Int. J. Anal. Appl. (2026), 24:383

International Journal of Analysis and Applications

A Flexible Sine-Transformed Extension of the Dagum Distribution

Ibrahim Abdullahi[1], **Wikanda Phaphan**[2,3,*]

[1]*Department of Mathematics and Statistics, Faculty of Science, Yobe State University, Damaturu, 500501, Nigeria*

[2]*Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand*

[3]*Research Group in Statistical Learning and Inference, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand*

[*]*Corresponding author:* wikanda.p@sci.kmutnb.ac.th

Abstract. This paper introduces a sine-transformed extension of the Dagum distribution, referred to as the SS–Dagum model, obtained by applying a bounded trigonometric transformation to the baseline cumulative distribution function. The proposed distribution preserves the heavy-tailed characteristics of the classical Dagum model while providing enhanced flexibility to capture diverse shapes in the density and hazard rate functions. Key statistical properties of the SS–Dagum distribution are derived, including closed-form expressions for the probability density, survival, and hazard rate functions. A hazard modulation representation relative to the baseline Dagum model is established, offering insight into the effect of the transformation on reliability behavior. In addition, an explicit upper-tail asymptotic result is obtained, showing that the proposed model preserves regular variation while modifying the tail index. Parameter estimation is performed using the maximum likelihood method, and the finite-sample performance of the estimators is assessed through Monte Carlo simulation. The practical usefulness of the SS–Dagum distribution is demonstrated through real data applications, where it provides improved fit compared to classical competing models. Overall, the SS–Dagum distribution offers a flexible and tractable alternative for modeling heavy-tailed and positively skewed data arising in economics, reliability, and related fields.

1. Introduction

Heavy-tailed probability distributions play a fundamental role in modeling income and wealth data, actuarial risk, finance, and reliability systems. Such data are typically characterized by strong

Received: Mar. 26, 2026.

2020 *Mathematics Subject Classification.* 62E15, 62F10, 62N01, 60E05.

Key words and phrases. Dagum distribution; sine-transformed distribution; heavy-tailed distribution; maximum likelihood estimation; hazard rate function; survival analysis.

https://doi.org/10.28924/2291-8639-24-2026-383
ISSN: 2291-8639

© 2026 the author(s).

right-skewness, high variability, and the presence of extreme observations, which cannot be adequately captured by classical light-tailed models. Among the available heavy-tailed models, the Dagum distribution [3] has gained considerable prominence due to its ability to represent income distributions with Pareto-type tails and flexible shape characteristics. It has been widely applied in economics, insurance, and survival analysis, particularly for modeling income inequality and lifetime data [6]. Despite its usefulness, the classical Dagum distribution may exhibit limitations when modeling datasets with complex curvature, varying hazard rate structures, or subtle deviations in tail behaviour. In response to these challenges, a large body of literature has focused on constructing generalized families of distributions through transformation or generator mechanisms. Notable examples include the beta-generated and Kumaraswamy-generated families, which introduce additional shape parameters to enhance flexibility [1,2].

However, despite these advances, existing sine- or trigonometric-generated families often suffer from one or more of the following limitations: (i) increased model complexity due to additional parameters, (ii) lack of clear interpretability of the transformed parameters, and (iii) limited insight into how the transformation affects tail behaviour and reliability structure. In particular, there is still a need for transformation-based extensions that preserve the heavy-tailed nature of the baseline model while providing interpretable modifications to the hazard rate and tail index.

Motivated by these gaps, we propose a new sine-transformed Dagum distribution, referred to as the SS–Dagum model. The proposed model is constructed by applying a bounded sine transformation to the cumulative distribution function of the classical Dagum distribution. Unlike many existing generator-based models, the SS–Dagum distribution retains the original parameter structure of the baseline model while introducing an additional layer of flexibility through the transformation. This approach allows for controlled modification of the distributional shape, including skewness, kurtosis, and hazard rate behaviour, without substantially increasing model complexity.

The novelty of the proposed SS–Dagum distribution lies in three key aspects. First, the sine transformation provides a bounded and smooth mechanism for reshaping the distribution, ensuring that the resulting model remains well-defined and analytically tractable. Second, we establish a hazard modulation representation that explicitly links the hazard function of the proposed model to that of the baseline Dagum distribution, offering clear interpretability of the transformation effect. Third, we derive an explicit upper-tail asymptotic result, showing that the proposed model preserves regular variation while modifying the tail index, thereby retaining heavy-tailed behaviour with enhanced flexibility.

The main contributions of this paper are summarized as follows. We define the SS–Dagum distribution and derive its fundamental statistical properties, including the probability density, cumulative distribution, survival, and hazard rate functions. We develop parameter estimation

procedures based on the maximum likelihood method and evaluate their finite-sample performance via Monte Carlo simulation. Finally, we demonstrate the practical usefulness of the proposed model through applications to real datasets and show that it provides improved fit compared to classical competing distributions.

