International Journal of Advanced Statistics and Probability, 12 (2) (2025) 48-67

International Journal of Advanced Statistics and Probability

Website: www.sciencepubco.com/index.php/IJASP
https://doi.org/10.14419/3b25xq49
Research paper

Derivation of The Generalized Gamma Distribution Via An Ordinary Differential Equation Approach: Theory and Applications

Mohammad Shakil [1] *, Tassaddaq Hussain [2], Mohammad Ahsanullah [3], B. M. G. Kibria [4]

[1] *Department of Mathematics, Miami Dade College, Hialeah, FL, USA*
[2] *Department of Statistics, Mirpur University of Science and Technology, Mirpur, Pakistan*
[3] *Department of Management Sciences, Emeritus Professor, Rider University, NJ, USA*
[4] *Dept. of Math & Stat, Florida International University, FL, USA*
**Corresponding author E-mail: mshakil@mdc.edu*

Received: August 6, 2025, Accepted: November 3, 2025, Published: November 8, 2025

Abstract

This paper introduces a novel approach for generating the Generalized Generalized Gamma (GGG) distribution by employing a second-order non-homogeneous non-linear differential equation with constant coefficients. The proposed framework yields a flexible univariate continuous probability distribution with a rich structure. We provide a comprehensive analysis of the distribution's formulation, key statistical properties, and characterizations via truncated moments. Parameter estimation is carried out through both the Maximum Likelihood Estimation (MLE) and the Method of Moments. A simulation study is conducted to assess the performance of these estimators. The practical applicability of the GGG distribution is demonstrated using real-life datasets, with model adequacy evaluated through multiple goodness-of-fit criteria, including the Anderson-Darling and Cramér–von Mises tests. The empirical results highlight the superior performance of the proposed distribution in modeling complex data patterns. Findings and implications are thoroughly discussed, underscoring the GGG distribution's potential in statistical modeling and applied probability.

Keywords: *Characterization; Differential Equation; Generalized Gamma Distribution; Truncated Moment.*

1. Introduction

Probability distributions play a central role in statistical analysis and mathematical modeling, enabling researchers to describe uncertainty, variability, and randomness in various real-world phenomena. Continuous probability distributions, in particular, are often used when dealing with data that can take any value within a range or interval. Over the years, many methods have been proposed to generate such distributions, with the differential equation approach standing out as one of the most powerful and flexible techniques.
Karl Pearson, in his work in the late 19th century, introduced the method of generating probability distributions through differential equations. This approach provides a framework for deriving distributions that satisfy certain conditions or properties, such as moments, skewness, and kurtosis. Since Pearson's pioneering work, the differential equation approach has evolved significantly, and it continues to be a rich area of research in modern statistical theory. In this research note, we examine the differential equation approach to generating a univariate continuous probability distribution, namely, the generalized generalized gamma distribution (GGGD).
In this Section, we begin by reviewing its historical origins, then present a summary of the mathematical foundations of the approach, and finally review current advancements and applications in the literature. The rest of the paper is organized as follows: Section 2 of the paper delves into the mathematical formulation of the generalized gamma distribution (GGGD) as a solution of a second-order non-homogeneous non-linear differential equation having constant coefficients. Section 3 contains the characterizations by truncated moments. In Sections 4 and 5, we provide parameter estimation and applications, respectively. Some conclusions are drawn in Section 6.

1.1. Current literature review

The differential equation approach to generating probability distributions has undergone significant development since Pearson's initial contributions. This section explores the major milestones in its evolution and the most recent contributions to the field.

1.1.1. Karl Pearson's pioneering work (1895, 1901)

Karl Pearson's work laid the groundwork for using differential equations to generate probability distributions. In his seminal papers, Pearson introduced what is now known as the Pearson Type I and Pearson Type II distributions. These distributions were derived using the following Pearson's first-order linear ordinary differential equation that satisfied boundary conditions related to the first and second moments:

$$\frac{1}{f(x)}\frac{df_x(x)}{dx} = -\frac{x+a}{bx^2+cx+d},$$

In this context, a, b, c, and d represent real parameters, ensuring that f serves as a probability density function. The above Pearson's system, through parameter tuning, leads to well-known distributions like the normal, beta, and gamma families. For details, see Pearson (1895, 1901, 1916). Pearson's work was a critical advancement in the study of univariate distributions, as it linked the properties of distributions to solutions of differential equations. By adjusting these parameters, Pearson's system can generate several classical distributions, such as the Normal, Beta, and Gamma distributions. This development marked a major shift in probability theory — linking distributional forms directly to the solutions of ODEs. It is well-known that the Pearson system comprises several families of distributions, known as Pearson Types I – VI, and Pearson Types VII – XII, which are governed by the above Pearson's first-order linear differential equation. Pearson's framework thus provided a unifying structure for modeling diverse datasets within a single theoretical system.
Obviously, Pearson's first-order linear differential equation is a first-order linear differential equation of the form,

$$\left(\frac{1}{f(x)}\right)\frac{df(x)}{dx} = \frac{P(x)}{Q(x)}, \; Q(x) \neq 0,$$

Where $f(x)$ Is the probability density function of a continuous random variable X, Whose specific value is x, $P(x)$ and $Q(x)$ Are polynomials in x, and $\frac{P(x)}{Q(x)}, Q(x) \neq 0,$ for any x, It is a rational function. Thus, by putting the Pearson or the generalized Pearson differential equations as a first-order linear differential equation of the form, $\left(\frac{1}{f(x)}\right)\frac{df(x)}{dx} = \frac{P(x)}{Q(x)}$, We could solve it by finding the corresponding integrating factor, $I(x) = e^{\int r(x)dx}$, where $r(x) = P(x) / Q(x)$, Which can easily be determined by the integration of the Rational Function, $r(x) = P(x) / Q(x)$, which appears as the exponent of the exponential function in the integrating factor. This could be achieved using "the techniques for integrating rational functions, $\frac{P(x)}{Q(x)}$, by decomposing these into partial fractions". After determining the integrating factor, $I(x) = e^{\int r(x)dx}$ We can easily find the normalizing constant "C", using the definition of the probability density function for a non-negative continuous random variable. X, and as such, we will have a nice probability density function, $f(x)$. Thus, using the first-order linear differential equation approach as described above, we could easily generate twelve different types of continuous probability density functions of the Pearson system from the above Pearson's first-order linear differential equation, for different values of the parameters a, b, c, and d. These probability density functions and the corresponding linear differential equations from which they were derived are provided in Table 1.1. Note that the flexibility of the Pearson system, achieved by adjusting its parameters, made it widely applicable in many fields.

Table 1.1: Pearson's Twelve Types of Distributions

Type	Linear Differential Equation: $\left(\frac{1}{f(x)}\right)\frac{df(x)}{dx} = \frac{P(x)}{Q(x)}$, $P(x)$ and $Q(x)$ Are polynomials in x, $(x) \neq 0$, for any x, and $y = f(x)$.	nsity Function: $f(x) = y_0 e^{\int r(x)dx}$,where $e^{\int r(x)dx} = I(x)$ is the Integrating Factor, $r(x) = \frac{P(x)}{Q(x)}$, and y_0 is the normalizing constant.
	$\frac{dy}{dx} = \frac{\nu(a_1+a_2)x}{(a_1+x)(a_2-x)}y$	$y = f(x) = y_0\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2}$,is ≥ 0, when $-a_1 \leq x \leq a_2$, where $\frac{m_1}{a_1} = \frac{m_2}{a_2}$ (= ν, say) > 0. (Beta distribution of the first kind)
II	$\frac{dy}{dx} = \frac{-2mx}{a^2-x^2}y$	$y = y_0\left(1-\frac{x^2}{a^2}\right)^m, -a \leq x \leq a.$
III	$\frac{dy}{dx} = \frac{-\nu x}{a+x}y$	$= y_0\left(1+\frac{x}{a}\right)^{\nu a}e^{-\nu x}, -a \leq x < \infty.$ (Gamma distribution)
IV	$\frac{dy}{dx} = \frac{-2mx-\nu a}{a^2+x^2}y$	$y = y_0\left(1+\frac{x^2}{a^2}\right)^{-m}e^{-\nu\tan^{-1}(\frac{x}{a})},$ $\nu, m > 0, -\infty < x < \infty.$
V	$\frac{dy}{dx} = \frac{q-px}{x^2}y$	$y = y_0 x^{-p}e^{-\frac{q}{x}},$ $p, q > 0, 0 \leq x < \infty.$
VI	$\frac{dy}{dx} = \frac{ha+(q-h)x}{x^2-ax}y$	$y = y_0(x-a)^q x^{-h},$ $0 < c \leq x < \infty, h > q+1 > 0.$ (Beta distribution of the second kind)
VII	$\frac{dy}{dx} = -\frac{(x+a)}{b_0}y$	$y = y_0 e^{-\frac{(x+a)^2}{2b_0}},$ $-\infty < x < \infty, b_0 > 0, -\infty < a < \infty$ (Normal distribution).
VIII	$\frac{dy}{dx} = \frac{-m}{a+x}y$	$= y_0\left(1+\frac{x}{a}\right)^{-m}, < m < 1, -a \leq x \leq 0.$
IX	$\frac{dy}{dx} = \frac{m}{a+x}y$	$= y_0\left(1+\frac{x}{a}\right)^{m}, s-a \leq x \leq 0.$
X	$\frac{dy}{dx} = \pm\frac{1}{\sigma}y$	$y = \frac{n}{\sigma}e^{\pm\frac{x}{\sigma}},$ $0 \leq x < \infty, \sigma > 0.$
XI	$\frac{dy}{dx} = \frac{-m}{x}y$	$= y_0 x^{-m}.$

XII	$\frac{dy}{dx} = \frac{p(a_1+a_2+2x)}{(a_1+x)(a_2-x)}$	$y = y_0 \left(\frac{a_1+x}{_2-x}\right)^p, -a_1 x \leq a_2.$

1.1.2. Generalized Pearson system and other systems of distributions

Following Pearson’s work, the Pearson system was expanded into a broader family of distributions parameterized to fit a wide variety of empirical data. Governed by second-order linear differential equations, this system includes families such as the Normal, Beta, Gamma, and Logistic distributions.

