Generated Family of Distributions: Properties, Regression Analysis, and Applications. *Entropy*, 25(9), 1359. doi: [10.3390/e25091359](https://doi.org/10.3390/e25091359)
- [58] Saeed, A., Saboor, A., Jamal, F., Alsadat, N., Balogun, O. S., Faal, A., & Elgarhy, M. (2025). Bounded sine hyperbolic distribution with applications to real datasets. *Kuwait Journal of Science*, 100467. doi: [10.1016/j.kjs.2025.100467](https://doi.org/10.1016/j.kjs.2025.100467)
- [59] Almetwally, E. M., Shah, Z., Ozonur, D., Khan, D. M., Jamal, F., El-Saeed, A. R., & Shaaban, M. (2025). A new flexible power-X family of distributions: Applications to reliability engineering data. *Journal of Radiation Research and Applied Sciences*, 18(2), 101525. doi: [10.1016/j.jrras.2025.101525](https://doi.org/10.1016/j.jrras.2025.101525)
- [60] Baaqeel, H., Alnashri, H., & Baharith, L. (2025). A New Lomax-G Family: Properties, Estimation and Applications. *Entropy*, 27(2), 125. doi: [10.3390/e27020125](https://doi.org/10.3390/e27020125)
- [61] Mohammad, S., & Mendoza, I. (2025). A New Hyperbolic Tangent Family of Distributions: Properties and Applications. *Annals of Data Science*, 12(2), 457–480. doi: [10.1007/s40745-024-00516-5](https://doi.org/10.1007/s40745-024-00516-5)
- [62] Onyekwere, C. K., Aguwa, O. C., & Obulezi, O. J. (2025). An updated Lindley distribution: Properties, estimation, acceptance sampling, actuarial risk assessment and applications. *Innovation in Statistics and Probability*, 1(1), 1–27.
- [63] Saboor, A., Jamal, F., Shafq, S., & Mumtaz, R. (2025). On the versatility of the unit logistic exponential distribution: Capturing bathtub, upside-down bathtub, and monotonic hazard rates. *Innovation in Statistics and Probability*, 1(1), 28–46.
- [64] Gemeay, A. M., Moakofi, T., Balogun, O. S., Ozkan, E., & Hossain, M. M. (2025). Analyzing real data by a new heavy-tailed statistical model. *Modern Journal of Statistics*, 1(1), 1–24.