2. The Classical Dagum Distribution

Let $X \sim \text{Dagum}(a, b, p)$ with $a > 0$, $b > 0$, and $p > 0$. The cdf is

$$G(x) = \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}, \qquad x > 0, \tag{2.1}$$

and the corresponding pdf is

$$g(x) = \frac{ap}{b}\left(\frac{x}{b}\right)^{-a-1}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p-1}, \qquad x > 0. \tag{2.2}$$

3. The SS–Dagum Distribution

Definition 3.1. *A random variable X is said to follow the* SS–Dagum *distribution, denoted $X \sim$ SS–Dagum(a, b, p), if its cdf is*

$$F(x) = \sin\left(\frac{\pi}{2}G(x)\right), \qquad x > 0, \tag{3.1}$$

where $G(x)$ is the Dagum cdf in (2.1).

$$F(x) = \sin\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right), \qquad x > 0. \tag{3.2}$$

3.1. Density and validity.

Differentiating (3.1) yields the pdf

$$f(x) = \frac{\pi}{2}\, g(x)\, \cos\left(\frac{\pi}{2}G(x)\right), \qquad x > 0, \tag{3.3}$$

where $g(x)$ is the Dagum pdf in (2.2). Substituting the explicit form of $g(x)$ gives

$$f(x) = \frac{\pi}{2}\frac{ap}{b}\left(\frac{x}{b}\right)^{-a-1}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p-1}\cos\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right), \qquad x > 0. \tag{3.4}$$

Theorem 3.1 (Validity)**.** *The function f in* (3.3) *is a valid probability density function.*

Proof. For $x > 0$, $0 < G(x) < 1$, hence $0 < F(x) < 1$. Moreover, $f(x) = \frac{d}{dx}F(x)$ and

$$\int_0^\infty f(x)\, dx = F(\infty) - F(0) = \sin\left(\frac{\pi}{2}\cdot 1\right) - \sin(0) = 1.$$

□

4. Reliability Properties

This section derives the survival and hazard rate functions of the SS–Dagum distribution in fully explicit form and establishes its structural hazard modulation and tail behavior properties.

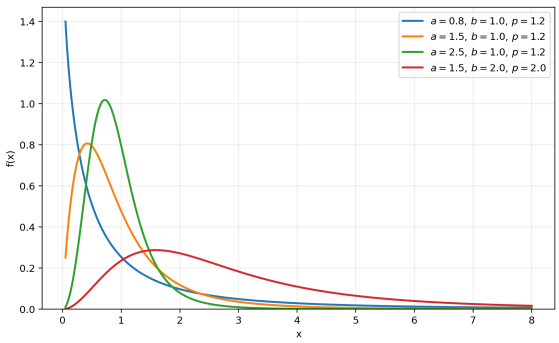


FIGURE 1. PDF

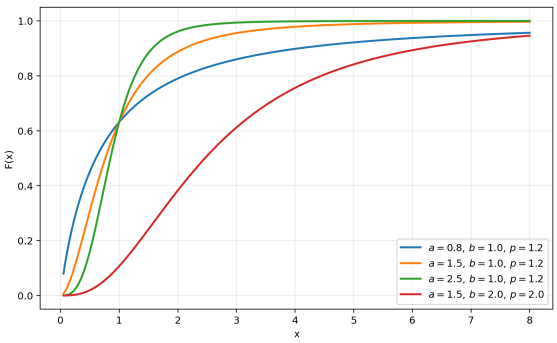


FIGURE 2. CDF

4.1. **Survival Function.** Recall that for $X \sim \text{SS–Dagum}(a, b, p)$, the survival function is

$$S(x) = 1 - \sin\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right), \qquad x > 0. \tag{4.1}$$

4.2. **Hazard Rate Function.** The SS–Dagum density is

$$\begin{aligned} f(x) = \frac{\pi}{2}\frac{ap}{b}\left(\frac{x}{b}\right)^{-a-1}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p-1} \\ \times \cos\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right), \qquad x > 0. \end{aligned} \tag{4.2}$$

Therefore, the hazard rate function $h(x) = f(x)/S(x)$ becomes

$$h(x) = \frac{\frac{\pi}{2}\frac{ap}{b}\left(\frac{x}{b}\right)^{-a-1}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p-1}\cos\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right)}{1 - \sin\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right)}, \qquad x > 0. \tag{4.3}$$

This explicit expression shows how the trigonometric factor alters the baseline Dagum risk dynamics.