In the 1940s and 1950s, researchers extended Pearson's differential equation approach by generalizing the family of distributions. One notable contribution was the development of the Pearson Type IV distribution, which added an extra parameter to the system, further enhancing its flexibility and capacity to model various types of data. Over time, this approach has been extended to more general forms; for example, the generalized Pearson system of distributions could be generated by the following differential equations:

$$\frac{1}{f(x)} \frac{df_X(x)}{dx} = \frac{\sum_{j=0}^{m} a_j x^j}{\sum_{j=0}^{n} b_j x^j},$$

Where $m, n \in N_0$ (set of non-negative integers), and the coefficients aj and bj are real parameters. By proper choices of these parameters, a vast majority of continuous probability density functions could be generated from the above generalized Pearson system of distributions. For a second-order non-homogeneous linear differential equation approach to generating cardioid distribution, we refer to Rao et al. (2011). In the interest of the readers, in what follows, for the sake of completeness, we provide in Table 1.2, a list of some systems (or families) of distributions generated as the solutions of the GPDE, developed by various authors and researchers, and also some other generalized distributions like Toranzos’ generalization of Pearson system, Burr, Dagum, Halphen, and Stoppa systems, among others. For details, see, for example, Singh (2018), Singh and Zhang (2020), Shakil et al. (2024), and references therein.

Table 1.2: Generalized Pearson and Other Systems of Distributions, cf. Singh, 2018; Singh and Zhang. 2020; and Shakil et al., 2024

Sl. No.	Systems of Distributions	Generalized Pearson and Other Linear Differential Equations
1	Halphen (1941, 1955)	Halphen system of distributions: In the 1940s, Halphen proposed a family or system of three distributions, Halphen type-A, Halphen type-B, and Halphen type inverse B. These distributions are more general than those of the Pearson family.
2	Burr (1942)	Burr system of distributions: Burr (1942) proposed a differential equation for a family of cumulative probability distributions, now known as the Burr family, that contains twelve distributions, and the last member of the family, known as the Burr XII distribution, is one of the most popular distributions in hydrometeorology and flood hydrology.
3	D’Addario (1949)	D’Addario system of distributions: D’Addario (1949) proposed a system of distributions that follows the concept of transformation and a probability generating function expressed as a differential equation. Two general distributions obtained from the D’Addario system are the Amoroso and Davis distributions.
4	Toranzos (1952)	tmof dstribions: oranzos (1952) proposed he GDE in th llo orm and stieh case, h $m = 1$: $\frac{1}{f(x)} \frac{df_X(x)}{dx} = \frac{Pm+1(x)}{Qm(x)}$, where $Pm + 1(x)$ ia olynomal of dege $m + 1$ d $Qm(x)$ is another oyomil ofdegree m, whichcan e expressed by the longdivsion of polynomal, $\frac{1}{f(x)} \frac{df_X(x)}{dx} = \alpha + \beta x + \frac{-1(x)}{Qm(x)}$, where α and β are constants.
5	Roy (1971)	Roy (1971) studied GPDE to derive five frequency curves whose parameters depend on the first seven population moments.
6	Das (1973)	Das (1973) studied GPDE when $m = 2$, $n = 3$, to derive five frequency curves whose parameters depend on the first seven population moments.
7	Dagum (1975, 1977)	The Dagum System of Distributions: The system, developed by Dagum (1975, 1977), is generated from a first-order differential equation that links the income elasticity of the cumulative distribution function (CDF) to the CDF itself. The motivation for this equation comes from the empirical observation that the income elasticity of the CDF for income and wealth is a decreasing and bounded function of the CDF.
8	Dunning and Hanson (1977)	Dunning and Hanson (1977) used GPDE in their paper on generalized Pearson distributions and nonlinear programming.
9	Esteban (1981)	Esteban's system of distributions: Esn (1981) proposed a system of distributions by charcterizing the elasticity of a PDF as etaas$\eta = -\frac{dlogf(x)}{dlogx} = -\frac{x}{f(x)} \cdot \frac{df(x)}{dx}$.
10	Cobb et al. (1983)	Cobb et al. (1983) extended Pearson's class of distributions to generate multimodal distributions by taking the polynomial in the numerator of GPDE of degree higher than one and the denominator, say $v(x)$, having one of the following forms: (I) $v(x) = 1, \quad -\infty < x < \infty$, (II) $v(x) = x, \quad 0 < x < \infty$, (III) $v(x) = x^2, \quad 0 < x < \infty$, (IV) $v(x) = x(1-x), \quad 0 < x < 1$.
11	Stoppa (1990)	Stoppa system of distributions: By using a slightly different definition of the elasticity of a distribution function used in econometrics as the derivative of the logarithm of distribution function with respect to the derivative of the logarithm of its argument, Stoppa (1990) proposed a system of distributions by assuming that (i) the distributions were either non-modal or unimodal and positively skewed; (ii) the domain of the distribution was $(x_0, x_0$ was the lower bound; and (iii) the elasticity of the distribution function $F(x)$was a monotonic decreasing function of $F(x)$which depended on the lative development of $1 - F(x)$and it diverged as x tended to x_0 and tended to zero as x tended to infinity.
12	Chaudhry and Ahmad (1993)	Chaudhry and Ahmad (1993) studied another class of generalized Pearson distribution based on GPDE (3.1), when $m = 4, n = 3, b_0 = b_1 = b_2 = 0, \frac{a_4}{2b_3} = -2\alpha, \frac{a_0}{2b_3} = 2\beta, b_3 \neq 0$.
13	Lefevre et al. (2002)	Lefevre et al. (2002) studied characterization problems based on some generalized Pearson distributions using GPDE.
14	Sankaran (2003)	Sankaran (2003) considered the following case of GPDE $\frac{1}{f(x)} \frac{df_X(x)}{dx} = \frac{a_0 + a_1 x + a_2 x^2}{b_0 + b_1 x + b_2 x^2}$, And proposed a new class of probability distributions.

15	Stavroyiannis and Stavroulakis (2007)	Stavroyiannis and Stavroulakis (2007) studied generalized Pearson distributions based on GPDE in the context of super-statistics with non-linear forces and various distributions.
16	Rossani and Scarfone (2009)	Rossani and Scarfone (2009) have studied the GPDE in the following form and used it to generate generalized Pearson distributions in order to study charged particles interacting with an electric and/or a magnetic field.
17	Shakil, Kibria, and Singh (2010)	(i)
18	Shakil, Singh, and Kibria (2010)	Shakil, Singh, and Kibria (2010) defined a new class based on the following differential equation. $\frac{df_X(x)}{dx} = \left(\frac{a_0 + a_1 x + a_2 x^2}{b_1 x}\right) f_X(x), \quad b_1 \neq 0,$ Which is a special case of the GPDE when $m = 2, n = 1$, and $b_0 = 0$.
19	Shakil and Kibria (2010)	Shakil and Kibria (2010) considered the GPDE in the following form $\frac{df_X(x)}{dx} = \left(\frac{a_0 + a_p x^p}{b_1 x + b_{p+1} x^{p+1}}\right) f_X(x), \quad b_1 \neq 0, b_{p+1} \neq 0, x > 0,$ When and $b_0 = b_2 = \cdots = b_p = 0$.
20	Hamedani (2011)	Hamedani (2011) considered the GPDE in the following form $\frac{1}{f(x)} \frac{df(x)}{dx} = \frac{\beta^2 p - \beta(p+1)x^p - 2\alpha\beta p x^{2p} - \alpha(p-1)x^{3p} + \alpha^2 p x^{4p}}{\beta x^{p+1} - \alpha x^{3p+1}},$ Where and $p > 0$.
21	Ahsanullah, Shakil, and Kibria (2013)	Ahsanullah, Shakil, and Kibria (2013) defined a new class of distributions as solutions of the GPDE in the following form. $\frac{df_X(x)}{dx} = \left(\frac{a_1 + a_2 x + a_3 x^2}{b_3 x^2 + b_4 x^3}\right) f_X(x),$ When $m = 2, n = 3$.
22	Stavroyiannis (2014)	Stavroyiannis (2014) defined a new class of distributions as solutions of the GPDE in the following form. $\frac{1}{f(x)} \frac{df_X(x)}{dx} = \frac{\sum_{j=0}^{5} a_j x^j}{\sum_{j=0}^{6} b_j x^j}$, when $m = 5, n = 6$.
23	Shakil and Singh (2015)	Shakil and Singh (2015) developed a new class of generalized Pearson distribution by considering the following case of GPDE $\frac{1}{f(x)} \frac{df(x)}{dx} = -\frac{a_1 + b_1 x}{a_2 + b_2 x}$, when $m = 1, n = 1$, $a_2 = a_3 = \cdots = a_m = 0$, $b_2 = b_3 = \cdots = b_n = 0$, and where the right-hand expression, $\frac{a_1 + b_1 x}{a_2 + b_2 x}$, the ratio of two polynomials of first degree in x, with $a_1, b_1 \geq 0; a_2, b_2 > 0; a_1 b_2 \neq a_2 b_1$ (except when $a_1, b_1 = 0$).
24	Ahsanullah, Hamedani, Shakil, Kibria, and George (2016)	Ahsanullah, Hamedani, Shakil, Kibria, and George (2016) considered the following cases of the GPDE $\frac{1}{f(x)} \frac{df(x)}{dx} = \frac{\sum_{j=0}^{m} a_j x^j}{\sum_{j=0}^{n} b_j x^j}$, involving the power series and derived some new classes of univariate exponential power series distributions for a real-valued continuous random variable, called exponential power series distributions.
25	Provost, Zareamoghaddama, Ahmed, and Ha (2022)	Provost, Zareamoghaddama, Ahmed, and Ha (2022) introduced a moment-based density approximation technique that is based on a generalization of Pearson's system of frequency curves, where, the derivative of the logarithm of a continuous density function is expressed as a ratio of polynomials whose coefficients are determined by solving a linear system, and a simple representation of the resulting density function is provided.
26	Ahsanullah and Shakil (2023)	Ahsanullah and Shakil (2023) defined a new class of distributions as a solution of the GPDE in the following form. $\frac{f'(x)}{f(x)} = \frac{a_0 + a_1 x + a_2 x^2}{b_0 + b_1 x + b_2 x^2 + b_3 x^3}.$

