A New Bivariate Family of Distributions Based on the Clayton Archimedean Copula and Dagum Distribution

Julius Kwaku Adu-Ntim[1]
Akoto Yaw Omari-Sasu[2]
Maxwell Akwasi Boateng[3*]
Isaac Adjei Mensah[4]

[1,2,4]*Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, Ghana.*

[3] *Department of Mathematics, Kwame Nkrumah University of Science and Technology, Ghana*

[1] ofgord@yahoo.com
[2] ayomari-sasu@knust.edu.gh
[3]boateng.ma@knust.edu.gh
[4]isaacadjeimensah@knust.edu.gh

https://doi.org/10.51867/ajernet.6.1.50

Abstract

This study introduces a novel bivariate distribution combining the Clayton Archimedean copula and the Dagum distribution, addressing challenges in modeling complex dependencies, skewness, heavy tails, and multimodal distributions. The proposed NBCDagE distribution leverages the Clayton copula's ability to capture asymmetric dependencies and the Dagum distribution's flexibility to model diverse data behaviors, making it suitable for reliability, finance, and survival analysis applications. Key statistical properties of the NBCDagE distribution, including the probability density function (PDF), cumulative distribution function (CDF), product and joint moments, and Shannon entropy, were derived and analyzed. The model demonstrates sensitivity to parameter changes, with higher parameter values leading to sharper PDFs and lighter tails, while lower values result in flatter PDFs and heavier tails. Joint moments and entropy analyses revealed the distribution's ability to adapt to varying data complexities, showcasing its robustness in capturing dependence structures and marginal characteristics. Visual representations, including contour plots and density curves, illustrate the flexibility of the NBCDagE model in handling a wide range of dependence patterns and data structures. The distribution's performance was further validated through theoretical derivations and numerical examples, highlighting its adaptability and precision in multivariate data modeling. In conclusion, the NBCDagE distribution provides a robust framework for analyzing bivariate data with intricate dependency structures. Its flexibility and statistical rigor make it a valuable tool for diverse applications, paving the way for future research in higher-dimensional extensions and practical implementations.

Keywords: Bivariate, Clayton Copula, Dagum Distribution, Statistical Properties, Joint Moments, Entropy

1 Introduction

Classical probability distributions have long been the cornerstone for modeling empirical data, providing a solid foundation for statistical inference, parameter estimation, and prediction (Oneto, 2020). These models, however, often struggle to capture the complexities inherent in real-world data, such as skewness, heavy tails, and multimodality, which are frequently observed in applied research (Plasad, 2020; Oneto, 2020). This limitation has led to a growing demand for more flexible models that can better reflect the diverse structures of real-world data (de Sousa Nevas, 2022). Generalized families of distributions, which extend classical models by introducing additional shape parameters or non-standard forms, have emerged as a solution. Distributions like the generalized Pareto, generalized exponential, and skewed distributions offer greater flexibility in modeling features such as asymmetry, heavy tails, and non-constant variance, leading to improved model fit and predictive accuracy (Coia, 2017; Zheng et al., 2018)

The Dagum distribution, introduced by Dagum (1977), has gained significant attention for its flexibility in modeling various types of real-world data, such as income distribution, meteorology, and reliability analysis. Its ability to capture diverse data behaviors, including skewness, heavy tails, and bathtub-shaped hazard functions, has made it a preferred alternative to traditional models like the Pareto and log-normal distributions (Dey, Al-Zahrani & Samerah Basloom, 2017). The Dagum distribution has been widely applied in survival and reliability studies by Febriantikasari et al. (2019) and Chama, Abdulkadir, and Akinrefon (2024), with recent research expanding its scope through extensions such as Weighted and Beta-Dagum distributions (Oluyede & Ye, 2014). Furthermore, advancements like Bayesian estimation and the inverted Dagum distribution have enhanced its analytical capabilities (Nassir & Ibrahim, 2020; Alotaibi et al., 2021). A new generalization of Dagum distribution with application to financial data sets by Ishak and Abiodun (2020) and comprehensive reviews of its applications highlight its growing importance in modeling complex data structures and multivariate dependencies (Dey, Al-Zahrani & Basloom, 2017; Ghalibaf, 2022).

Copulas are a fundamental tool for modeling the dependence structure between random variables by allowing the separation of marginal distributions from their interdependencies. This separation enables the construction of multivariate distributions where the marginals and their dependencies are specified independently (Hao & Singh, 2016). Over time, copulas have become essential for representing multivariate data dependencies, driving significant research and practical applications.

Several bivariate distributions have been derived using copulas, such as the Weibull distribution using the Farlie-Gumbel-Morgenstern (FGM) copula, and other models like the Ali-Mikhail-Haq (AMH), Gumbel-Hougaard, and Gumbel-Barnett copulas (Peres et al., 2018; Saraiva et al., 2018; Kularatne et al., 2021). The Clayton copula is used to create the generalized bivariate Rayleigh distribution, as proposed by El-Sherpieny and Almetwally (2019). The bivariate Fréchet distribution can be constructed using the FGM or AMH copulas, while the generalized inverted Kumaraswamy distribution is based on the Marshall-Olkin method (El-Sherpieny et al., 2022). Additionally, Samanthi and Sepanski (2022) introduced families of bivariate distributions using four different copulas, such as the Kumaraswamy, bivariate exponentiated half-logistic from the Marshall-Olkin class, and the bivariate Lindley distribution derived from the FGM copula. Most recently a new bivariate distribution based on copulas via the Lomax distribution by Aldhufairi et al. (2024)

This study introduces a new bivariate distribution that combines the Clayton Archimedean copula with the Dagum distribution. Archimedean copulas are valued for their closed-form expressions, making them ideal for modeling complex dependencies in multivariate data. The proposed model effectively addresses dependence structures and data distribution features by integrating the Clayton copula's flexibility in capturing dependencies with the versatile hazard function of the Dagum distribution. The study examines the statistical properties of the new distribution, including product moments, joint moments, and joint entropy, offering a robust framework for analyzing complex bivariate data. This distribution has significant potential for enhancing modeling in finance, reliability, and survival analysis, where dependence and distribution flexibility are crucial.

This study is structured as follows: Section 2 provides an overview of general copula-based bivariate distributions, while Section 3 introduces the Archimedean copula, Dagum distributions, and the derivation of the bivariate distributions based on the Clayton and Dagum models. Section 4 focuses on the methods and materials, explicitly addressing the statistical properties of interest. Section 5 presents the discussion, and the final section addresses the study's conclusion.

.

2 The General Copula based Bivariate distribution

2.1 Sklar's Theorem

Let $Y = (Y_1, ..., Y_d)$ be a random vector with marginal cumulative distribution functions (cdfs) $F_1(y_1), ..., F_d(y_d)$ and let $F(y_1, ..., y_d)$ be their joint cdfs. Define

$$u_i = F_i(yi) = P(Yi \leq y_i), i = 1, ..., d. \tag{1}$$

Then, there is a copula function $C(\cdot)$ such that

$$C(u_1, ..., u_d) = P(Y_1 \leq y_1, ..., Y_d \leq y_d) = F(y_1, ..., y_d). \tag{2}$$

By differentiating (2), the joint probability density function (pdf) follows as:

$$f(y_1, ..., y_d) = c(F_1(y_1), ..., F_d(y_d)) \prod_{i=1}^{d} f_i(y_i), i = 1, ..., d. \tag{3}$$

In the bivariate case, we have the following:

$$f(y_1, y_2) = f_1(y_1; \phi) .f_2(y_2; \phi) .c(F_1(y_1; \phi), F_2(y_2; \phi))$$

(4)

Where:

$$f(y_1, y_2) = \frac{\partial^2}{\partial y_1 \partial y_2} F(y_1, y_2)$$

The functions $G_1(y_1; \xi) = Z_1$ and $G_2(y_2; \xi) = Z_2$ represent the cumulative distribution functions (CDFs) of the baseline distributions, characterized by the parameter vector ξ . According to Al-Shomrani (2023) the probability density function (PDF) of a bivariate distribution family, constructed using any copula, is derived as:

$$f(Z_1, Z_2) = (f_1(Z_1) . f_2(Z_2)) \times c(F_1(Z_1), F_2(Z_2))\, dZ_1 dZ_2 \tag{5}$$