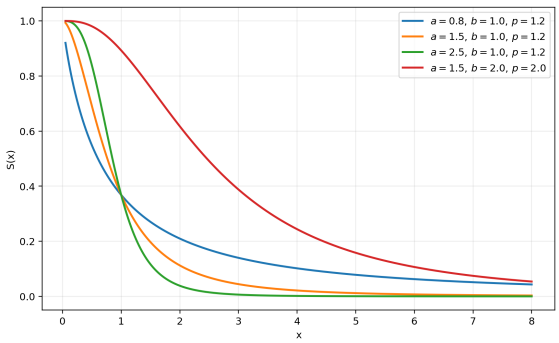


FIGURE 3. Survival Function

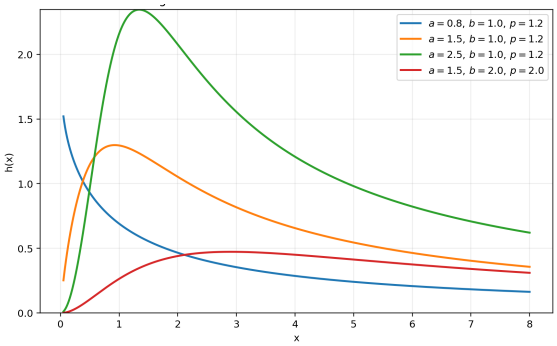


FIGURE 4. Hazard Function

4.3. **Structural Hazard Modulation.** The baseline Dagum survival function is

$$\overline{G}(x) = 1 - \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}.$$

The baseline Dagum hazard function is

$$h_D(x) = \frac{\frac{ap}{b}\left(\frac{x}{b}\right)^{-a-1}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p-1}}{1 - \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}}.$$

Theorem 4.1 (Hazard modulation representation)**.** *Let* $Y \sim$ SS–Dagum(a, b, p) *and* $X \sim$ Dagum(a, b, p). *Then*

$$h_{SS}(x) = h_D(x)\, R(x),$$

where

$$R(x) = \frac{\pi}{2}\overline{G}(x)\frac{\cos\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right)}{1 - \sin\left(\frac{\pi}{2}\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right)}.$$

Proof. Substituting the explicit forms of $f(x)$ and $S(x)$ into $h_{SS}(x) = f(x)/S(x)$ and factorizing the baseline hazard $h_D(x)$ yields the result. □

4.4. **Tail Behavior and Asymptotic Structure.** Let $G(x)$ denote the cumulative distribution function of the Dagum distribution. Then its survival function is

$$\overline{G}(x) = 1 - G(x) = 1 - \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}.$$

Since the SS–Dagum distribution is defined by

$$F(x) = \sin\left(\frac{\pi}{2}G(x)\right),$$

we can write

$$F(x) = \sin\left(\frac{\pi}{2}\left(1 - \overline{G}(x)\right)\right) = \cos\left(\frac{\pi}{2}\overline{G}(x)\right).$$

Therefore, the survival function of the SS–Dagum distribution is

$$\overline{F}(x) = 1 - F(x) = 1 - \cos\left(\frac{\pi}{2}\overline{G}(x)\right).$$

Theorem 4.2 (Tail behavior)**.** *As* $x \to \infty$,

$$\overline{F}(x) \sim \frac{\pi^2}{8}\,\overline{G}(x)^2.$$

In particular, since

$$\overline{G}(x) \sim Cx^{-ap},$$

it follows that

$$\overline{F}(x) \sim C^* x^{-2ap},$$

where $C^* = \frac{\pi^2}{8}C^2$. *Hence, the tail index changes from ap (Dagum) to* 2*ap (SS–Dagum).*

Proof. Let

$$t = \frac{\pi}{2}\overline{G}(x).$$

As $x \to \infty$, we have $\overline{G}(x) \to 0$, hence $t \to 0$.

Using the standard Taylor expansion

$$1 - \cos t \sim \frac{t^2}{2}, \quad t \to 0,$$

we obtain

$$\overline{F}(x) = 1 - \cos\Big(\frac{\pi}{2}\overline{G}(x)\Big) \sim \frac{1}{2}\Big(\frac{\pi}{2}\overline{G}(x)\Big)^2 = \frac{\pi^2}{8}\,\overline{G}(x)^2.$$