1.1.3. Modern extensions and new approaches

In recent years, the differential equation approach has seen several innovations. One area of development is the introduction of higher-order differential equations to generate more complex distributions. For instance, new families of distributions, such as the generalized gamma and generalized normal distributions, have been derived using systems of nonlinear differential equations, giving rise to Toranzos' generalization of the Pearson system, Burr, Dagum, Halphen, and Stoppa systems. A key advancement is the integration of differential equations with the method of moments matching. Researchers have used this approach to generate distributions that match specified moments up to any order, offering a higher degree of flexibility for modeling data that exhibit non-normal characteristics such as skewness and heavy tails. For details, see, for example, Singh (2018), Singh and Zhang (2020), and references therein. Furthermore, for a second-order non-homogeneous linear differential equation approach to generating cardioid distribution, we refer to Rao et al. (2011).

1.1.4. The use of differential equations in stochastic processes

Another significant area of research involves using differential equations to generate distributions arising from stochastic processes. For example, the Langevin equation (Langevin, 1908), a stochastic differential equation, is frequently used to model physical processes such as Brownian motion. The probability distributions derived from these equations, such as the Wiener distribution, play an essential role in fields like statistical physics and finance.

Recent work has also explored the connections between partial differential equations (PDEs) and probability distributions. The Fokker-Planck equation, which is a type of PDE, has been employed to describe the evolution of probability densities in systems subject to random influences.

1.1.5. Comparative analysis and literature gaps

A comparative study could be designed to evaluate the performance of classical Pearson types (I, III, IV, VII) and selected generalized Pearson families against modern flexible alternatives such as the generalized gamma, generalized beta, and metalog distributions. Real datasets from hydrology (flood peaks), finance (returns), and biomedical data (survival times) can be used alongside simulated data. Evaluation should include maximum likelihood and Bayesian estimation, model selection via AIC/BIC, and predictive assessment using cross-validation and tail-focused metrics. A comprehensive comparative analysis and literature gap are presented in Table 1 below:

Table 1.3: Comparative Analysis

Aspect	Classical Approaches (Pearson, 1895–1950s)	Modern Approaches (1980s – Present)	Identified Gaps / Future Directions
Type of Differential Equation	First or second-order linear ODEs	Nonlinear / higher-order ODEs and PDEs	Lack of a unified framework linking ODE and PDE-based methods
Distribution Families	Normal, Beta, Gamma, Logistic, Pearson Types I–IV	Burr, Dagum, Halphen, Generalized Gamma, Cardioid, etc.	Need for multi-parameter systems accommodating multimodality.
Application Areas	Biometrics, demography, and early economics	Finance, environmental sciences, and reliability modeling	Limited exploration in AI-driven generative modeling
Estimation Methods	Analytical moment fitting	MLE, Bayesian, and computational optimization	Integration of machine learning with DE-based generation remains underexplored.
Computational Tools	Manual derivations	Symbolic computation, simulation, and numerical solvers	Need for automated DE-based distribution discovery algorithms

Summary of Gaps

1) Limited research connects stochastic PDEs and distribution generation frameworks.
2) Few studies address multi-modal and asymmetric structures using DE-based derivations.
3) Insufficient integration between computational intelligence and differential equation modeling.
4) Lack of comparative evaluations of DE-based methods versus transformation-based or compounding approaches.

2. Mathematical Formulation

2.1. Derivation of PDF and CDF:

Theorem 2.1: *For a positive continuous random variable X, the solution of a second-order nonlinear non-homogeneous ordinary differential equation of the following form*

$$\left(\frac{1}{y}\right)\frac{d^2y}{dx^2} - \left\{\left(\frac{1}{y}\right)\frac{dy}{dx}\right\}^2 = g(x),$$

Where $g(x)$ is an arbitrary function of x, admits a probability density function $y = f(x)$, given by

$$y = C \cdot \exp\left(q(x)\right),\ x > 0,$$

Where C Is the normalizing constant, and $q(x) = \int\left(\int g(x)dx\right)dx$.

Proof. For a positive continuous random variable XConsider a second-order nonlinear non-homogeneous ordinary differential equation in the following form:

$$\left(\frac{1}{y}\right)\frac{d^2y}{dx^2} - \left\{\left(\frac{1}{y}\right)\frac{dy}{dx}\right\}^2 = g(x), \quad (1)$$

Where $y = f(x)$ Is the probability density function of a continuous random variable X, and $g(x)$ is an arbitrary function of x. The equation (1) can easily be reduced as follows:

$$\frac{d}{dx}\left(\frac{y'}{y}\right) = g(x),$$

Or

$$\frac{d}{dx}\left(\ln(y)\right)' = g(x),$$

Or

$$\frac{d^2\ln(y)}{dx^2} = g(x). \quad (2)$$

Integrating equation (1) with respect to x Once, we obtain

$$\frac{d}{dx}\ln(y) = q_1(x) + k_1, \quad (3)$$

Where $q_1(x) = \int g(x)dx$, and k_1 It is an arbitrary constant. Now, integrating equation (3) with respect to x , we obtain

$$\ln(y) = q(x) + k_1x + k_2, \quad (4)$$

Where $q(x) = \int q_1(x)dx$, and k_2 It is another arbitrary constant. Exponentiating both sides of the equation (4) gives:

$$y = C \cdot \exp\left(q(x)\right) \cdot \exp\left(k_1 x\right),\ x > 0, \tag{5}$$

Where $C = \exp(k_2)$. Since y is intended to be a valid probability density function (pdf) in (5) satisfying the requirement $\int y(x)\,dx = 1$, and conditions such as $y(x) > 0$ Or behavior at boundaries, we may implicitly set $k_1 = 0$ In (5), thus fixing one of the constants, that is, $C = \exp(k_2)$. While this may still lead to a valid solution if our y is intended to be a PDF, where the final normalization condition removes extra constants - mathematically, the general solution should account for both constants unless additional conditions or constraints (like normalizability or boundary conditions) eliminate them. Thus, in view of this, we set $k_1 = 0$ In (5), fix the other constant. $C = \exp(k_2)$ So that it is the normalizing constant. Hence, the solution (5) of equation (1) is reduced to

$$y = C \cdot \exp\left(q(x)\right),\ x > 0,$$

Which is a valid and required probability density function, with $C = \exp(k_2)$ as the normalizing constant, and where $q(x) = \int(\int g(x)dx)dx$, and $g(x)$ is an arbitrary function of x. This completes the proof of Theorem 2.1.

Theorem 2.2: *For a positive continuous random variable X, the solution of a second-order nonlinear non-homogeneous ordinary differential equation of the following form*

$$\left(\frac{1}{y}\right)\frac{d^2y}{dx^2} - \left\{\left(\frac{1}{y}\right)\frac{dy}{dx}\right\}^2 = \frac{a_p(1-p)\,x^p - a_0}{b_1 x^2}, \tag{6}$$

Where $\frac{a_p(1-p)\,x^p - a_0}{b_1 x^2}$ is a rational function of x, $b_1 x^2 \neq 0$, for any x, the coefficients a_0, a_p, and b_1 ($\neq 0$) They are real parameters, and we assume that $p > 0$. The equation (6) admits a PDF. $y = f(x)$ of the continuous random variable X, Whose specific value is x, given by

$$y = f(x) = Cx^{\nu - 1}\exp\left(-\alpha x^p\right),\ x > 0,\ \alpha > 0, \nu > 0, p > 0, \tag{7}$$

Where $\alpha = -\frac{a_p}{pb_1}$, $\nu = \frac{a_0 + b_1}{b_1}$, $b_1 \neq 0$, $p > 0$, and C Is the normalizing constant given by

$$C = \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)}. \tag{8}$$

Proof. In equation (6), let $g(x) = \frac{a_p(1-p)\,x^p - a_0}{b_1 x^2}$, so that

$$\left(\frac{1}{y}\right)\frac{d^2y}{dx^2} - \left\{\left(\frac{1}{y}\right)\frac{dy}{dx}\right\}^2 = g(x), \tag{9}$$

Which is obviously the equation (1) of Theorem 2.1. Thus, in view of the arguments as provided in Theorem 2.1, the equation (9) admits a PDF $y = f(x)$, given by

$$y = C \cdot \exp\left(q(x)\right),\ x > 0, \tag{10}$$

Where C Is the normalizing constant, $q(x) = \int(\int g(x)dx)dx$, and $g(x) = \frac{a_p(1-p)\,x^p - a_0}{b_1 x^2}$. Now, the right side of the equation (10) can easily be expressed as

$$q(x) = \int(\int g(x)dx)dx = \int\left(\int \frac{a_p(1-p)\,x^p - a_0}{b_1 x^2}dx\right)dx$$

$$= \int\left(\frac{a_0 - a_p x^p}{b_1 x}\right)dx = \int\left(\frac{a_0}{b_1 x} - \frac{a_p x^p}{b_1 x}\right)dx$$

$$= \ln(C) + \frac{a_0}{b_1}\ln(x) - \frac{a_p x^p}{b_1 p}, \tag{11}$$

Where C It is a constant of integration to be determined. Substituting the expression (11) for $q(x)$ In equation (10), we obtain

$$y = f(x) = Cx^{\frac{a_0}{b_1}}\exp\left(-\frac{a_p x^p}{b_1 p}\right). \tag{12}$$