The survival function for the bivariate distribution based on any copula is also derived as:

$$S(Z_1, Z_2) = \bar{C}\left(\bar{F_1}(Z_1), \bar{F_2}(Z_2)\right) = \bar{F_1}(Z_1) + \bar{F_2}(Z_2) - 1 + C\left(1 - \bar{F_1}(Z_1), 1 - \bar{F_2}(Z_2)\right) \tag{6}$$

2.2 Product Moment of a Bivariate Distribution

Given the two random variables Y_1 and Y_2 with the pdf of the copula-based bivariate dostribution, the r_1^{th} and r_2^{th} Joint moment is derived by the general expression;

$$E_{Y_1 Y_2}(Y_1^{r_1}, Y_2^{r_2}) = \int_{Y_1}^{\cdot} \int_{Y_2}^{\cdot} y_1^{r_1} y_1^{r_2} f(y_1, y_2)\, dy_1 dy_2$$

By substituting equation (5) into the above expression we have;

$E_{Y_1 Y_2}(Y_1^{r_1}, Y_2^{r_2}) = \int_{Y_1}^{\cdot} \int_{Y_2}^{\cdot} y_1^{r_1} y_1^{r_2} f_1(Z_1) . f_2(Z_2) \times C(F_1(Z_1), F_2(Z_2))\, dZ_1 dZ_2$

We let $F_1(Z_1) = u_1$ and $F_2(Z_2) = u_2$ Then we have that

$$\begin{aligned} E_{Y_1 Y_2}(Y_1^{r_1}, Y_2^{r_2}) &= \int_0^1 \int_0^1 \left\{G_1^{-1}\left(F_1^{-1}(u_1)\right)\right\}^{r_1} \left\{G_2^{-1}\left(F_2^{-1}(u_2)\right)\right\}^{r_2} C(u_1, u_2)\, du_1 du_2 \\ &= E_{u_1, u_2}\left(\left\{G_1^{-1}\left(F_1^{-1}(u_1)\right)\right\}^{r_1} \left\{G_2^{-1}\left(F_2^{-1}(u_2)\right)\right\}^{r_2}\right) \end{aligned} \tag{7}$$

2.3 Joint Moment Generating Function of a Bivariate Distribution

For a bivariate distribution family constructed using any copula, the joint moment generating function of the random variables. Y_1 and Y_2 can be expressed as follows:

$M_{Y_1, Y_2}(r_1, r_2) = \int_{Y_1}^{\cdot} \int_{Y_2}^{\cdot} \exp(r_1 y_1 + r_2 y_2) f(y_1, y_2) dy_1 dy_2$

By substituting equation (5), we have;

$M_{Y_1, Y_2}(r_1, r_2) = \int_{Y_1}^{\cdot} \int_{Y_2}^{\cdot} \exp\{r_1 y_1 + r_2 y_2\} f_1(Z_1) f_2(Z_2) c(F_1(Z_1), F_2(Z_2))\, dZ_1 dZ_2$

Then,

$M_{Y_1, Y_2}(r_1, r_2) = \int_0^1 \int_0^1 \exp\left\{r_1 G_1^{-1}\left(F_1^{-1}(u)\right) + r_2 G_2^{-1}\left(F_2^{-1}(u)\right)\right\} C(u, v) du dv$

$$= M_{U,V}\left\{\exp\left(r_1 G_1^{-1}\left(F_1^{-1}(u)\right) + r_2 G_2^{-1}\left(F_2^{-1}(u)\right)\right)\right\} \tag{8}$$

2.4 Joint Shannon Entropy of a Bivariate Distribution

The Joint Shannon Entropy of a bivariate distribution quantifies the uncertainty or randomness associated with two jointly distributed random variables, Y_1 and Y_2 (Sosa-Cabrera et al., 2019) . It is mathematically defined as:

$$E\left[-\log f(y_1, y_2)\right] = -\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(y_1, y_2)\log[f(y_1, y_2)]\, dy_1 dy_2$$

By substituting Equation (5), we have

$$K = E\left\{-\log f(y_1, y_2)\right\} = -\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(Z_1, Z_2)\log\left[f_1(Z_1).f_2(Z_2)c(F_1(Z_1), F_2(Z_2))\right]\, dZ_1 dZ_2 \tag{9}$$

$$= -\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(Z_1, Z_2)\left[\log c(u_1, u_2) + \log f_1(Z_1) + \log f_1(Z_1)\right]\, dZ_1 dZ_2$$

$$= -\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} c(u_1, u_2)\log c(u_1, u_2)\, du_1 du_2 - \int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(Z_1, Z_2)\left[\log f_1(Z_1) + \log f_1(Z_1)\right] dZ_1 dZ_2$$

Let

$$-\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} c(u_1, u_2)\log c(u_1, u_2)\, du_1 du_2 = K(U_1, U_2)$$

We have that:

$$-\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(Z_1, Z_2)\left[\log f_1(Z_1)\right] dZ_1 dZ_2 = -\int_{Y_1}^{\cdot} \log f(Z_1)\left[\int_{Y_2}^{\cdot} f(Z_1, Z_2) dZ_2\right] dZ_1$$

$$= -\int_{Y_1}^{\cdot} f(Z_1)\log f(Z_1) dZ_1 = K(Y_1)$$

Similarly,

$$-\int_{Y_2}^{\cdot} \log f(Z_2)\left[\int_{Y_1}^{\cdot} f(Z_1, Z_2) dZ_1\right] dZ_2$$

$$= -\int_{Y_2}^{\cdot} f(Z_2)\log f(Z_2) dZ_2 = K(Y_2)$$

Hence, $K = K(Y_1) + K(Y_2) + K(U_1, U_2)$ (10)

3 Archimedean Copula

According to Nelsen (2006), an Archimedean copula is a function. C which takes $[0,1]^n$ to $[0,1]$. These are copulas which take the form:

$$C\left(u_1, ..., u_n\right) = \varphi^{-1}\left(\varphi\left(u_1\right) + ... + \varphi\left(u_n\right)\right)$$

Here, $\varphi\left(t\right)$ is called the generator function of the Archimedean copula and φ^{-1} is its inverse function. Suppose $\varphi : [0,1] \to [0,\infty)$ is a continuous, strictly decreasing function that maps the interval 0 to 1 onto the non-negative real line and has $\varphi(1) = 0$ and $\varphi(0) \leq \infty$. Then its pseudo-inverse is defined as:

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0) \\ 0, & \varphi(0) < t \leq \infty \end{cases}$$

A strict generator has $\varphi(0) = \infty$, i.e., where the pseudo-inverse function is the inverse function. A bivariate copula (with $n = 2$), as described, qualifies as a copula only when the function φ exhibits convexity. For a multivariate Archimedean copula (with $n > 2$) defined similarly, being a copula requires that the generator function is strictly decreasing and has an entirely monotonic inverse. A function $f(t)$ is considered completely monotonic over the interval $[a, b]$ if it meets the following criteria:

$$\frac{(-1)^k d^k}{dt^k} f(k) \geq 0 \quad \forall k \in \mathbb{N}, t \in (a, b)$$

A bivariate Archimedean copula has a Kendall's tau, ρ_T, as follows:

$$\rho_T = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt$$

In the case of a two-dimensional Archimedean copula, the distribution function is expressed as:

$$C(u, \theta) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2); \theta), \quad u = (u_1, u_2)^T$$

where θ the control parameter measures the degree of dependency on the variables.