□

4.5. **Moments: existence and an explicit Beta-function representation.**

Theorem 4.3. *Let $Y \sim$ SS–Dagum(a, b, p). Then*

$$\mathbb{E}(Y^r) < \infty \quad \textit{for all } 0 \le r < 2ap,$$

while

$$\mathbb{E}(Y^r) = \infty \quad \textit{for } r \ge 2ap.$$

Proof. From Theorem 4.2, the survival function satisfies

$$\overline{F}(x) \sim C\,x^{-2ap}, \qquad x \to \infty,$$

for some constant $C > 0$. Differentiating gives the tail behavior of the density

$$f(x) \asymp x^{-2ap-1}.$$

Therefore,

$$\int_0^\infty x^r f(x)\,dx$$

converges if and only if

$$r < 2ap.$$

This establishes the result. □

Theorem 4.4 (A computable Beta/Gamma-series for $\mathbb{E}(Y^r)$)**.** *Let $Y \sim$ SS–Dagum(a, b, p) and assume $0 \le r < a$. Then*

$$\mathbb{E}(Y^r) = \frac{\pi}{2}\,b^r\,\Gamma\Big(1 - \frac{r}{a}\Big)\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}\Big(\frac{\pi}{2}\Big)^{2k}\frac{1}{(2k+1)}\frac{\Gamma\Big((2k+1)p + \frac{r}{a}\Big)}{\Gamma((2k+1)p)}. \tag{4.4}$$

Equivalently, using $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$,

$$\mathbb{E}(Y^r) = \frac{\pi}{2}\,b^r\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}\Big(\frac{\pi}{2}\Big)^{2k}\frac{1}{(2k+1)}\frac{B\Big((2k+1)p + \frac{r}{a},\, 1 - \frac{r}{a}\Big)}{B((2k+1)p,\, 1)}. \tag{4.5}$$

Proof. From (3.3), expand

$$\cos z = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!}, \quad z = \frac{\pi}{2} G(x),$$

to obtain

$$f(x) = \frac{\pi}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \left(\frac{\pi}{2}\right)^{2k} g(x)\, G(x)^{2k}.$$

For $0 \le r < a$, dominated convergence justifies termwise integration:

$$\mathbb{E}(Y^r) = \frac{\pi}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \left(\frac{\pi}{2}\right)^{2k} \int_0^{\infty} x^r g(x) G(x)^{2k}\, dx.$$

Since $G(x)^{2k} = (1 + (x/b)^{-a})^{-2kp}$,

$$g(x)G(x)^{2k} = \frac{ap}{b} \left(\frac{x}{b}\right)^{-a-1} \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-(2k+1)p-1} = \frac{1}{2k+1}\, g_{(2k+1)p}(x),$$

where $g_{(2k+1)p}$ is the Dagum pdf with parameter $p' = (2k+1)p$.

Hence,

$$\int_0^{\infty} x^r g(x) G(x)^{2k}\, dx = \frac{1}{2k+1}\, \mathbb{E}(X_{p'}^r), \quad X_{p'} \sim \text{Dagum}(a, b, p').$$

Using the standard Dagum moment formula (valid for $r < a$),

$$\mathbb{E}(X_{p'}^r) = b^r\, \Gamma\!\left(1 - \frac{r}{a}\right) \frac{\Gamma\left(p' + \frac{r}{a}\right)}{\Gamma(p')},$$

we obtain (4.4). The Beta representation follows directly. □

Remark 4.1 (Tail-driven moment extension)**.** *The existence range in Theorem 4.3 extends to $r < 2ap$ because the sine transformation effectively squares the baseline Dagum survival function in the tail region.*

The explicit series representation is stated for $r < a$ to ensure a clean dominated convergence argument and numerical stability. Extension of the series to the full range $r < 2ap$ would require a more delicate tail-splitting argument.

4.6. **Quantile Function.** The quantile function is useful for statistical inference and simulation. In this section, we derive the quantile function of the SS–Dagum distribution based on its cumulative distribution function given in equation (3.2). Let $u \in (0, 1)$ and set $u = F(x)$. From (3.2), we have

$$u = \sin\!\left(\frac{\pi}{2} \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}\right).$$

Taking inverse sine on both sides gives

$$\arcsin(u) = \frac{\pi}{2} \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p}.$$

Hence,

$$\left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-p} = \frac{2}{\pi} \arcsin(u).$$

$$1+\left(\frac{x}{b}\right)^{-a}=\left(\frac{2}{\pi}\arcsin(u)\right)^{-1/p}.$$

Therefore,

$$\left(\frac{x}{b}\right)^{-a}=\left(\frac{2}{\pi}\arcsin(u)\right)^{-1/p}-1.$$

Solving for x, we obtain the quantile function

$$Q(u)=b\left[\left(\frac{2}{\pi}\arcsin(u)\right)^{-1/p}-1\right]^{-1/a}, \qquad 0<u<1. \tag{4.6}$$

4.6.1. *Random Variate Generation.* Random samples from the SS–Dagum distribution can be generated using the inverse transform method:

(1) Generate $U \sim \text{Uniform}(0,1)$.