Letting $\alpha = -\frac{a_p}{pb_1}$, $\nu = \frac{a_0 + b_1}{b_1}$, $b_1 \neq 0$, $p > 0$In equation (12), we obtain

$$y = f(x) = Cx^{\nu - 1}\exp\left(-\alpha x^p\right),\ x > 0,\ \alpha > 0, \nu > 0, p > 0, \tag{13}$$

Where $\alpha = -\frac{a_p}{pb_1}$, $\nu = \frac{a_0 + b_1}{b_1}$, $b_1 \neq 0$, $p > 0$, and C It is a constant of integration. Now, integrating both sides of the equation (13) with respect to x , from $x = 0$ to ∞, using the formula from Gradshteyn and Ryzhik (2000), equation 3.478.1, page 342, we obtain

$$\int_0^\infty x^{\nu-1}\exp(-\alpha x^p)\,dx = \frac{\alpha^{-\frac{\nu}{p}}}{p}\Gamma\left(\frac{\nu}{p}\right), (\alpha > 0,\ \nu > 0, p > 0),$$

Where $\Gamma(.)$ denotes the gamma function, and noting that $\int_0^\infty f(x) = 1$, (since f(x) defines a probability density function, we easily obtain

$$C = \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)}. \tag{14}$$

Substituting the expression (14) for the constant C In equation (13), we have

$$y = f(x) = \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} x^{\nu-1}\exp\left(-\alpha x^p\right),\ x > 0,\ \alpha > 0, \nu > 0, p > 0, \tag{15}$$

Which is the required solution of the equation (6), and represents the necessary PDF of the suggested generalized class of the generalized gamma distribution (GGGD), where the coefficients: $p > 0$ indicates the shape parameter, $\nu > 0$, responsible for skewness, and $\alpha > 0$ Serves as a scale parameter. Its cumulative distribution function is specified and given by the expression. $G(x) = \int_0^x g(t)dt$ Which gives the probability of failure before a specific value x So for the equation (15), it is expressed, in terms of the incomplete gamma function, as

$$F(x) = \int_0^x \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} t^{\nu-1}\exp\left(-\alpha t^p\right)dt = \frac{\gamma\left(\frac{\nu}{p}, \alpha x^p\right)}{\Gamma\left(\frac{\nu}{p}\right)}, x > 0,\ \alpha > 0, \nu > 0, p > 0, \tag{16}$$

Which, on differentiation, yields equation (15). In this regard, it is noted that $\frac{\partial\gamma(a,t)}{\partial t} = t^{a-1}e^{-t}$. The possible shapes of the PDF (15) of $X \sim GGGD(\alpha, \nu, p)$ Are provided below for some selected values of the parameters in Figure 1 (panels a and b) and Figure 2 (panels c and d), respectively.

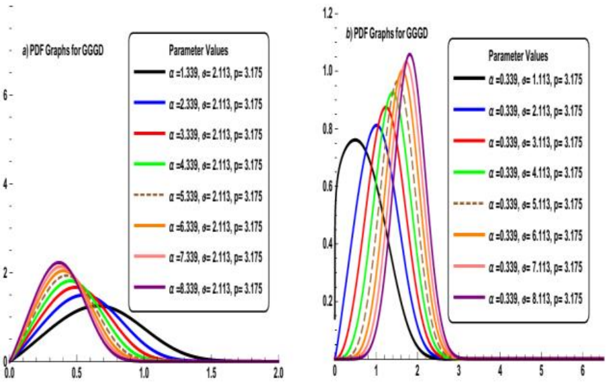

Fig. 1: (Panels a, b): PDF Graphs of GGGD for Different Values of α, ν, p.

The pattern of the probability curve shown in Figure 1 (panels a and b) indicates that the form of the distribution is significantly affected by ν, influencing both the position of the maximum and the dispersion. The GGGD is capable of representing asymmetric distributions, offering versatility for use in fields such as finance, reliability assessment, and survival analysis. Smaller ν values produce more pronounced peaks, concentrating the likelihood near lower x values, whereas larger ν values move the distribution to the right, distributing the probability over a wider interval.

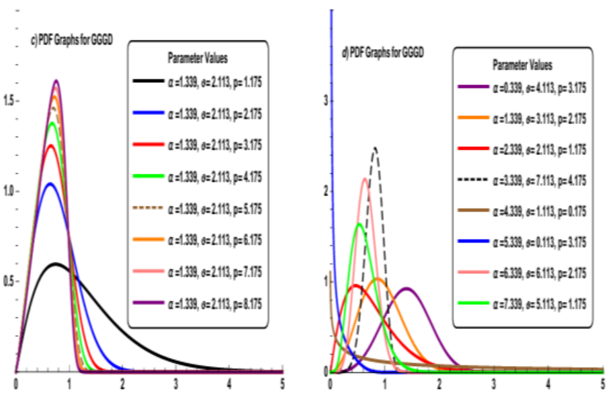

Fig. 2: (Panels C and D): PDF Graphs of GGGD for Different Values of A, N, P.

In Figure 2 (panel d) on the right, the illustration demonstrates the impact of adjusting α while maintaining constant ν and p. As α rises, the distribution's maximum shifts slightly to the right and becomes more prominent. The rate at which the distribution diminishes varies, suggesting that α plays a role in concentrating prob- ability around lower values. Conversely, Figure 2 (panel d) on the right examines how different combinations of α, ν, and p influence the PDF. Some curves display symmetrical characteristics; for instance, the black curve (corresponding to the lowest α) has the tallest peak at a larger x, implying that the probability mass is centered at higher values relative to other distributions. Other curves feature peaks nearer to zero, reflecting a more pronounced skewness.

Furthermore, the modal value of the PDF (15) can be obtained by solving the equation. $\frac{df_X(x)}{dx} = 0$, which after simplification yields $\frac{(-\alpha p x^p + (\nu - 1))f_X(x)}{x} = 0$, or $\alpha p x^p - (\nu - 1) = 0$. Now solving for x, we get $x = \sqrt[p]{\frac{(\nu - 1)}{\alpha p}}$, provided $\nu > 1$. However, the second derivative of the PDF (15) is less than zero, i.e., $\frac{d^2 f(x)}{dx^2} < 0 \Rightarrow \frac{f(x)(-p(\nu-1))}{x^2} < 0$. This implies that the proposed distribution, GGGD, is unimodal, which can also be visualized from Figures 1 and 2, respectively.
Remarks: In what follows, we discuss the case $p = -1$ Separately.
Proposition 2.1. (The Case $p = -1$):
Let the ODE from Theorem 2.2 be given by:
$\frac{1}{y}\frac{d^2y}{dx^2} - \frac{1}{y}\left(\frac{dy}{dx}\right)^2 = \frac{(ap(1-p)x^p - a_0)}{b_1 x^2}, x > 0$, and take $p = -1$. Define $\alpha = \frac{-a_p}{pb_1}$ and $\nu = \frac{-a_0 + b_1}{b_1}$.
For $p = -1$These reduce to $\alpha = \frac{a_p}{b_1}$ and $\nu = (a_0 + b_1)/b_1$. Then the ODE admits a candidate density:

$$f(x) = C\,x^{\nu-1}\exp\left(-\frac{\alpha}{x}\right), x > 0,$$

Which is a proper PDF if and only if α > 0 and ν < 0. When these hold, the normalization constant is

$$C = \frac{1}{\alpha^\nu \Gamma(-\nu)}.$$

Equivalently, in terms of (a_P, a_0, b_1), The required inequalities are:

$$\frac{a_{P,}}{b_1} > 0 \text{ and } (a_0 + b_1)/b_1 < 0.$$

Proof. With $p = -1$, the inner computation of Theorem 2.2 yields (up to a multiplicative constant):

$$y(x) = C\,x^{\nu-1}\exp\left(-\frac{\alpha}{x}\right), x > 0, \text{ where } \alpha = \frac{a_p}{b_1} \text{ and } \nu = (a_0 + b_1)/b_1.$$

To be a valid probability density function, we require $\int_0^\infty y(x)\,dx = 1$. Compute the integral (assuming α > 0):

$$\int_0^\infty x^{\nu-1}\exp\left(-\frac{\alpha}{x}\right)dx = 1.$$

Let $u = \alpha/x \Rightarrow x = \alpha/u$ and $dx = -\alpha\,u^{-2}\,du$. Then:

$$\int_0^\infty x^{\nu-1}\exp\left(-\frac{\alpha}{x}\right)dx = \alpha^\nu \int_0^\infty u^{-\nu-1}\,e^{-u}du = \alpha^\nu\,\Gamma(-\nu).$$

This integral is finite when $\alpha > 0$ and $\nu < 0$, since $\Gamma(-\nu)$ is finite for $-\nu > 0$. Therefore, the normalization constant is: $C = 1/(\alpha^\nu\,\Gamma(-\nu))$, and $f(x) = C\,x^{\nu-1}\exp\left(-\frac{\alpha}{x}\right)$ is a proper pdf if α > 0 and ν < 0.
The sign constraints in parameter form are: $\alpha = a_p/b_1 > 0 \Rightarrow a_p$ and b_1 Have the same sign.

$$\nu = (a_0 + b_1)/b_1 < 0 \Rightarrow a_0 < -b_1 \text{ if } b_1 > 0, \text{or } a_{10} > -b_1 \text{ if } b_1 < 0.$$

For integrability:
As $x \to 0^+$, $e^{-\alpha/x} \to 0$ super-exponentially ⇒ integrable and as $x \to \infty$, $e^{-\alpha/x} \to 1$, And the tail behaves like $x^{\nu-1}$; integrability ⇒ $\nu < 0$. Hence, the proposition is proved.