3.1 The Clayton Archimedean Copula

The Clayton copula has a generating function:

$$\varphi(t) = \frac{t^{-\theta} - 1}{\theta}$$

with distribution:

$$C(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-\frac{1}{\theta}} \tag{11}$$

Moreover, density functions are also derived as:

$$c(u_1, u_2) = (\theta + 1)\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-\frac{(2\theta+1)}{\theta}} (u_1 u_2)^{-(\theta+1)} \tag{12}$$

When θ tends to 0, the variables are independent; otherwise $\theta > 0$.
Where $\varphi(0) = \infty$ and $\varphi^{-1}(t) = (1 + \theta\, t)^{-\frac{1}{\theta}}$, the Clayton Lower and Upper tail dependence become:

$$\theta_L = \lim_{u \to \infty} \left(1 + \frac{\theta u}{1 + \theta u}\right)^{-\dfrac{1}{\theta}} = 2^{-\frac{1}{\theta}}$$

(13)

and

$$\theta_U = 2 - \lim_{u \to 0} \frac{1 - (1 + 2\theta u)^{-\frac{1}{\theta}}}{1 - (1 + \theta u)^{-\frac{1}{\theta}}} = 0 \tag{14}$$

3.2 The Dagum Distribution

For a continuous random variable S to follow a three-parameter **Type I Dagum** distribution,

Denoted $S \sim D_{aq}(p, q, \psi)$, its probability density function (pdf) is given by:

$$f(s; p, q, \psi) = pq\psi s^{-p-1}(1 + \psi s^{-p})^{-q-1} \tag{15}$$

where $\psi > 0$ is a scale parameter and p, q are shape parameters.

The Cumulative distribution function (CDF) is given by:

$$F(s; p, q, \psi) = (1 + \psi s^{-p})^{-q}, \quad s > 0, p, q, \psi > 0 \tag{16}$$

The Quartile function is given as:

$$Q(u) = \inf\{s \in R : F(s) \geq u\}, \quad 0 < u < 1$$

$$F^{-1}(u) = Q(u) = \left\{ \frac{1}{\psi} \left[u^{-\frac{1}{q}} - 1 \right] \right\}^{-\frac{1}{p}} \tag{17}$$

3.3 The NBCDag-G Model

With the baseline cdf given as $G_1(y_1; \xi)$ and $G_2(y_2; \xi)$, substituted into Equation (16) to give a marginal distribution which is, respectively,

$$F_1(G_1(y_1; \xi)) = (1 + \psi_1 G_1(y_1; \xi)^{-p_1})^{-q_1} \tag{18}$$

and

$$F_2(G_2(y_2; \xi)) = (1 + \psi_2 G_2(y_2; \xi)^{-p_2})^{-q_2} \tag{19}$$

The **NBCDag-G** cumulative distribution function (**cdf**) is therefore derived by substituting Equation (18) and (19) into Equation (11) and simplifies to:

$$F_{NBCDag-G}(y_1,y_2)=\left[[(1+\psi G_1(y_1;\xi)^{-p_1})^{-q_1}]^{-\theta}+[(1+\psi G_2(y_2;\xi)^{-p_2})^{-q_2}]^{-\theta}-1\right]^{-\frac{1}{\theta}} \tag{20}$$

The associated joint pdf after differentiating (20) is of the form :

$$\begin{aligned} f_{NBCDag-G}(y_1,y_2) &= (\theta+1)\left[[(1+\psi_1 G_1(y_1;\xi)^{-p_1})^{-q_1}]^{-\theta}+[(1+\psi_2 G_2(y_2;\xi)^{-p_2})^{-q_2}]^{-\theta}-1\right]^{-\frac{2\theta+1}{\theta}} \\ &\times\left[(1+\psi_1 G_1(y_1;\xi)^{-p_1})^{-q_1}\times(1+\psi_2 G_2(y_2;\xi)^{-p_2})^{-q_2}\right]^{-(\theta+1)} \\ &\times[p_1q_1\psi_1 g_1(y_1;\xi)(1+\psi_1 G_1(y_1;\xi)^{-p_1})^{-q_1-1}G_1(y_1;\xi)^{-p_1-1}] \\ &\times[p_2q_2\psi_2 g_2(y_2;\xi)(1+\psi_2 G_2(y_2;\xi)^{-p_2})^{-q_2-1}G_2(y_2;\xi)^{-p_2-1}] \end{aligned} \tag{21}$$

Where

$$g_1(y_1;\xi)=\frac{dG_1(y_1;\xi)}{dy_1},\quad g_2(y_2;\xi)=\frac{dG_2(y_2;\xi)}{dy_2}$$

3.4 The NBCDagE Model

By taking our baseline distribution to be an exponential distribution and assuming that $Y_1\sim g_1(y_1;\xi)\equiv\exp(y_1;\omega_1)$ and $Y_2\sim g_2(y_2;\xi)\equiv\exp(y_2;\omega_2)$, where $\xi=(\omega_1,\omega_2)^T$, with cdfs given as:

$$G_1(y_1;\omega_1)=1-e^{-\omega_1 y_1},\quad G_2(y_2;\omega_2)=1-e^{-\omega_2 y_2}$$

and the respective pdfs given as:

$$g_1(y_1;\omega_1)=\omega_1 e^{-\omega_1 y_1},\quad g_2(y_2;\omega_2)=\omega_2 e^{-\omega_2 y_2}$$

The New Bivariate Clayton Dagum Exponential (**NBCDagE**) Distribution has the cdf and pdf after substituting the above into equations (20) and (21), respectively, as:

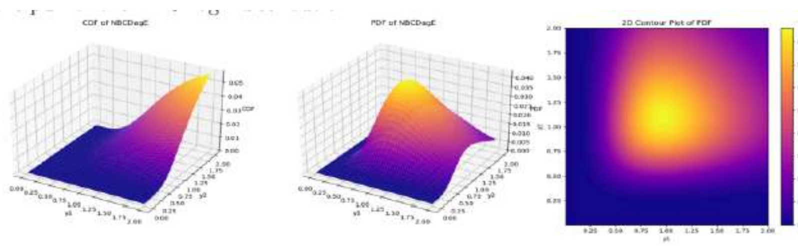
$$F_{NBCDagE}(y_1,y_2)=\left[\left[(1+\psi_1(1-e^{-\omega_1 y_1})^{-p_1})^{-q_1}\right]^{-\theta}+\left[(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2}\right]^{-\theta}-1\right]^{-\frac{1}{\theta}} \tag{22}$$

$$\begin{aligned} f_{NBCDagE}(y_1,y_2) &= (\theta+1)\left[\left[(1+\psi_1(1-e^{-\omega_1 y_1})^{-p_1})^{-q_1}\right]^{-\theta}+\left[(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2}\right]^{-\theta}-1\right]^{-\frac{2\theta+1}{\theta}} \\ &\times\left[\left[(1+\psi_1(1-e^{-\omega_1 y_1})^{-p_1})^{-q_1}\right]\times\left[(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2}\right]\right]^{-(\theta+1)} \end{aligned}$$

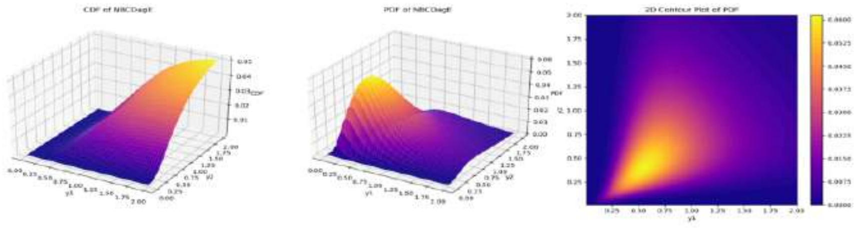
$$\times[p_1 q_1 \psi_1 \omega_1 e^{-\omega_1 y_1}(1+\psi_1(1-e^{-\omega_1 y_1})^{-p_1})^{-q_1-1}(1-e^{-\omega_1 y_1})^{-p_1-1}]$$

$$\times[p_2 q_2 \psi_2 \omega_2 e^{-\omega_2 y_2}(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2-1}(1-e^{-\omega_2 y_2})^{-p_2-1}] \tag{23}$$

Figure *1(a)(b)(c)(d)* presents the bivariate cdfs, the joint densities, and the contour plot of the pdf for the NBCDagE distribution.