(2) Compute

$$X=b\left[\left(\frac{2}{\pi}\arcsin(U)\right)^{-1/p}-1\right]^{-1/a}.$$

Then X follows the SS–Dagum(a,b,p) distribution.

4.7. **Stochastic ordering relative to Dagum.**

Theorem 4.5. *Let $X \sim$ Dagum(a,b,p) with cdf G, and let $Y \sim$ SS–Dagum(a,b,p) with cdf $F(x)=\sin\left(\frac{\pi}{2}G(x)\right)$. Then*

$$F(x) \geq G(x) \quad \textit{for all } x>0,$$

and hence $Y \leq_{\text{st}} X$ (the SS–Dagum distribution is stochastically smaller than the classical Dagum distribution).

Proof. For $u \in [0,1]$, define $\phi(u)=\sin(\frac{\pi}{2}u)-u$. Then $\phi(0)=\phi(1)=0$ and $\phi'(u)=\frac{\pi}{2}\cos(\frac{\pi}{2}u)-1$. Since $\phi'(0)=\frac{\pi}{2}-1>0$ and ϕ is concave on $[0,1]$, it follows that $\phi(u)\geq 0$ for all $u\in[0,1]$. Thus $\sin(\frac{\pi}{2}u)\geq u$, and substituting $u=G(x)$ gives $F(x)\geq G(x)$. Therefore $\Pr(Y>x)=1-F(x)\leq 1-G(x)=\Pr(X>x)$ for all x, i.e., $Y\leq_{\text{st}} X$. □

5. Parameter Estimation

In this section, we consider the estimation of the parameters $\boldsymbol{\theta}=(a,b,p)$ of the SS–Dagum distribution using the method of maximum likelihood.

Let $X_1, X_2, \ldots, X_n$ be a random sample from the SS–Dagum(a,b,p) distribution with probability density function given in equation (5) and cumulative distribution function in equation (3). The likelihood function is

$$L(a,b,p)=\prod_{i=1}^{n} f(x_i;a,b,p).$$

Taking logarithms, the log-likelihood function becomes

$$\ell(a,b,p)=\sum_{i=1}^{n}\log f(x_i;a,b,p).$$

Using the density in equation (5), we obtain

$$\begin{aligned}\ell(a,b,p) = n\log\left(\frac{\pi}{2}\right) + n\log a + n\log p - n\log b \\ - (a+1)\sum_{i=1}^{n}\log\left(\frac{x_i}{b}\right) \\ - (p+1)\sum_{i=1}^{n}\log\left(1+\left(\frac{x_i}{b}\right)^{-a}\right) \\ + \sum_{i=1}^{n}\log\left[\cos\left(\frac{\pi}{2}\left(1+\left(\frac{x_i}{b}\right)^{-a}\right)^{-p}\right)\right].\end{aligned}$$

The maximum likelihood estimates (MLEs) of (a,b,p) are obtained by numerically maximizing the log-likelihood function $\ell(a,b,p)$, since the likelihood equations do not admit closed-form solutions. The maximization of $\ell(a,b,p)$ can be carried out using numerical optimization algorithms such as the BFGS method. The optimization is performed subject to the parameter constraints

$$a > 0, \quad b > 0, \quad p > 0.$$

Suitable initial values may be obtained from the classical Dagum model.

Under standard regularity conditions, the asymptotic distribution of the MLE is

$$\widehat{\boldsymbol{\theta}} \sim N\left(\boldsymbol{\theta}, \mathbf{I}^{-1}(\boldsymbol{\theta})\right),$$

where $\mathbf{I}(\boldsymbol{\theta})$ is the observed Fisher information matrix defined by

$$\mathbf{I}(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^{\top}}.$$

The standard errors of the estimators are obtained as the square roots of the diagonal elements of $\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}})$.

Approximate $100(1-\alpha)\%$ confidence intervals for the parameters are given by

$$\widehat{\theta}_j \pm z_{\alpha/2}\sqrt{\mathrm{Var}(\widehat{\theta}_j)}, \quad j = 1,2,3,$$

where $z_{\alpha/2}$ is the standard normal quantile.

6. Simulation Study

6.1. **Monte Carlo Simulation Design.** In this section, a Monte Carlo simulation study is conducted to evaluate the finite-sample performance of the maximum likelihood estimators (MLEs) of the parameters of the SS–Dagum distribution.

Let $X \sim$ SS-Dagum(a,b,p) with cumulative distribution function given by

$$F(x) = \sin\left(\frac{\pi}{2}\left(1+\left(\frac{x}{b}\right)^{-a}\right)^{-p}\right), \qquad x > 0,$$

where $a > 0$, $b > 0$, and $p > 0$.