2.2. Distributional relationships

For some special values of the parameters in $X \sim GGGD(\alpha, \nu, p)$We obtain various special cases of the GGGD, namely, generalized gamma, chi-squared, chi, Erlang, exponential, gamma, Weibull, Rayleigh, Maxwell-Boltzmann, half normal, and log-normal. Moreover, a comprehensive comparison between GGGD, Burr Type XII, and Dagum is stated in Table 2.1.

Table 2.1: Comparative Summary of GGGD, Burr (XII), and Dagum Distributions

Feature	GGGD (Generalized Generalized Gamma)	Burr Type XII	Dagum
Functional Form	(x) = C x^(ν − 1) e^(−αx^p), where C = pα^(ν/p)/Γ(ν/p)	F(x) =1−(1+(x/λ)^c)^(−k)	F(x)=(1+(x/b)^(−a))^(−p)
Tail Behavior	Exponential / stretched-exponential (light tail)	Power-law / Pareto-like (heavy tail)	Power-law / Pareto-like (heavy tail)
Moment Existence	All moments finite; closed-form via the Gamma function	Finite if tail index > moment order	Fne· p > moment order
Hazard Rate Shape	Increasing, decreasing, bathtub, or unimodal (depends on p, ν)	Very flexible; can reproduce most shapes	Similar flexibility; often used for long-tail failure rates
Special / Nested Cases	Gamma, Weibull, Exponential, Rayleigh, Maxwell, Normal	Pareto, Log-logistic, Weibull	Burr, Log-logistic, Fisk, Pareto
Tail Index Control	Exponential rate p; light-tail decay	Product c·k controls heaviness	Product a·p controls heaviness

Moment Computation	E[X^r] =α^(−r/p)Γ((ν+r)/p)/Γ(ν/p)	Closed-form using Beta functions	Closed-form using Beta functions
Estimation (MLE)	Smooth likelihood, numerically stable	May face flat likelihood in tail; sensitive to extremes	Similar to Burr, sensitive to small samples
Quantile / Simulation	No closed-form inverse; requires numerical inversion	Closed-form quantile → easy sampling	Closed-form quantile → easy sampling
Empirical Domains	Reliability, survival, hydrology, biomedical, and environmental	Income, insurance losses, risk, reliability (heavy tails)	Income/wealth, demography, actuarial risk
Computational Complexity	Moderate (requires incomplete gamma)	Low (simple algebraic CDF)	Low (simple algebraic CDF)
Goodness-of-Fit Tools	AIC/BIC, QQ, PP, residual, tail plots	AIC/BIC, QQ, tail index diagnostics	AIC/BIC, Lorenz, and Gini analysis
Model Strengths	Smooth hazard flexibility, closed-form moments, exponential-type tail	High flexibility, closed-form quantile, heavy-tail modeling	Ideal for power-law data, economic interpretation
Main Limitation	Cannot capture power-law tails	May overfit the tail or be unstable for small n	Similar sensitivity; may underfit light tails
Best Use Case	Light-tailed lifetime or environmental data	Heavy-tailed income, risk, or survival data	Socio-economic and actuarial heavy-tail analysis
Extension Possibilities	Mixtures, spliced models (GGGD + Burr tail), Bayesian inference	Hybrid models, truncated variants	Hierarchical Bayesian, multivariate Dagum

2.3. Reliability analysis

The GGG distribution serves as a useful model for lifetime data within reliability engineering. Its adaptability enables it to capture various failure rate behaviors. For instance, the cumulative hazard function, $H(x) = -\ln(S(x))$, and Mill's Ratio, $m(x) = \frac{S(x)}{f(x)}$, are key tools for analyzing the overall risk of failure or event occurrence. These functions are closely linked to the survival function, making them vital for modeling time-to-event data and conducting statistical inference in survival analysis. For $X \sim GGGD(\alpha, \nu, p)$, the reliability (or survival) function, $R(x)$, hazard rate (or failure) function, $H(x)$, cumulative hazard function, $H(x)$, and Mill's Ratio, $m(x)$, are derived, respectively, as follows:

$$S(x) = 1 - F(x) = 1 - \frac{1}{\Gamma\left(\frac{\nu}{p}\right)} \gamma\left(\frac{\nu}{p},\ \alpha x^{p}\right), \tag{11}$$

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{\frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} x^{\nu-1} \exp\left(-\alpha x^{p}\right)}{1 - \frac{1}{\Gamma\left(\frac{\nu}{p}\right)} \gamma\left(\frac{\nu}{p}, \alpha x^{p}\right)}, \tag{12}$$

$$H(x) = -\ln(S(x)) = -\ln\left(1 - \frac{1}{\Gamma\left(\frac{\nu}{p}\right)} \gamma\left(\frac{\nu}{p},\ \alpha x^{p}\right)\right), \tag{13}$$

And

$$m(x) = \frac{S(x)}{f(x)} = \frac{1 - \frac{1}{\Gamma\left(\frac{\nu}{p}\right)} \gamma\left(\frac{\nu}{p}, \alpha x^{p}\right)}{\frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} x^{\nu-1} exp\left(-\alpha x^{p}\right)} \tag{14}$$

Where $x > 0,\ \alpha > 0, \nu > 0,\ p > 0$. The possible shapes of the hazard rate (or failure) function (hf) are provided for some selected values of the parameters in Figures 3 and 4, respectively.

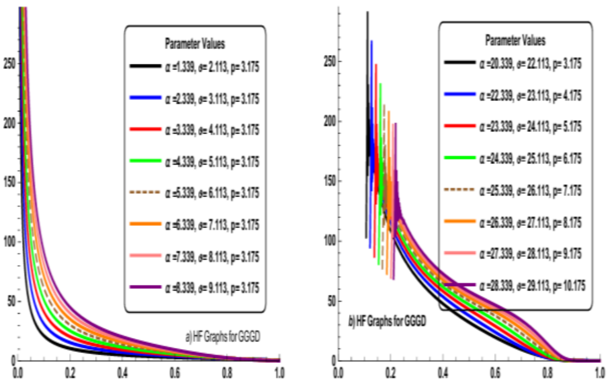

Fig. 3: hf Graphs of GGGD for Different Values of α, ν, p.

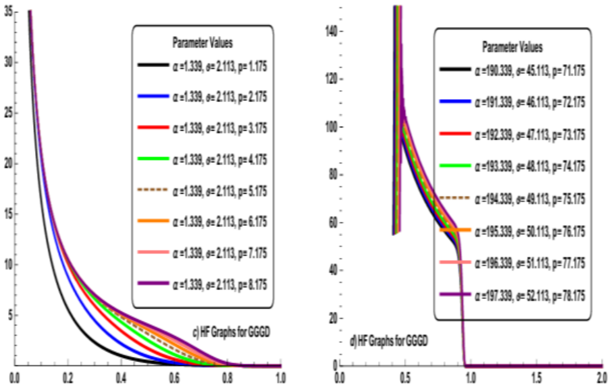

Fig. 4: hf Graphs of GGGD for Different Values of α, ν, p.

Based on Figures 3 and 4, it is clear that the GGGD shows a declining hazard rate, indicating that items or systems described by this distribution are more prone to failure at the beginning, with the risk diminishing as time progresses. The parameter values play a significant role in influencing the rate of this decline, which is valuable in the fields of reliability and survival analysis. For larger values of x, the hazard rate levels off, implying that the likelihood of failure becomes nearly constant. Initially, the hazard function is very high and then decreases gradually, eventually reaching a steady state at lower levels. This behavior points to a decreasing hazard rate, where the chance of failure or an event occurring diminishes over time. Additionally, the hazard function begins at a high point and undergoes some fluctuations before stabilizing, which may reflect an initial phase of high variability in risk before settling into a consistent pattern.

2.4. Moments and moment generating function

In this section, we will evaluate the formulas for the moments and the moment generating function of the distribution, which are crucial in statistical analysis, particularly in practical applications. It is well understood that key features of a distribution, such as central tendency, variability, skewness, and kurtosis, can be examined via its moments.

Proposition 2.2. If the random variable X Is GGG distributed? $X \sim \mathrm{GGGD}(\alpha, \nu, p)$, then it's kth The moment around zero can be expressed as

$$E(X^k) = \frac{\alpha^{-\frac{k}{p}}\Gamma\left(\frac{k+\nu}{p}\right)}{\Gamma\left(\frac{\nu}{p}\right)}. \tag{15}$$

Proof. The moment definition implies

$$E(X^k) = \int_0^\infty x^k \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} x^{\nu-1} \exp\left(-\alpha x^p\right) dx. \tag{16}$$

Using Equation 3.478.1, Page 342, of Gradshteyn and Ryzhik (2000) in equation (16), we obtain

$$E(X^k) = \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} \frac{\alpha^{-\frac{k+\nu}{p}}}{p} \Gamma\left(\frac{k+\nu}{p}\right). \tag{17}$$

On simplifying equation (17), we obtain (15).