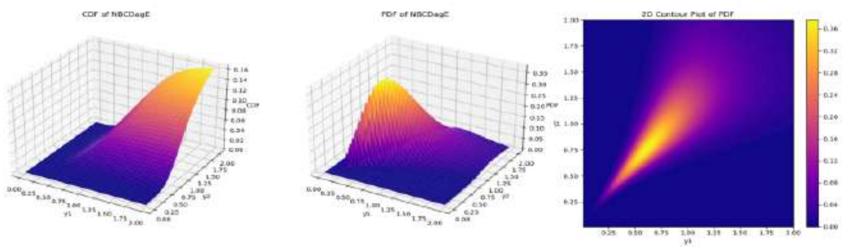


$(a)\omega_1 = 1.5, \quad \omega_2 = 1.5, \quad \psi_1 = 1.8, \quad \psi_2 = 1.8, \quad p_1 = 5.2, \quad p_2 = 5.5, q_1 = 1.1, \quad q_2 = 1.3, \quad \theta = 0.02$

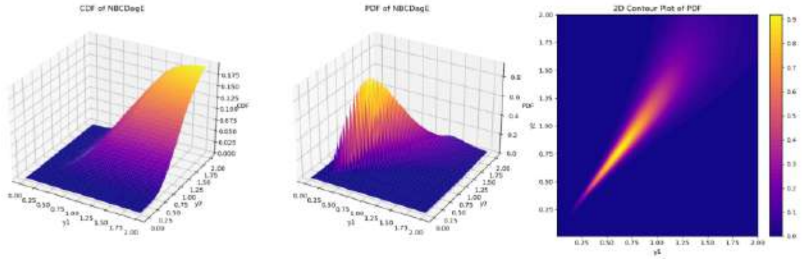


$(b)\omega_1 = 1.5, \quad \omega_2 = 1.5, \quad \psi_1 = 1.8, \quad \psi_2 = 1.8, \quad p_1 = 5.2, \quad p_2 = 1.5, q_1 = 1.1, \quad q_2 = 2.3, \quad \theta = 0.5$

These plots captured in Figure 1 (a)(b)(c) and (d) analyzes the NBCDagE distribution through its PDF, CDF, and contour plots, focusing on the impact of varying parameters. When parameter values are increased, the peak of the PDF becomes sharper, indicating a higher probability density around specific values. Conversely, decreasing parameters tend to flatten the PDF, spreading the probability density over a broader range of values. Particularly, higher θ values produce a sharper and narrower PDF peak, indicating concentrated probability density, and lead to quicker accumulation in the CDF with closer contour lines. Lower θ values result in a flatter, wider PDF, slower accumulation in the CDF, and more dispersed contour lines. The tail behavior is affected as well. Higher θ values lead to lighter tails, while lower values result in heavier tails.



$(c)\omega_1 = 1.5, \quad \omega_2 = 1.5, \quad \psi_1 = 1.8, \quad \psi_2 = 1.8, \quad p_1 = 5.2, \quad p_2 = 5.5, q_1 = 1.1, \quad q_2 = 1.3, \quad \theta = 2.0$



$(d)\omega_1 = 1.5, \quad \omega_2 = 0.5, \quad \psi_1 = 1.8, \quad \psi_2 = 1.8, \quad p_1 = 5.2, \quad p_2 = 2.5, q_1 = 2.1, \quad q_2 = 1.3, \quad \theta = 5.0$

Figure 1: Plot of the cdf, pdf, and contour plot of the pdf of the NBCDagE distribution with parameter values.

4 Methods and Materials

4.1 Some Statistical Properties

Here, we derive the marginal distribution for the **NBCDagE** distribution. Thus, the Dagum distribution with an exponential baseline based on the Clayton copula with $p_i, q_i, \psi_i, \omega_i > 0$ where $i = 1, 2$ and $\theta > 0$. We let $Y_i \sim NBCDagE(p_i, q_i, \psi_i, \omega_i, \theta)$ where $Y_1 \sim DagE(p_1, q_1, \psi_1, \omega_1)$ and $Y_2 \sim DagE(p_2, q_2, \psi_2, \omega_2)$, then we can derive the following:

i The marginal distribution of Y_1 and Y_2

$$f(y_1) = \int_0^\infty f(y_1, y_2) dy_2 = p_1 q_1 \psi_1 \omega_1 e^{-\omega_1 y_1} (1 - e^{-\omega_1 y_1})^{-p_1 - 1} (1 + \psi_1 (1 - e^{-\omega_1 y_1})^{-p_1})^{-q_1 - 1}$$

$$f(y_2) = \int_0^\infty f(y_1, y_2) dy_1$$
$$= p_2 q_2 \psi_2 \omega_2 e^{-\omega_2 y_2} (1 - e^{-\omega_2 y_2})^{-p_2 - 1} (1 + \psi_2 (1 - e^{-\omega_2 y_2})^{-p_2})^{-q_2 - 1}$$

ii The conditional density of Y_1 given Y_2

$$f(y_1 \mid y_2) = \frac{f(y_1, y_2)}{f(y_2)} = (\theta + 1) \left\{ \left[[\left(1 + \psi_1 (1 - e^{-\omega_1 y_1})^{-p_1}\right)^{-q_1}]^{-\theta} + \left(1 + \psi_2 (1 - e^{-\omega_2 y_2})^{-p_2}\right)^{-q_2} \right]^{-\theta} - 1 \right\}^{-\frac{2\theta + 1}{\theta}}$$

$$\times \left[\left[(1 + \psi_1 (1 - e^{-\omega_1 y_1})^{-p_1})^{-q_1} \right] \times \left[(1 + \psi_2 (1 - e^{-\omega_2 y_2})^{-p_2})^{-q_2} \right] \right]^{-(\theta + 1)}$$

$$\times \left[p_1 q_1 \psi_1 \omega_1 e^{-\omega_1 y_1} (1 + \psi_1 (1 - e^{-\omega_1 y_1})^{-p_1})^{-q_1 - 1} (1 - e^{-\omega_1 y_1})^{-p_1 - 1} \right]$$

$$\left(e^{-\omega_2 y_2}\right)^{-p_2(\theta + 1)} \times \left[p_1 q_1 \psi_1 \omega_1 e^{-\omega_1 y_1} (1 + \psi_1 (1 - e^{-\omega_1 y_1})^{-p_1})^{-q_1 - 1} (1 - e^{-\omega_1 y_1})^{-p_1 - 1} \right]$$

iii The conditional distribution of Y_1 given Y_2

$$F(Y_1|Y_2) = \frac{\left[[(1+\psi_1(1-e^{-\omega_1 y_1})^{-p_1})^{-q_1}]^{-\theta} + [(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2}]^{-\theta} - 1\right]^{-\frac{1}{\theta}}}{(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2}}$$

iv The conditional reliability of Y_1 given Y_2

$$S(Y_1|Y_2) = 1 - \left\{\frac{[(1+\psi_1(1-e^{-\omega_1 y_1})^{-p_1})^{-q_1}]^{-\theta} + [(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2}]^{-\theta} - 1}{(1+\psi_2(1-e^{-\omega_2 y_2})^{-p_2})^{-q_2}}\right\}^{-\frac{1}{\theta}}$$

4.2 Product Moments of the NBCDagE Distribution

The product moment measures the expected value of the product of powers of two random variables, Y_1 and Y_2. Mathematically, it is given by: $E_{r_1,r_2}(Y_1^{r_1} Y_2^{r_2})$
where r_1 and r_2 are the powers applied to the random variables. The product moment provides insights into the joint behavior of the two random variables, capturing their interdependence, scaling, and distributional properties. In practical applications, the product moment helps describe complex relationships in data and aids in understanding joint risk or variability in multivariate contexts. The product moment for the NBCDagE using equation (7) is derived by:
Given that:

$$F_1(s) = (1+\psi_1 s^{-p_1})^{-q_1} \quad \text{implies that}$$

$$F_1^{-1}(u) = \left\{\frac{1}{\psi_1}\left[u^{-\frac{1}{q_1}} - 1\right]\right\}^{-\frac{1}{p_1}}$$

and

$$G_1(s) = 1 - e^{-\omega_1 s} \text{also implies that}$$

$$G_1^{-1}(s) = \frac{\log(1-s)}{-\omega_1}.$$

Let

$$u_1 = F_1(G_1(y_1;\xi)) = (1+\psi(G_1(y_1;\xi)^{-p_1})^{-q_1}$$

by inserting the baseline distribution, we have

$$u_1 = (1+\psi_1(1-e^{-\omega_1 y_1})^{-p_1})^{-q_1}$$

this implies that:

$$y_1 = G_1^{-1}(F_1^{-1}(u_1)) = \frac{1}{\omega_1 p_1}\log\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}})$$