Random samples were generated using the inverse transform method based on the quantile function. Let $U \sim \text{Uniform}(0,1)$, then the random variate X is obtained as

$$X = b\left[\left(\frac{2}{\pi}\arcsin(U)\right)^{-1/p} - 1\right]^{-1/a}.$$

The simulation study was carried out for sample sizes $n \in \{50, 100, 250, 500\}$, and for each sample size, $R = 2000$ replications were generated. The true parameter values were set to

$$(a, b, p) = (1.5, 2.0, 1.2).$$

For each generated sample, the parameters were estimated using the method of maximum likelihood by numerically maximizing the log-likelihood function.

6.2. **Performance Measures.** The performance of the estimators was assessed using the following criteria:

- **Mean Estimate:**
$$\overline{\hat{\theta}} = \frac{1}{R}\sum_{r=1}^{R}\hat{\theta}_r.$$
- **Bias:**
$$\text{Bias}(\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{R}\hat{\theta}_r - \theta.$$
- **Standard Deviation (SD):**
$$\text{SD}(\hat{\theta}) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\hat{\theta}_r - \overline{\hat{\theta}}\right)^2}.$$
- **Root Mean Squared Error (RMSE):**
$$\text{RMSE}(\hat{\theta}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}(\hat{\theta}_r - \theta)^2}.$$
- **Convergence Rate:** The proportion of successful optimizations among the R replications.

6.3. **Simulation Results.** The simulation results are presented in Table 1 and Table 2.

The results show that the maximum likelihood estimators perform satisfactorily for the SS–Dagum distribution. As the sample size increases, the mean estimates approach the true parameter values, while the bias, standard deviation, and RMSE decrease consistently. This indicates that the estimators are consistent and become more efficient for larger sample sizes.

For small sample size ($n = 50$), relatively large variability is observed, particularly for the parameters a and p. However, the estimation accuracy improves substantially for moderate sample sizes ($n = 100$), and becomes highly stable for larger samples ($n = 250$ and $n = 500$).

In addition, the convergence rate of the numerical optimization procedure is very high across all scenarios, reaching nearly 100% for moderate and large sample sizes. This confirms the reliability and stability of the maximum likelihood estimation procedure for the proposed model.

Overall, the simulation results demonstrate that the SS–Dagum distribution provides reliable parameter estimation, especially for moderate and large sample sizes.

TABLE 1. Monte Carlo simulation results for the SS–Dagum distribution under $(a, b, p) = (1.5, 2.0, 1.2)$

n	Parameter	True	Mean	Bias	SD	RMSE	Conv. Rate
	a	1.5	2.1907	0.6907	5.1217	5.1668	0.9925
	b	2.0	2.0406	0.0406	1.2499	1.2503	0.9925
50	p	1.2	3.1526	1.9526	9.2107	9.4131	0.9925
	a	1.5	1.5745	0.0745	0.3719	0.3792	0.9995
	b	2.0	1.9915	-0.0085	0.8472	0.8471	0.9995
100	p	1.2	1.5426	0.3426	1.8873	1.9183	0.9995
	a	1.5	1.5260	0.0260	0.1871	0.1884	1.0000
	b	2.0	1.9899	-0.0101	0.5276	0.5276	1.0000
250	p	1.2	1.2856	0.0856	0.4194	0.4282	1.0000
	a	1.5	1.5095	0.0095	0.1252	0.1257	1.0000
	b	2.0	1.9921	-0.0079	0.3612	0.3611	1.0000
500	p	1.2	1.2349	0.0349	0.2423	0.2448	1.0000

TABLE 2. Monte Carlo simulation results for the SS–Dagum distribution under $(a, b, p) = (1.0, 1.5, 2.0)$

n	Parameter	True	Mean	Bias	SD	RMSE	Conv. Rate
50	a	1.0	1.2605	0.2605	2.4699	2.4778	0.9885
50	b	1.5	1.5594	0.0594	1.8331	1.8590	0.9885
50	p	2.0	3.2767	1.2767	11.8422	12.7348	0.9885
100	a	1.0	1.0464	0.0464	0.2013	0.2042	0.9960
100	b	1.5	1.5102	0.0102	1.2119	1.2210	0.9960
100	p	2.0	2.3585	0.3585	2.6050	2.6295	0.9960
250	a	1.0	1.0197	0.0197	0.1035	0.1053	0.9995
250	b	1.5	1.4962	-0.0038	0.7894	0.7894	0.9995
250	p	2.0	2.1083	0.1083	0.8235	0.8306	0.9995
500	a	1.0	1.0098	0.0098	0.0704	0.0711	1.0000
500	b	1.5	1.4987	-0.0013	0.5621	0.5621	1.0000
500	p	2.0	2.0489	0.0489	0.4692	0.4717	1.0000