Corollary 1. If $X \sim \mathrm{GGGD}(\alpha, \nu, p)$, then, for $k = 1, 2$ In (15), the mean and variance are, respectively, given as

$$E(X) = \frac{\alpha^{-\frac{1}{p}}\Gamma\left(\frac{\nu+1}{p}\right)}{\Gamma\left(\frac{\nu}{p}\right)} \tag{18}$$

Corollary 2. Central Moment: The kth central moment, β_k, of $X \sim \mathrm{GGGD}(\alpha, \nu, p)$ is given by

$$\beta_k = E[X - E(X)]^k = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \left(E(X)\right)^j E(X^{k-j}), \tag{19}$$

Where $\left(E(X)\right)^j$ and $E(X^{k-j})$ They are given by equation (15). For $k = 2$ In (19), and using equation (18), the variance (second moment) is obtained as

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{\alpha^{-\frac{2}{p}}\left[-\left(\Gamma\left(\frac{\nu+1}{p}\right)\right)^2 + \Gamma\left(\frac{\nu}{p}\right)\Gamma\left(\frac{\nu+2}{p}\right)\right]}{\left[\Gamma\left(\frac{\nu}{p}\right)\right]^2}. \tag{20}$$

Corollary 2. Skewness and Kurtosis: Skewness and kurtosis are derived analytically by using the moments. The distribution of $X \sim \mathrm{GGGD}(\alpha, \nu, p)$ Provides significant flexibility in these measures,
making it suitable for modeling asymmetrical data. For $X \sim \mathrm{GGG}(\alpha, \nu, p)$, the measure of skewness, γ_1, and kurtosis, γ_2, are, respectively, given as

$$\gamma_1 = \frac{\beta_3}{(\beta_2)^{3/2}}, \tag{21}$$

And

$$\gamma_2 = \frac{\beta_4}{(\beta_2)^{2}}, \tag{22}$$

Where the β_2, β_3 and β_4, are obtained from (19) by taking $k = 2,3,4$, Respectively. Figure 5 below displays the skewness and kurtosis of the GGG distribution as a function of α, ν, p.

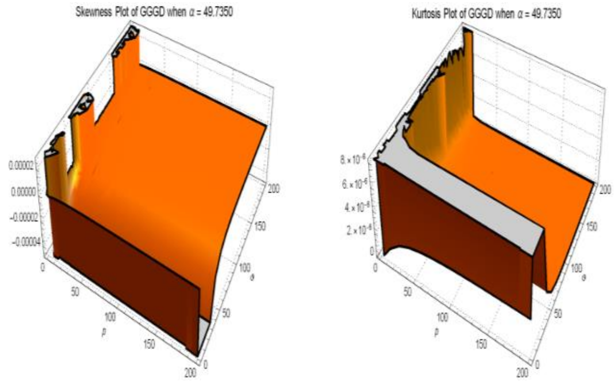

Fig. 5: Skewness and Kurtosis of GGGD.

Note that since skewness evaluates the degree of asymmetry within a distribution, the plot displayed on the left side of Figure 5 indicates that skewness depends on the parameters ν and p. The prominent peaks suggest that skewness attains notably high values for specific parameter combinations, reflecting pronounced asymmetry. Conversely, in the flatter regions, skewness approaches zero, indicating a distribution that is more symmetric. Additionally, kurtosis assesses the "peakedness" or "tailedness" of the distribution. Elevated values in certain areas on the right side of Figure 5 point to the existence of heavy tails, characteristic of leptokurtic distributions. Meanwhile, the lower regions imply that, for some parameter settings, the distribution exhibits lighter tails, indicative of platykurtic behavior.

Proposition 2.3. If the random variable X is GGG distributed, $X \sim GGGD(\alpha, \nu, p)$, then its moment generating function is given by

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{\alpha^{-\frac{k}{p}} \Gamma\left(\frac{k+\nu}{p}\right)}{\Gamma\left(\frac{\nu}{p}\right)}. \tag{23}$$

Proof. The moment generating function's definition implies

$$M_X(t) = E(e^{tX}) = \int_0^{\infty} e^{tx} \frac{p\, \alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} x^{\nu-1} \exp\left(-\alpha x^p\right) dx,$$

From which, using the Taylor series for e^{tx}: for all $x \in \mathbb{R}$, the integral becomes

$$M_X(t) = \frac{p\, \alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_0^{\infty} x^{k+\nu-1} \exp\left(-\alpha x^p\right) dx. \tag{24}$$

Using Equation 3.478.1, Page 342, of Gradshteyn and Ryzhik (2000) in equation (24), and simplifying, we obtain (23).

Proposition 2.4. If the random variable X is GGG distributed, $X \sim GGGD(\alpha, \nu, p)$, then it's jth incomplete moment, $I_j(x)$, is given by

$$I_j(x) = \frac{\alpha^{-\frac{j}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} \gamma\left(\frac{\nu+j}{p}, \alpha x^p\right), j > 0, x > 0, \alpha > 0, \nu > 0, p > 0. \tag{25}$$

Proof. The jth The incomplete moment's definition implies

$$I_j(x) = \int_0^{x} t^j \left[\frac{p\, \alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} t^{\nu-1} \exp\left(-\alpha t^p\right)\right] dt. \tag{26}$$

Substituting $\alpha t^p = u$ in (26), and using the definition of the incomplete gamma function, $\gamma(b, x) = \int_0^x t^{b-1} e^{-t} dt,\ b > 0$ the *jth* incomplete moment of $X \sim GGGD(\alpha, \nu, p)$ is easily obtained as

$$I_j(x) = \frac{\alpha^{-\frac{j}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} \gamma\left(\frac{\nu+j}{p}, \alpha x^p\right), j > 0, x > 0, \alpha > 0, \nu > 0, p > 0.$$

2.5. Quantile

For any $0 < q < 1$, the quantile of order q (that is, the 100qth percentile) of $X \sim GGGD(\alpha, \nu, p)$ is a number x_q such that the area under $f_X(x)$ to the left of x_q is q, which is any root of the equation $F(x_q) = \int_0^{x_q} f_X(u)du = q$. Thus, the percentage points x_q of the distribution of $X \sim GGGD(\alpha, \nu, p)$ are computed by numerically solving the equation $F(x_q) = \int_0^{x_q} f_X(u)du = q$, for different sets of values of the parameters as provided in Table 2.2.

Table 2.2: Percentile Points of $X \sim GGGD(\alpha, \nu, p)$

Parameters / Percentiles q		0.75	0.80	0.85	0.90	0.95	0.99
$\alpha = 0.2$, $\nu = 1.5, p = 2$	x_q	2.273864	2.479476	2.725375	3.043132	3.528525	4.474624
$\alpha = 0.5$, $\nu = 1.5, p = 2$	x_q	1.438118	1.568159	1.723680	1.924645	2.231635	2.830001
$\alpha = 1$, $\nu = 1.5, p = 2$	x_q	1.016903	1.108856	1.218825	1.360930	1.578004	2.001113
$\alpha = 2$, $\nu = 1.5, p = 2$	x_q	0.719059	0.784079	0.861839	0.962323	1.115818	1.415000
$\nu = 2$, $\alpha = 0.5$, $p = 1.5$	x_q	2.379834	2.594070	2.854552	3.198173	3.738461	4.844288
$\nu = 3$, $\alpha = 0.5$, $p = 2$	x_q	2.026905	2.154444	2.305872	2.500278	2.795483	3.368214
$\nu = 3.5$, $\alpha = 1$, $p = 2.5$	x_q	1.299076	1.366355	1.445224	1.544983	1.693560	1.973172
$\nu = 4$, $\alpha = 2$, $p = 3$	x_q	0.971822	1.014622	1.064345	1.126586	1.218035	1.386527

2.6. Entropy

The Maximum Entropy Principle (MEP) is a technique used to determine probability distributions by choosing the distribution that maximizes entropy subject to specified constraints. This approach guarantees that no extra assumptions or biases are incorporated beyond the known information. The principle was proposed by Jaynes (1957) within the context of information theory and statistical mechanics. The entropy of an absolutely continuous random variable X It was proposed by Shannon (1948). It measures the degree of uncertainty or variability and is used to quantify the amount of information contained in a random observation regarding its parent distribution (population). A higher entropy value signifies greater unpredictability or uncertainty in the data. In this context, the Shannon entropy of the random variable $X \sim GGGD(\alpha, \nu, p)$ With probability density function (15) is derived as follows:

$$H[X] = E[-\ln(f(X)] = -\int_0^\infty \ln[f(x)]f(x)\, dx$$

$$= -\int_0^\infty \ln\left[\frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma(\frac{\nu}{p})} x^{\nu-1}\exp(-\alpha x^p)\right]\left[\frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma(\frac{\nu}{p})} x^{\nu-1}\exp(-\alpha x^p)\right] dx$$

$$= \alpha E(X^p) - (\nu - 1)E(\ln X) - \ln\left[\frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma(\frac{\nu}{p})}\right]. \quad (27)$$

The figure (6) below illustrates the behavior of Shannon entropy (27). Obviously, it tends to rise as both p and ν Increase, reflecting greater unpredictability at higher parameter values.

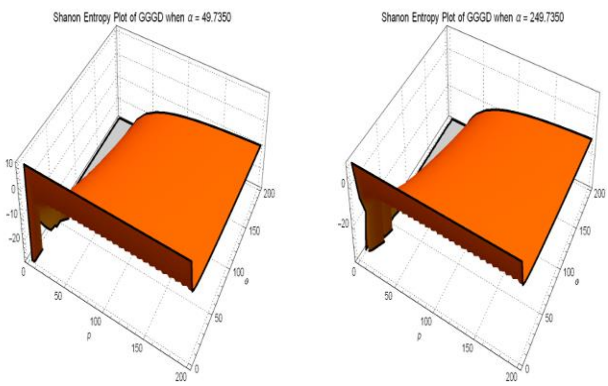

Fig. 6: Shannon Entropy Plots of GGG Distribution.

Furthermore, the entropy seems to level off at larger values of p and ν, implying that beyond a certain threshold, the degree of randomness remains relatively unchanged. The left graph corresnds to $\alpha = 4949$he right graph corresnds to $\alpha = 2450$. As α grows, the overall entropy landscape shifts slightly, revealing varying degrees of uncertainty. Typically, higher αvalues result in elevated entropy, indicating a broader spread in the distribution. A pronounced decline at low p and ν values suggests that for very small shape parameters, the entropy diminishes significantly, implying the distribution becomes more deterministic. This scenario may represent cases where the distribution approaches a degenerate or exponentially dominated form. Elevated entropy values point to a more dispersed distribution with increased randomness, whereas lower entropy indicates a more concentrated distribution. The entropy behavior is influenced by pνand α, with larger parameters generally producing higher entropy. Overall, the plots effectively demonstrate how the uncertainty associated with the GGGD varies with different parameter configurations.