Similarly,

$$y_2 = G_2^{-1}(F_2^{-1}(u_2)) = \frac{1}{\omega_2 p_2}\log\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})$$

$$E_{Y_1,Y_2}(Y_1^{r_1}, Y_2^{r_2}) = \frac{(-1)^{r_1+r_2}}{(\omega_1 p_1)^{r_1}(\omega_2 p_2)^{r_2}} E(U_1, U_2)\left[\left(\log\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}})\right)^{r_1}\left(\log\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})\right)^{r_2}\right]$$

$$E_{Y_1,Y_2}(Y_1^{r_1}, Y_2^{r_2}) = \frac{(-1)^{r_1+r_2}}{(\omega_1 p_1)^{r_1}(\omega_2 p_2)^{r_2}} \int_0^1\int_0^1 \left[\left(\log\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}})\right)^{r_1}\left(\log\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})\right)^{r_2}\right] c(u_1,u_2)\, du_1 du_2$$

$$E_{Y_1,Y_2}(Y_1^{r_1}, Y_2^{r_2}) = \frac{(-1)^{r_1+r_2}}{(\omega_1 p_1)^{r_1}(\omega_2 p_2)^{r_2}} \int_0^1\int_0^1 \left[\left(\log\left(\frac{1-u_1^{-\frac{1}{q_1}}}{\psi_1}\right)\right)^{r_1}\left(\log\left(\frac{1-u_2^{-\frac{1}{q_2}}}{\psi_2}\right)\right)^{r_2}\right]$$
$$\times(\theta+1)(u_1^{-\theta}+u_2^{-\theta}-1)^{\frac{-(2\theta+1)}{\theta}}(u_1u_2)^{-(\theta+1)}\, du_1 du_2$$

The product moment of the distributions, $E_{Y_1,Y_2}(Y_1^{r_1}Y_2^{r_2})$, exhibits substantial sensitivity to parameter changes **(Table 1)**. At baseline values, with $\theta = 0.5$ and moderate q_1, q_2, p_1, p_2, the moments grow steadily and remain modest. Increasing q_1, q_2, and θ significantly amplifies the moments, with higher values reflecting greater variability and heavier distribution tails. Extreme growth occurs when θ is further increased, indicating heightened tail dependencies and sensitivity to large values in the integration domain.
Additionally, reducing scaling parameters $(\omega_1, \omega_2, p_1, p_2)$ dramatically increases the coefficient, leading to exponential growth in moments. These findings underscore how parameter adjustments can drastically affect the magnitude and behavior of moments, highlighting the model's sensitivity and the need for careful parameter calibration in practical applications.

Table 1: **Product Moment of NBCDagE Distribution with altering parameter values.**
Parameters: $\omega_1 = 7.5, \omega_2 = 5.3, \psi_1 = 3.1, \psi_2 = 2.2, p_1 = 2.7, p_2 = 3.8, q_1 = 4.05, q_2 = 2.05, \theta = 0.5$

$r_1 \downarrow, r_2 \rightarrow$	1	2	3	4	5
1	1.2951	1.4807	1.6929	1.9355	2.2128
2	1.4727	1.6837	1.9249	2.2008	2.5161
3	1.6745	1.9115	2.1888	2.5025	2.8611
4	1.9041	2.1716	2.4889	2.8455	3.2533
5	2.1651	2.4754	2.8301	3.2356	3.6992

4.3 Joint moment generating function of the NBCDagE Distribution

The Joint Moment Generating Function (MGF) of a distribution is a tool used to summarize all moments of a random variable. For the NBCDagE, its joint MGF is derived based on the definition and specific characteristics of the NBCDagE distribution.
By using equation (8) we have that:

Parameters: $\omega_1 = 7.5, \omega_2 = 5.3, \psi_1 = 3.1, \psi_2 = 2.2, p_1 = 2.7, p_2 = 1.8, q_1 = 10.05, q_2 = 7.05, \theta = 1.5$

$r_1 \downarrow, r_2 \rightarrow$	1	2	3	4	5
1	2.7353	6.6019	15.9344	38.4595	92.8262
2	3.1102	7.5069	18.1189	43.7314	105.5507
3	3.5367	8.5359	20.6024	49.7261	120.0195
4	4.0214	9.7060	23.4265	56.5425	136.4716
5	4.5726	11.0365	26.6378	64.2933	155.1791

Parameters: $\omega_1 = 7.5, \omega_2 = 5.3, \psi_1 = 3.1, \psi_2 = 2.2, p_1 = 1.7, p_2 = 1.8, q_1 = 10.05, q_2 = 7.05, \theta = 10.0$

$r_1 \downarrow, r_2 \rightarrow$	1	2	3	4	5
1	4.3458	10.4891	25.3167	61.1047	147.4830
2	7.8483	18.9429	45.7207	110.3520	266.3468
3	14.1737	34.2098	82.5692	199.2901	481.0087
4	25.5971	61.7812	149.1159	359.9077	868.6773
5	46.2269	111.5737	269.2957	649.9750	1568.7870

Parameters: $\omega_1 = 5.5, \omega_2 = 2.3, \psi_1 = 3.1, \psi_2 = 2.2, p_1 = 2.7, p_2 = 1.8, q_1 = 10.05, q_2 = 7.05, \theta = 1.5$

$r_1 \downarrow, r_2 \rightarrow$	1	2	3	4	5
1	8.5950	47.8039	265.8756	1478.7471	8224.4953
2	13.3271	74.1229	412.2568	2292.8896	12752.5928
3	20.6646	114.9322	639.2298	3555.2683	19773.6903
4	32.0417	178.2095	991.1657	5512.6652	30660.3398
5	49.6827	276.3250	1536.8642	8547.7311	47540.7686

$$M_{Y_1Y_2}(t_1,t_2) = \mathbb{E}_{U_1,U_2}\left\{\exp\left(-\frac{t_1}{\omega_1 p_1}\log\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}}) - \frac{t_2}{\omega_2 p_2}\log\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})\right)\right\}$$

$$= \mathbb{E}_{U_1,U_2}\exp\left\{\log\left(\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}})\right)^{\frac{t_1}{\omega_1 p_1}} + \log\left(\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})\right)^{\frac{t_2}{\omega_2 p_2}}\right\}$$

$$= \mathbb{E}_{U_1,U_2}\left[\left(\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}})\right)^{\frac{t_1}{\omega_1 p_1}}\left(\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})\right)^{\frac{t_2}{\omega_2 p_2}}\right]$$

$$M_{Y_1Y_2}(t_1,t_2) = \int_0^1\int_0^1\left(\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}})\right)^{\frac{t_1}{\omega_1 p_1}}\left(\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})\right)^{\frac{t_2}{\omega_2 p_2}} c(u_1.u_2)du_1du_2$$

$$M_{Y_1Y_2}(t_1,t_2)=\int_0^1\int_0^1\left(\frac{1}{\psi_1}\left(1-u_1^{-\frac{1}{q_1}}\right)\right)^{\frac{t_1}{\omega_1 p_1}}\left(\frac{1}{\psi_2}\left(1-u_2^{-\frac{1}{q_2}}\right)\right)^{\frac{t_2}{\omega_2 p_2}}$$

$$\times(\theta+1)\left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{\frac{(2\theta+1)}{\theta}}(u_1u_2)^{-(\theta+1)}\,du_1du_2$$

This report presents an analysis of the Joint Moment Generating Function (MGF) for NBCDagE under different sets of parameter values. The changes in the MGF values are observed by altering parameters such as ω_1 , ω_2 , ψ_1, ψ_2 , p_1, p_2 , q_1 , q_2 , and θ. The table 2 display the MGF values for varying values of t_1 and t_2 (ranging from 1 to 5) in each case. Below, we summarize the effects of changing the parameters on the MGF values:

Table 2: **Joint Moment Generating of NBCDagE Distribution with altering parameter values.**

Parameters: $\omega_1=10.5,\omega_2=10.3,\psi_1=5.1,\psi_2=8.2,p_1=4.7,p_2=10.8,q_1=10.05,q_2=9.05,\theta=0.5$

$t_1\downarrow,t_2\rightarrow$	1	2	3	4	5
1	1.9539	2.4033	2.9559	3.6357	4.4718
2	3.1156	3.8321	4.7133	5.7973	7.1305
3	4.9679	6.1104	7.5156	9.2440	11.3699
4	7.9215	9.7432	11.8939	14.7399	18.1296
5	12.6311	15.5359	19.1087	23.5032	28.9084

Parameters: $\omega_1=10.5,\omega_2=10.3,\psi_1=5.1,\psi_2=3.2,p_1=4.7,p_2=5.8,q_1=10.05,q_2=9.05,\theta=5.5$

$t_1\downarrow,t_2\rightarrow$	1	2	3	4	5
1	2.3373	3.4365	5.0525	7.4285	10.9217
2	3.7270	5.4796	8.0564	11.8450	17.4151
3	5.9428	8.7374	12.8460	18.8872	27.7609
4	9.4760	13.9321	20.4838	30.1163	44.2786
5	15.1098	22.2153	32.6621	48.0215	70.6038

The joint moment-generating function (MGF) exhibits complex, exponential growth patterns influenced by various parameters (Table 2). As t_1 and t_2 increase, the MGF values increase across all cases, but the parameters significantly impact the rate of increase θ, ψ, p, and others. High values of θ lead to exponential growth in MGF values, especially when t_1 and t_2 are larger. The analysis of these values offers valuable insights into how different parameter configurations affect the behavior of the MGF. It provides guidance for selecting appropriate parameter values based on the desired rate of change in MGF values.

Parameters: $\omega_1 = 8.5, \omega_2 = 9.3, \psi_1 = 2.1, \psi_2 = 8.2, p_1 = 4.7, p_2 = 7.8, q_1 = 7.3, q_2 = 9.05, \theta = 10.5$

$t_1 \downarrow, t_2 \rightarrow$	1	2	3	4	5
1	2.4371	3.3475	4.5981	6.3159	8.6754
2	4.3369	5.9571	8.1826	11.2395	15.4833
3	7.7178	10.6010	14.5614	20.0135	27.4735
4	13.7343	18.8652	25.9129	35.6936	48.8907
5	24.4410	33.5718	46.1136	63.4039	87.0040

Parameters: $\omega_1 = 8.5, \omega_2 = 10.3, \psi_1 = 10.1, \psi_2 = 5.2, p_1 = 5.7, p_2 = 1.8, q_1 = 7.3, q_2 = 5.5, \theta = 10.5$

$t_1 \downarrow, t_2 \rightarrow$	1	2	3	4	5
1	5.5523	19.2241	66.5612	230.4601	797.9397
2	8.9304	30.9204	107.0583	370.6757	1283.4191
3	14.3638	49.7329	172.1940	596.2008	2064.2733
4	23.1030	79.9912	276.9596	958.9386	3320.2068
5	37.1592	128.6591	445.4663	1542.3721	5340.2733

4.4 Joint Shannon Entropy of the NBCDagE Distribution

The Joint Shannon Entropy of the NBCDagE distribution quantifies the uncertainty or information content of the system described by the joint PDF of U_1 and U_2 (Sosa-Cabrera et al., 2019). For a joint distribution $P(X, Y)$ of random variables X and Y, the Joint Shannon Entropy (9) is developed for the NBCDagE distribution as follows:

$$f_1(s) = p_1 q_1 \psi_1 s^{p_1 - 1}(1 + \psi_1 s^{p_1})^{-q_1 - 1}$$

$$f_2(s) = p_2 q_2 \psi_2 s^{p_2 - 1}(1 + \psi_2 s^{p_2})^{-q_2 - 1}$$

$$g_1(s) = \omega_1 e^{-\omega_1 s}$$

$$g_2(s) = \omega_2 e^{-\omega_2 s}$$

$$\frac{d}{dy_1} G_1(y_1; \omega_1) = g_1(y_1; \omega_1)$$

$$\frac{d}{dy_2} G_2(y_2; \omega_2) = g_2(y_2; \omega_2)$$

$$F_1^{-1}(u_1) = G_1(y_1;\omega_1), \quad u_2 = F_2(G_2(y_2;\omega_2))$$

$$F_1^{-1}(u_2) = G_2(y_2;\omega_2)$$

$$E[-\log f(Y_1,Y_2)] = -\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(Z_1,Z_2)[\log c(u_1,u_2) + \log f_1(Z_1) + \log f_1(Z_1)]dZ_1 dZ_2$$

$$= -\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} c(u_1,u_2)\log c(u_1,u_2)du_1du_2 - \int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(Z_1,Z_2)[\log f_1(Z_1) + \log f_1(Z_1)]dZ_1dZ_2$$

Let

$$-\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} c(u_1,u_2)\log c(u_1,u_2)du_1du_2 = K(U_1,U_2)$$

and

$$-\int_{Y_1}^{\cdot}\int_{Y_2}^{\cdot} f(Z_1,Z_2)[\log f_1(Z_1)]dz_1dz_2 = -\int_{Y_1} \log f_1(Z_1)\left[\int_{Y_2} f(Z_1,Z_2)dZ_2\right]dZ_1$$

$$= -\int_{Y_1}^{\cdot} f(Z_1)\log f(Z_1)dZ_1 = K(Y_1)$$

Similarly,

$$-\int_{Y_2}^{\cdot} \log f_2(Z_2)\left[\int_{Y_1}^{\cdot} f(Z_1,Z_2)dZ_1\right]dz_2$$

$$= -\int_{Y_2}^{\cdot} f(Z_2)\log f(Z_2)dZ_2 = K(Y_2)$$

Hence,

$$K = K(Z_1) + K(Z_2) + K(U_1,U_2)$$

$$\int_0^1\int_0^1 \left[(\theta+1)(u_1^{-\theta}+u_2^{-\theta}-1)^{-\frac{(2\theta+1)}{\theta}}(u_1u_2)^{-(\theta+1)}\right]\log\left[(\theta+1)(u_1^{-\theta}+u_2^{-\theta}-1)^{-\frac{(2\theta+1)}{\theta}}(u_1u_2)^{-(\theta+1)}\right]du_1du_2$$

$$K(Y_1) = -\int_{Y_1}^{\cdot} f(Z_1)\log f(Z_1)\,dZ_1$$

$$K(U_1) = -\int_0^1 \left[f_1(F_1^{-1}(u_1))g_1(G_1^{-1}(F_1^{-1}(u_1)))\log f_1(F_1^{-1}(u_1))\right] du_1$$

$$K(U_2) = -\int_0^1 \left[f_1(F_1^{-1}(u_2))g_1(G_2^{-1}(F_2^{-1}(u_2)))\log f_1(F_1^{-1}(u_2))\right] du_2$$

$$f_1(F_1^{-1}(u_1)) = f_1\left(\left\{\frac{1}{\psi_1}\left[u_1^{-\frac{1}{q_1}}-1\right]\right\}^{-\frac{1}{p_1}}\right) = p_1q_1\psi_1\left(\left\{\frac{1}{\psi_1}\left[u_1^{-\frac{1}{q_1}}-1\right]\right\}^{-\frac{1}{p_1}}\right)^{-p_1-1}$$

$$\left(1+\psi_1\left(\left\{\frac{1}{\psi_1}\left[u_1^{-\frac{1}{q_1}}-1\right]\right\}^{-\frac{1}{p_1}}\right)^{-p_1}\right)^{-q_1-1}$$

Which is simplified as:

$$= p_1q_1\psi_1\left\{\frac{1}{\psi_1}\left[u_1^{-\frac{1}{q_1}}-1\right]\right\}^{\frac{p_1+1}{p_1}} u_1^{\frac{q_1+1}{q_1}}$$