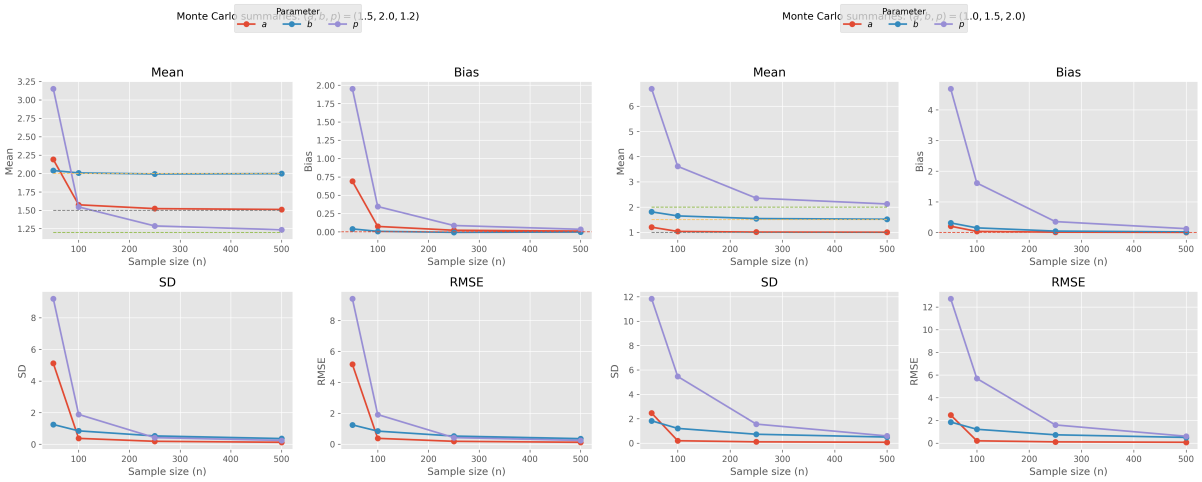


FIGURE 5. Monte Carlo summaries of the maximum likelihood estimators for the SS–Dagum distribution under two parameter settings: $(a, b, p) = (1.5, 2.0, 1.2)$ (left) and $(a, b, p) = (1.0, 1.5, 2.0)$ (right). Each panel displays the mean estimate, bias, standard deviation, and RMSE across different sample sizes.

The simulation results presented in Table 2 and Figure 5 demonstrate that the maximum likelihood estimators perform well for the SS–Dagum distribution. As the sample size increases, the mean estimates converge to the true parameter values, while the bias, standard deviation, and RMSE decrease steadily. This behavior confirms the consistency and efficiency of the estimators. The convergence rate is also very high across all sample sizes, indicating the numerical stability of the estimation procedure.

7. APPLICATION

The Engel dataset is widely used in empirical studies of household consumption and income expenditure relationships, particularly in the analysis of Engel curves. Engel-type data have been extensively employed to investigate consumption behaviour, income elasticity, and welfare patterns across both developed and developing economies. For instance, recent studies such as [4] and [7] utilize Engel-type household expenditure data to estimate consumption functions and assess the validity of Engel's law under different economic settings. These studies highlight the importance of flexible statistical models in capturing nonlinear expenditure patterns and heterogeneity in household behaviour. Moreover, Engel curve analysis remains a fundamental tool in policy-oriented research, including poverty measurement, taxation, and income distribution analysis [5]. Motivated by these applications, the Engel dataset provides an appropriate real-world benchmark for assessing the performance of the proposed SS–Dagum distribution, particularly due to its positive skewness and heavy-tailed characteristics.

TABLE 3. Maximum likelihood estimates (MLEs), standard errors (SE), and 95% confidence intervals for the reduced model set

Model	Parameter	Estimate	SE	CI Lower	CI Upper
Weibull	β	2.3509	0.1717	2.0374	2.7128
	λ	705.0353	21.5923	663.9595	748.6523
Dagum	a	3.7442	0.4649	2.9354	4.7759
	b	463.4285	99.1109	304.7457	704.7384
	p	1.7403	0.8176	0.6929	4.3706
SS–Dagum	a	2.2678	0.3024	1.7463	2.9452
	b	273.9130	153.3627	91.4155	820.7396
	p	6.1246	6.6515	0.7289	51.4661

TABLE 4. Model comparison based on log-likelihood and information criteria for the reduced model set

Model	LogLik	AIC	BIC	CAIC
SS–Dagum	**-1609.9513**	**3225.9027**	**3236.2814**	**3239.2814**
Dagum	-1611.5926	3229.1852	3239.5639	3242.5639
Weibull	-1640.3402	3284.6803	3291.5995	3293.5995

TABLE 5. Goodness-of-fit statistics for the reduced model set

Model	KS Statistic	p-value	AD Statistic
SS–Dagum	**0.0462**	**0.6807**	**0.3161**
Dagum	0.0480	0.6331	0.4326
Weibull	0.0887	0.0466	4.6242

The SS–Dagum distribution provides the best overall fit to the Engel data, as indicated by the highest log-likelihood and the lowest AIC, BIC, and CAIC values. It also achieves the smallest KS and Anderson–Darling statistics, confirming superior agreement with the empirical distribution. Although the Dagum model performs competitively, the SS–Dagum model exhibits greater flexibility. In contrast, the Weibull distribution shows noticeable lack-of-fit, particularly in the tail region.