3. Characterization

Characterizing probability distributions involves identifying unique conditions that distinguish a
distribution from all others, as highlighted by Nagaraja (2006) and Koudou and Ley (2014), who emphasize its importance for understanding stochastic models and driving mathematical innovation. Key works on distribution characterization include contributions from Galambos and Kotz (1978), Gl¨Glanzel et al. (1984), Ahsanullah (2017), and Shakil et al. (2018, 2021), among others. Our focus here would be on the characterization of $X \sim GGGD(\alpha, \nu, p)$ Using truncated moments.

3.1. Characterization by truncated moments

In this subsection, we provide two new characterization results of $X \sim GGGD(\alpha, \nu, p)$ By truncated moments. The first characterization result (Theorem 3) is based on a relation between the left truncated moment and the failure rate function. The second characterization result (Theorem 4) is based on a relation between right right-truncated moment and the reversed failure rate function. For this, we will also need the following assumptions and lemmas, see, for example, Shakil et al. (2018, 2021).

Assumption 3.1: Assume that the random variable. X is absolutely continuous with CDF, F(x), and PDF, f (x). Let $\omega = \inf\{x \mid F(x) > 0\}$, and $\delta = \sup\{x \mid F(x) < 1\}$, represent the infimum and supremum of the support, respectively. Additionally, f (x) is differentiable for all x and the expected value E(X) Is finite.

Lemma 3.1. If the random variable X satisfies Assumption 5.1 with $\omega = 0$ and $\delta = \infty$, and if $E(X \mid X \le x) = g(x)\tau(x)$, where $\tau(x) = \frac{f(x)}{F(x)}$ and $g(x)$ is a continuous differentiable function of x with the condition that $\int_0^x \frac{u - g'(u)}{g(u)} du$ is finite, then $f(x) = ce^{\int_0^x \frac{u - g'(u)}{g(u)} du}$, where c Is a constant determined by the condition? $\int_0^\infty f(x)\,dx = 1$.

Proof. See Shakil et al. (2018).

Lemma 3.2: *If the random variable X satisfies Assumption 5.1 with $\omega = 0$ and $\delta = \infty$, and if $E(X \mid X \ge x) = \tilde{g}(x) r(x)$, where $r(x) = \frac{f(x)}{1 - F(x)}$ and $\tilde{g}(x)$ is a continuous differentiable function of x with the condition that $\int_x^\infty \frac{u + [\tilde{g}(u)]'}{\tilde{g}(u)} du$ is finite, then $f(x) = ce^{-\int_x^\infty \frac{u + [\tilde{g}(u)]'}{\tilde{g}(u)} du}$, where c Is a constant determined by the condition? $\int_0^\infty f(x)\,dx = 1$.*

Proof. See Shakil et al. (2018).

Theorem 3: *If the random variable X satisfies Assumption 5.1 with $\omega = 0$ and $\delta = \infty$, then $E(X \mid X \le x) = g(x)\frac{f(x)}{F(x)}$, where*

$$g(x) = \frac{\gamma\left(\frac{\nu + j}{p}, \alpha x^p\right)}{p\,\alpha^{\frac{\nu}{p} + \frac{j}{p}} x^{\nu - 1} \exp(-\alpha x^p)}, \tag{28}$$

If and only if X has the PDF

$$f_X(x) = \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} x^{\nu - 1} \exp(-\alpha x^p),\ x > 0,\ \alpha > 0, \nu > 0, p > 0.$$

Proof. Suppose that $E(X|X \le x) = g(x)\frac{f(x)}{F(x)}$. Then, since $E(X|X \le x) = \frac{\int_0^x u f(u)\,du}{F(x)}$, we have $g(x) = \frac{\int_0^x u f(u)\,du}{f(x)}$. Now, if the random variable X Satisfies Assumption 1 and has the distribution with the pdf (15), then, using (25), we obtain

$$g(x) = \frac{\int_0^x u f(u)\,du}{f(x)} = \frac{\gamma\left(\frac{\nu + j}{p}, \alpha x^p\right)}{p\,\alpha^{\frac{\nu}{p} + \frac{j}{p}} x^{\nu - 1} \exp(-\alpha x^p)}$$

Consequently, the proof of the “if” part of Theorem 3 follows from Lemma 1.

Conversely, suppose that.

$$g(x) = \frac{\gamma\left(\frac{\nu + j}{p}, \alpha x^p\right)}{p\ \alpha^{\frac{\nu}{p} + \frac{j}{p}} x^{\nu - 1} \exp\left(-\alpha x^p\right)}. \quad (29)$$

Now, from Lemma 1, we obtain

$$g(x) = \frac{\int_0^x u\, f(u)\, du}{f(x)},$$

Or

$$\int_0^x u\, f(u)\, du = f(x)\, g(x). \quad (30)$$

Differentiating equation (30) with respect to x, and using the Fundamental Theorem of Calculus, we obtain

$$x f(x) = f^{/}(x) g(x) + f(x) g^{/}(x)$$

From which, using the expressions for the PDF f(x) as given in (15) and $f^{/}(x) = \frac{\left(-\alpha p x^p + (\nu - 1)\right) f_X(x)}{x}$, respectively, after simplification, we have

$$\frac{x - g^{/}(x)}{g(x)} = \frac{\left(-\alpha p x^p + (\nu - 1)\right)}{x}. \quad (31)$$

Since, by Lemma 1, we have

$$\frac{x - g^{/}(x)}{g(x)} = \frac{f^{/}(x)}{f(x)} \text{, see Shakil et al. (2018),} \quad (32)$$

Therefore, from (31) and (32), we obtain

$$\frac{f^{/}(x)}{f(x)} = \frac{\left(-\alpha p x^p + (\nu - 1)\right)}{x},$$

Or,

$$\frac{d \ln\left(f(x)\right)}{dx} = \frac{(\nu - 1)}{x} - \alpha p x^{p-1}. \quad (33)$$

Integrating (33) with respect to x And simplifying, we obtain

$$\ln\left(f(x)\right) = \ln\left(c x^{\nu - 1} \exp(-\alpha x^p)\right),$$

Or

$$f(x) = c x^{\nu - 1} \exp(-\alpha x^p), \quad (34)$$

Where c Denotes the normalizing constant. Integrating (34) with respect to x from $x = 0$ to $x = \infty$, and using the condition $\int_0^\infty f(x)\, dx = 1$, we obtain

$$\frac{1}{c} = \int_0^\infty x^{\nu - 1} \exp(-\alpha x^p)\, dx. \quad (35)$$

Using the following formula of Gradshteyn and Ryzhik (2000), Equation 3.478.1, Page 342 in (35),

$$\int_0^\infty x^{\nu - 1} \exp(-\alpha x^p)\, dx = \frac{\alpha^{-\frac{\nu}{p}}}{p} \Gamma\left(\frac{\nu}{p}\right), (\alpha > 0,\ \nu > 0, p > 0),$$

We obtain

$$\frac{1}{c} = \frac{\alpha^{-\frac{\nu}{p}}}{p} \Gamma\left(\frac{\nu}{p}\right),$$

Which, on using in (34), gives

$$f_X(x) = \frac{p\ \alpha^{\frac{\nu}{p}}}{\Gamma\left(\frac{\nu}{p}\right)} x^{\nu - 1} \exp\left(-\alpha x^p\right),\ x > 0,\ \alpha > 0, \nu > 0, p > 0,$$

Which is the required PDF of $X \sim GGGD(\alpha, \nu, p)$. This completes the proof of Theorem 3.

Theorem 4: If the random variable X satisfies Assumption 1 with $\omega = 0$ and $\delta = \infty$, then $E(X|X \geq x) = \tilde{g}(x)\frac{f(x)}{1 - F(x)}$, where $\tilde{g}(x) = \frac{(E(X) - g(x)f(x))}{f(x)}$, where $g(x)$ is given by (28) and $E(X)$ is given by (18), if and only if X has the PDF

$$f_X(x) = \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma(\frac{\nu}{p})} x^{\nu - 1}\exp(-\alpha x^p),\ x > 0,\ \alpha > 0, \nu > 0, p > 0.$$

Proof. Suppose that $E(X|X \geq x) = \tilde{g}(x)\frac{f(x)}{1 - F(x)}$. Then, since $E(X|X \geq x) = \frac{\int_x^\infty uf(u)du}{1 - F(x)}$, we have $\tilde{g}(x) = \frac{\int_x^\infty uf(u)du}{f(x)}$. Now, if the random variable X Satisfies Assumption 1 and has the PDF (15), then, using Theorem 3 and equation (30), we obtain

$$\tilde{g}(x) = \frac{\int_x^\infty uf(u)du}{f(x)} = \frac{\int_0^\infty uf(u)du - \int_0^x uf(u)du}{f(x)}$$

$$= \frac{(E(X) - g(x)f(x))}{f(x)},$$

Where $f(x)$ denotes the PDF of $X \sim GGGD(\alpha, \nu, p)$ given by (15), $g(x)$ is given by (28) and $E(X)$ It is given by (18). Consequently, the proof of the "if" part of Theorem 4 follows from Lemma 2.
Conversely, suppose that. $\tilde{g}(x) = \frac{(E(X) - g(x)f(x))}{f(x)}$. Now, from Lemma 2, we obtain

$$\tilde{g}(x) = \frac{\int_x^\infty uf(u)du}{f(x)},$$

Or

$$\int_x^\infty uf(u)du = f(x)\,\tilde{g}(x).$$

Differentiating the above equation with respect to x , we obtain

$$-xf(x) = f^{/}(x).\tilde{g}(x) + f(x).\left(\tilde{g}(x)\right)^{/}.$$

Proceeding in the same way as in Theorem 3 and following the similar arguments, we obtain

$$f(x) = cx^{\nu - 1}\exp(-\alpha x^p),$$

Where $c = \frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma(\frac{\nu}{p})}$. This completes the proof of Theorem 4.