Similarly,

$$f_2\left(\left\{\frac{1}{\psi_2}\left[u_2^{-\frac{1}{q_2}}-1\right]\right\}^{-\frac{1}{p_2}}\right) = p_2q_2\psi_2\left\{\frac{1}{\psi_2}\left[u_2^{-\frac{1}{q_2}}-1\right]\right\}^{\frac{p_2+1}{p_2}} u_2^{\frac{q_2+1}{q_2}}$$

Also,

$$g_1\left(\frac{1}{\omega_1p_1}\log\frac{1}{\psi_1}(1-u_1^{-\frac{1}{q_1}})\right) = \omega_1\psi_1^{-\frac{1}{p_1}}(1-u_1^{-\frac{1}{q_1}})^{-\frac{1}{p_1}},$$

and

$$g_2\left(\frac{1}{\omega_2p_2}\log\frac{1}{\psi_2}(1-u_2^{-\frac{1}{q_2}})\right) = \omega_2\psi_2^{-\frac{1}{p_2}}(1-u_2^{-\frac{1}{q_2}})^{-\frac{1}{p_2}}.$$

$$K(U_1) = -\int_0^1\left[p_1q_1\psi_1\left\{\frac{1}{\psi_1}\left[u_1^{-\frac{1}{q_1}}-1\right]\right\}^{\frac{p_1+1}{p_1}} u_1^{\frac{q_1+1}{q_1}}\right]\left[\omega_1\psi_1^{-\frac{1}{p_1}}\left(1-u_1^{-\frac{1}{q_1}}\right)^{-\frac{1}{p_1}}\right]$$

$$\log\left[p_1q_1\psi_1\left\{\frac{1}{\psi_1}\left[u_1^{-\frac{1}{q_1}}-1\right]\right\}^{\frac{p_1+1}{p_1}}u_1^{\frac{q_1+1}{q_1}}\right]du_1.$$

Simplified as:

$$K(U_1)=-p_1q_1\omega_1\psi_1^{-\frac{2}{p_1}}\int_0^1 u_1\left[1-u_1^{-\frac{1}{q_1}}\right]\log\left[p_1q_1\psi_1^{-\frac{2}{p_1}}\left\{\left[u_1^{-\frac{1}{q_1}}-1\right]\right\}^{\frac{p_1+1}{p_1}}u_1^{\frac{q_1+1}{q_1}}\right]du_1.$$

Similarly,

$$K(U_2)=-p_2q_2\omega_2\psi_2^{-\frac{2}{p_2}}\int_0^1 u_2\left[1-u_2^{-\frac{1}{q_2}}\right]\log\left[p_2q_2\psi_2^{-\frac{2}{p_2}}\left\{\left[u_2^{-\frac{1}{q_2}}-1\right]\right\}^{\frac{p_2+1}{p_2}}u_2^{\frac{q_2+1}{q_2}}\right]du_2.$$

Finally,

$$K=E\{-\log[f(Y_1,Y_2)]\}=K(U_1)+K(U_2)+K(U_1,U_2).$$

Table 3: **Joint Shannon Entropy for the NBCDagE distribution with varying parameter values .**

Parameters: $\omega_1=5.0,\omega_2=5.0,\psi_1=2.0,\psi_2=2.0,q_1=2.0,q_2=2.0,\theta=1$

$p_1\downarrow,p_2\rightarrow$	0.5	2.5	5	10
0.5	0.2208	2.1127	1.2441	-5.1262
2.5	2.1127	4.0046	3.1360	-3.2343
5	1.2441	3.1360	2.2674	-4.1029
10	-5.1262	-3.2343	-4.1029	-10.4732

Parameters: $\omega_1=5.0,\omega_2=5.0,\psi_1=2.5,\psi_2=2.5,p_1=5.0,p_2=5.0,\theta=2.5$

$q_1\downarrow,q_2\rightarrow$	0.5	2.5	5	10
0.5	1.6286	2.0761	2.4664	2.9611
2.5	2.0761	2.5237	2.9140	3.4087
5	2.4664	2.9140	3.3042	3.7989
10	2.9611	3.3042	3.7989	4.2936

The entropy analysis in **Table 3** across various parameter combinations reveals that as parameters such as ω_1, ω_2, q_1, and q_2 increase, the entropy values generally rise, indicating greater system complexity and disorder. Conversely, increasing values of θ and p_2 tend to decrease entropy, signaling more excellent system stability. Negative entropy values, particularly for higher p_1 and p_2, suggest instability or complex behavior, while

Parameters: $p_1 = 5.0, p_2 = 5.0, \psi_1 = 2.0, \psi_2 = 2.0, q_1 = 2.0, q_2 = 2.0, \theta = 1$

$\omega_1 \downarrow, \omega_2 \rightarrow$	0.5	2.5	5	10
0.5	0.0529	0.5450	1.1602	2.3904
2.5	0.5450	1.0371	1.6523	2.8826
5	1.1602	1.6523	2.2674	3.4977
10	2.3904	2.8826	3.4977	4.7280

Parameters: $\omega_1 = 5.0, \omega_2 = 5.0, p_1 = 5.0, p_2 = 5.0, q_1 = 2.0, q_2 = 2.0, \theta = 1$

$\psi_1 \downarrow, \psi_2 \rightarrow$	0.5	2.5	5	10
0.5	-3.2259	-0.2750	0.1059	0.2182
2.5	-0.2750	2.6760	3.0568	3.1692
5	0.1059	3.0568	3.4377	3.55001
10	0.2182	3.4377	3.55001	3.6623

Parameters: $\omega_1 = 5.0, \omega_2 = 5.0, \psi_1 = 2.5, \psi_2 = 2.5, p_1 = 5.0, p_2 = 5.0, q_1 = 2.0, q_2 = 2.0$

θ	0.5	2.5	5	10
Shannon Entropy	2.7970	2.3307	1.9107	1.3803

positive values reflect higher system diversity and order. The results highlight a trend where higher ω_1 , ω_2, and q_1 , q_2 values lead to more complex, disordered systems. In contrast, larger θ values contribute to more stable, less complex systems.

5 Discussion

The findings of this study contribute to the growing body of research on multivariate distributions based on copulas by introducing a new bivariate family combining the Clayton Archimedean copula with the Dagum distribution. Copula-based approaches have become indispensable in modeling complex dependencies because they can separate marginal distributions from dependency structures (Sun et al., 2019; Fang & Pan, 2021). This study aligns with earlier works, such as those by Oh and Patton (2017), which demonstrated the utility of copulas in capturing high-dimensional dependencies. This study expands on existing methodologies by leveraging the Clayton copula's flexibility in modeling asymmetric dependencies and the Dagum distribution's versatile hazard function, offering a model that effectively handles dependence and marginal distribution complexities.

In comparison with other bivariate distributions derived from copulas, such as those using the Farlie–Gumbel–Morgenstern (FGM), Ali–Mikhail–Haq (AMH), and Gumbel-Hougaard copulas, the NBCDagE distribution offers greater adaptability. For instance, the Clayton copula, known for its ability to model strong lower-tail dependence by Li and Kang (2018), complements the Dagum distribution's ability to capture heavy tails and varying hazard functions, as also identified by Okorie et al.(2019). This combination is particularly advantageous over simpler copula-based models like the bivariate Weibull distribution derived from the FGM copula, which may be limited in capturing extreme dependence or tail behavior, according to Cooray (2019). Additionally, the sensitivity of the NBCDagE distribution's entropy and joint moments to parameter variations underscores its potential for applications requiring precise modeling of variability and dependence, such as in reliability analysis and survival studies.

Furthermore, the study's findings align with recent advancements in copula-based multivariate modeling, such as the use of Marshall-Olkin methods to construct generalized bivariate distributions like the inverted Kumaraswamy or bivariate Lindley distributions by Peres et al. (2018) and Tahir et al. (2020). While effective in specific applications, these models may lack the comprehensive flexibility offered by the NBCDagE distribution, particularly in capturing diverse dependence structures and tail behaviors. The detailed examination of the NBCDagE distribution's statistical properties, including entropy and joint moments, contributes a robust framework for analyzing complex bivariate data. This paper highlights the versatility of copula-based models. It sets a foundation for future research exploring extensions to higher dimensions and diverse application areas, such as finance, environmental modeling, and health sciences.