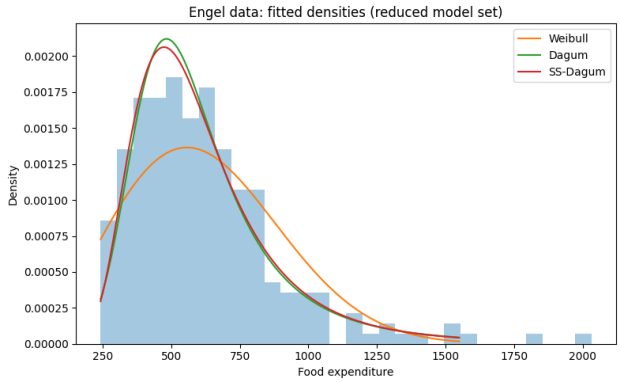


FIGURE 6. Histogram with fitted probability density functions for the reduced model set

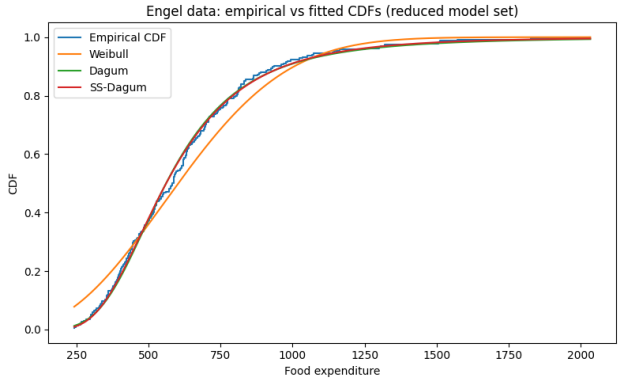


FIGURE 7. Empirical and fitted cumulative distribution functions for the reduced model set

8. CONCLUSION

This paper introduced a sine-transformed extension of the Dagum distribution, referred to as the SS–Dagum model. The proposed distribution preserves the heavy-tailed characteristics of the baseline model while providing additional flexibility to capture diverse shapes in the density and hazard rate functions.

Key statistical properties were derived, and parameter estimation was performed using the maximum likelihood method. A simulation study confirmed the satisfactory performance of the estimators. Applications to real data demonstrated that the SS–Dagum distribution provides improved fit compared to classical competing models.

Overall, the proposed model offers a flexible and tractable alternative for modeling heavy-tailed and positively skewed data. Future work may focus on extending the proposed model to regression frameworks by linking its parameters to covariates, thereby enabling its application in econometric and survival analysis settings.

Acknowledgments: This research budget was allocated by the National Science, Research and Innovation Fund (NSRF), and King Mongkut's University of Technology North Bangkok with (Project no. KMUTNB-FF-69-B-19).

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] A. Alzaatreh, C. Lee, F. Famoye, A New Method for Generating Families of Continuous Distributions, METRON 71 (2013), 63–79. https://doi.org/10.1007/s40300-013-0007-y.

[2] G.M. Cordeiro, M. de Castro, A New Family of Generalized Distributions, J. Stat. Comput. Simul. 81 (2011), 883–898. https://doi.org/10.1080/00949650903530745.

[3] C. Dagum, A New Model of Personal Income Distribution: Specification and Estimation, Écon. Appl. 30 (1977), 413–437. https://doi.org/10.3406/ecoap.1977.4213.

[4] H. Daoud, An Econometric Analysis of Engel's Curve: Household Commodity Groups in Jordan, Indian J. Econ. Bus. 20 (2021), 1085–1091.

[5] P. De Vreyer, S. Lambert, M. Ravallion, Unpacking Household Engel Curves, National Bureau of Economic Research, 2020. https://doi.org/10.3386/w26850.

[6] C. Kleiber, S. Kotz, Statistical Size Distributions in Economics and Actuarial Sciences, Wiley, 2003. https://doi.org/10.1002/0471457175.

[7] M.O.A. Ndanshau, An Econometric Analysis of Engel's Curve: The Case of Peasant Households in Northern Tanzania, Utafiti 4 (2001), 57–70.