6 Conclusion

This study presents a new bivariate distribution family combining the Clayton Archimedean copula with the Dagum distribution, aiming to model complex dependencies and diverse data structures. We analyzed the impact of varying parameters($\omega_1, \omega_2, \psi_1, \psi_2, p_1, p_2, q_1, q_2 a \text{and} \theta$) on the distribution's behavior, including its PDF, CDF, joint moments, and entropy. The analysis of the NBCDagE distribution through its PDF, CDF, and contour plots shows that increasing parameters sharpens the PDF peak and concentrates the probability density, while decreasing parameters flatten the PDF and spread the density over a broader range. Higherθ values lead to lighter tails and quicker accumulation in the CDF. In comparison, lower θ values result in heavier tails and slower accumulation, demonstrating the distribution's flexibility in capturing varying data characteristics. The results show that increasing parameters such as $\omega_1, \omega_2, q_1, \text{and} q_2$ leads to higher entropy values, indicating greater complexity and disorder. In contrast, higher θ,p_1 and p_2 values result in lower entropy, signaling more stable systems. Joint moments are sensitive to parameter changes, with higher values resulting in increased variability and heavier tails. The model effectively captures various dependence structures, making it suitable for accurate economic modeling, reliability, and survival analysis. Future research may extend this framework to higher-dimensional settings, refine parameter estimation methods, improve computational efficiency, and explore applications in diverse industries.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could

have appeared to influence the work reported in this paper.

Funding sources

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Al-Shomrani, A. A. (2023). New bivariate family of distributions based on any copula function: Statistical properties. *Heliyon*, 9(4).

[2] Aldhufairi, F. A. A., & Sepanski, J. H. (2024). New Bivariate Copulas via Lomax Distribution Generated Distortions. *AppliedMath*, 4(2), 641-665.

[3] Alotaibi, R., Rezk, H., Dey, S., & Okasha, H. (2021). Bayesian estimation for Dagum distribution based on progressive type I interval censoring. *PLOS ONE*, 16(6), e0252556.

[4] Bechiri, S., & Remita, M. R. (2024). On Dagum-Pareto mixture distribution: properties, simulation, and application in insurance. *Studies in Engineering and Exact Sciences*, 5(2), e7530-e7530.

[5] Chama, A. F., Abdulkadir, S. S., & Akinrefon, A. A. (2024). Cure Rate Model in Survival Analysis Using Cubic Transmuted Dagum Distribution. *European Journal of Applied Science, Engineering and Technology*, 2(4), 117-125.

[6] Coia, V. (2017). Forecasting of nonlinear extreme quantiles using copula models (Doctoral dissertation, University of British Columbia).

[7] Cooray, K. (2019). A new extension of the FGM copula for negative association. *Communications in Statistics-Theory and Methods*, 48(8), 1902-1919.

[8] Dagum, C. (1977). A New Model for Personal Income Distribution: Specification and Estimation. *Economie Appliquée*, 30, 413-437.

[9] de Sousa Neves, G. (2022). Models for Forecasting Value at Risk: A Comparison of the Predictive Ability of Different VaR Models to Capture Market Losses Incurred During the 2020 Pandemic Recession (Master's thesis, Universidade NOVA de Lisboa (Portugal)).

[10] Dey, S., Al-Zahrani, B., & Basloom, S. (2017). Dagum distribution: Properties and different methods of estimation. *International Journal of Statistics and Probability*, 6(2), 74-92.

[11] El-Sherpieny, E. S. A., Muhammed, H. Z., & Almetwally, E. M. (2022). Bivariate Chen distribution based on copula function: Properties and application of diabetic nephropathy. *Journal of Statistical Theory and Practice*, 16(3), 54.

[12] El-Sherpieny, E. S., & Almetwally, E. M. (2019, December). Bivariate generalized rayleigh distribution based on Clayton Copula. In *Proceedings of the annual conference on statistics (54rd), computer science and operation research, faculty of graduate studies for statistical research, Cairo University* (pp. 1-19).

[13] Fang, G., & Pan, R. (2021). On multivariate copula modeling of dependent degradation processes. *Computers & Industrial Engineering*, 159, 107450.

[14] Febriantikasari, E., Adnan, A., Yendra, R., & Muhaijir, M. N. (2019). Using Dagum Distribution to Simulated Concentration PM10 in Pekanbaru City, Indonesia. *Applied Mathematical Sciences*, 13(9), 439-448.

[15] Ghalibaf, M. B. (2022). Stress-strength reliability with dependent variables based on copula function. *Journal of Dynamics & Games*, 9(3).

[16] Hao, Z., & Singh, V. P. (2016). Review of dependence modeling in hydrology and water resources. *Progress in Physical Geography*, 40(4), 549-578.

[17] Ishaq, A. I., & Abiodun, A. A. (2020). A new generalization of Dagum distribution with application to financial data sets. In *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)* (pp. 1-6). IEEE.

[18] Kularatne, T. D., Li, J., & Pitt, D. (2021). On the use of Archimedean copulas for insurance modelling. *Annals of Actuarial Science*, 15(1), 57-81.

[19] Li, F., & Kang, Y. (2018). Improving forecasting performance using covariate-dependent copula models. *International Journal of Forecasting*, 34(3), 456-476.

[20] Nassir, L. M., & Ibrahim, N. A. (2020). The Bayesian estimator for probabilistic dagum distribution. *International Journal of Innovation, Creativity and Change*, 12(5), 42-74.

[21] Nelsen, R. B. (2006). *An introduction to copulas.*

[22] Oh, D. H., & Patton, A. J. (2017). Modeling dependence in high dimensions with factor copulas. *Journal of Business & Economic Statistics*, 35(1), 139-154.

[23] Okorie, I. E., Akpanta, A. C., & Osu, B. O. (2019). Flexible heavy tail distributions for Surface ozone for selected sites in the United States of America. *Ozone: Science & Engineering*, 41(5), 473-488.

[24] Oluyede, B. O., & Ye, Y. (2014). Weighted Dagum and related distributions. *Afrika Matematika*, 25, 1125-1141.

[25] Oneto, L. (2020). *Model selection and error estimation in a nutshell.* Springer International Publishing.

[26] Peres, M. V. D. O., Achcar, J. A., & Martinez, E. Z. (2018). Bivariate modified Weibull distribution derived from Farlie-Gumbel-Morgenstern copula: a simulation study. *Electronic Journal of Applied Statistical Analysis*, 11(2), 463-488.

[27] Prasad, A. S. (2020). *Dependence: From classical copula modeling to neural networks.*

[28] Samanthi, R. G. M., & Sepanski, J. (2022). On bivariate Kumaraswamy-distorted copulas. *Communications in Statistics-Theory and Methods*, 51(8), 2477-2495.

[29] Saraiva, E. F., Suzuki, A. K., & Milan, L. A. (2018). Bayesian computational methods for sampling from the posterior distribution of a bivariate survival model, based on AMH copula in the presence of right-censored data. *Entropy*, 20(9), 642.

[30] Shankar, P. M., & Shankar, P. M. (2021). Multiple Random Variables and Their Characteristics. *Probability, Random Variables, and Data Analytics with Engineering Applications*, 233-336.

[31] Sosa-Cabrera, G., García-Torres, M., Gómez-Guerrero, S., Schaerer, C. E., & Divina, F. (2019). A multivariate approach to the symmetrical uncertainty measure: Application to feature selection problem. *Information Sciences*, 494, 1-20.

[32] Sun, F., Zhang, W., Wang, N., & Zhang, W. (2019). A copula entropy approach to dependence measurement for multiple degradation processes. *Entropy*, 21(8), 724.

[33] Tahir, M. H., Hussain, M. A., Cordeiro, G. M., El-Morshedy, M., & Eliwa, M. S. (2020). A new Kumaraswamy generalized family of distributions with properties, applications, and bivariate extension. *Mathematics*, 8(11), 1989.

[34] Zhang, Y., Kim, C. W., Beer, M., Dai, H., & Soares, C. G. (2018). Modeling multivariate ocean data using asymmetric copulas. *Coastal Engineering*, 135, 91-111.