4. Parameter Estimation, Simulation Study, and Inference

In this section, we will estimate the parameters of the distribution using four techniques: maximum likelihood estimation (MLE) and the method of moments (MoM). Additionally, we have conducted simulation studies to evaluate the effectiveness of these estimation methods by generating synthetic data and assessing important metrics such as bias and mean square error (MSE).

4.1. Maximizing the likelihood function

The likelihood function is given below:

$$L(\alpha, \upsilon, p) = \left(\frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma(\frac{\upsilon}{p})}\right)^n \prod_{i=1}^n l\,n\,((x_i))^{\nu-1}\, e^{-\alpha\sum_{i=1}^n (x_i)^p}, x_i > \alpha, p, \alpha, \nu > 0. \tag{36}$$

Now, taking the logarithm of the likelihood function in (36) to simplify maximization:

$$\ell\,(\Theta) = l\,n\big(L(\alpha, \upsilon, p)\big) = nl\,n\left(\frac{p\,\alpha^{\frac{\nu}{p}}}{\Gamma(\frac{\upsilon}{p})}\right) + (\nu - 1)\sum_{i=1}^n l\,n\,\big((x_i)\big) - e^{-\alpha\sum_{i=1}^n (x_i)^p}, \tag{37}$$

Where $x_i > \alpha, p, \alpha, \nu > 0$
The derivatives of Equation (37) with respect to p, α, and ν are:

$$\frac{\partial\ell\,(\Theta)}{\partial\,p} = \frac{n}{p} - \frac{n\,\nu\,l\,n\,(\alpha)}{p^2} + \frac{n\,\nu\,\psi\left(0,\frac{\upsilon}{p}\right)}{p^2} - \alpha\sum_{i=1}^n \alpha\sum_{i=1}^n\big(l\,n(x_i)\big)\,(x_i)^p = 0, x_i > \alpha, p, \alpha, \nu > 0, \tag{38}$$

$$\frac{\partial\ell\,(\Theta)}{\partial\,\alpha} = \frac{n\,\nu}{p\,\alpha} - \sum_{i=1}^n (x_i)^p = 0\,, x_i > \alpha, p, \alpha, \nu > 0, \tag{39}$$

$$\frac{\partial \ell(\Theta)}{\partial \alpha} = \frac{n \ln(\alpha)}{p} - \frac{n\,\psi\left(0,\frac{\nu}{p}\right)}{p} + \sum_{i=1}^{n} \ln(x_i) = 0, x_i > \alpha, p, \alpha, \nu > 0. \tag{40}$$

To determine the values of p, α, and ν that maximize the likelihood function, we need to solve equations (38), (39), and (40). Since these equations can be complex and difficult to solve analytically, numerical methods such as the Newton-Raphson algorithm are commonly employed to approximate p. Consequently, the Maximum Likelihood Estimators (MLEs) for p, α, and ν can be obtained either through analytical solutions or iterative numerical optimization techniques.

4.1. Simulation of maximum likelihood estimation (MLE)

MLE aims to identify the parameter p that maximizes the likelihood based on observed data. It is known for its asymptotic efficiency, meaning it performs well with large sample sizes. However, for small samples, MLE may exhibit bias, and its variance estimates can be unstable. To examine this behavior, a simulation study was conducted using three data sets:

- Set-I: p = 0.0234, α = 0.2565, ν = 0.5426
- Set-II: p = 1.3235, α = 1.5435, ν = 1.2425
- Set-III: p= 54.3235, α = 0.5435, ν = 71.2425

The simulation involved:

1) Generating 1000 samples of sizes n = 5, 25, and 50 from the specified distribution.
2) Estimating p via the log-likelihood function.
3) Computing the Bias = $E(\hat{p} - p)$ and Mean Squared Error (MSE) = $E(\hat{p} - p)^2$.
4) Repeating these steps for all values of p.

Results from Table 4.1 reveal that:

- For Set-I, the bias for p, α, and ν remains constant regardless of sample size, indicating that the MLE does not improve with larger samples.
- For Set-II, the bias decreases markedly as n increases, suggesting improved estimation accuracy with more data.
- In Set-III, the bias is substantial for small n but diminishes with larger samples, especially for p and ν, indicating stabilization of the estimates as the sample size grows.

Regarding MSE:

- For Set-I, it remains unchanged across different n, implying limitations of MLE in accurately estimating very small parameters.
- For Set-II, the MSE declines significantly as n increases, demonstrating better estimation with larger samples.
- In Set-III, the high MSE at small n reduces considerably with increased sample size, confirming that MLE performs more reliably with larger datasets.

Table 4.1: Mean Bias and MSE of MLEs Against Various Sample Sizes

Sample Size	Bias($\hat{p}$)	MSE($\hat{p}$)	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Bias($\hat{\nu}$)	MSE($\hat{\nu}$)
Set-I						
n = 25	0.9766	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
n = 50	0.9760	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
n = 100	0.9760	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
n = 200	0.9760	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
Set-II						
n = 25	1129.403	683.858	0.444	2.064×10^{6}	1.053×10^{9}	1.593
n = 50	209.599	0.939	0.278	4.119×10^{6}	1.130×10^{9}	1.074
n = 100	4.826	11.571	0.115	2.219×10^{4}	5.306×10^{4}	0.177
n = 200	0.035	0.994	0.052	1.164×10^{-1}	2.148×10^{2}	0.060
Set-III						
n = 25	380.222	0.029	47.807	7.993×10^{6}	6.188×10^{9}	7,448,514
n = 50	69.438	0.024	41.818	8.912×10^{5}	6.227×10^{8}	6,477,084
n = 100	-14.138	0.020	35.818	8.192×10^{2}	7.692×10^{4}	2,262,882
n = 200	-14.592	0.018	31.407	5.468×10^{2}	5.447×10^{4}	1,678,309

Summary:

- MLE performs poorly for very small parameters (Set-I), leading to consistent bias and MSE across all sample sizes.
- MLE estimates improve with larger sample sizes for moderate and large parameter values (Set-II and Set-III).
- For large parameter values (Set-III), MLE suffers from high bias and MSE at small sample sizes but performs well as *n* increases.
- Overall, MLE is asymptotically unbiased, meaning that its bias and error reduce with larger samples.

4.2. Method of moments (MoM)

The Method of Moments (MoM) is a statistical estimation technique that determines parameters by matching sample moments (e.g., mean, variance) with their theoretical counterparts. It is simpler but less efficient than MLE.

Simulation Study:

Using three data sets:

- Set-I: p = 0.0234, α = 0.2565, ν = 0.5426
- Set-II: p = 1.3235, α = 1.5435, ν = 1.2425
- Set-III: p = 54.3235, α = 0.5435, ν = 71.2425

Results from Table 4.2 indicate:

- Set-I: Strong positive bias for p, consistent negative bias for α, and extreme instability for ν.
- Set-II: Bias for p decreases with larger samples, while α remains stable, but ν is unstable for small samples.

- Set-III: Severe underestimation for p, large fluctuations for α, and extreme instability for ν.

Conclusion: MoM is highly unreliable for GGGD parameter estimation, particularly for small samples and extreme parameter sets, favoring MLE instead.

Table 4.2: Mean Bias and MSE of MoMs Against Various Sample Sizes

			Set-I			
Sample Size	Bias($\hat{p}$)	MSE($\hat{p}$)	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Bias($\hat{\nu}$)	MSE($\hat{\nu}$)
n = 25	0.9766	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
n = 50	0.9760	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
n = 100	0.9760	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
n = 200	0.9760	0.7435	0.4574	0.958×10^{-1}	5.5279×10^{-10}	0.2092
			Set-II			
n = 25	0.884	3.262	NaN	1411.557	1.977×10^{1}	NaN
n = 50	0.071	2.533	-1.113	83.916	$\times\ 10^{01}$	1.282
n = 100	0.502	2.081	-1.091	118.359	6.945	1.235
n = 200	0.012	1.801	-1.082	75.621	4.869	1.208
			Set-III			
n = 25	-53.803	6.0×10^{39}	-7.1×10^{01}	2918.494	3.7×10^{82}	5.0×10^{3}
n = 50	-53.293	2.105×10^{174}	-7.123×10^{1}	3284.494	NaN	5.0×10^{03}
n = 100	-53.557	2.7×10^{81}	-7.1×10^{1}	2965.391	7.5×10^{165}	5.0×10^{3}
n = 200	-53.968	1.3×10^{4}	-7.1×10^{1}	2913.192	1.2×10^{11}	5.0×10^{3}

5. Real data Applications

In this section, we provide the flexibility of the distribution of $X \sim GGGD(\alpha, \nu, p)$ By fitting it to two real data sets and comparing its performance with various well-known distributions. Model comparison is conducted using AIC, BIC, and K-S Statistics, with parameter estimation performed through the MLE method.

5.1. Goodness of fit

In this subsection, various discrimination criteria based on the log-likelihood function are employed, where q represents the number of parameters, $\widehat{\Theta}$ Denotes the MLEs of Θ, n is the sample size, and k is the number of classes. $z_j = F_X(xj), x_j, j = 1,2,3, \ldots$ Are the ordered observations. The Akaike Information Criterion (AIC) by Akaike (1973) optimizes model selection by balancing fit and complexity, while the Bayesian Information Criterion (BIC) by Schwarz (1978) imposes a stricter complexity penalty, favoring simpler models in larger samples. These are defined as follows:

AIC=$2q-2\ell(\widehat{\Theta})$, and BIC=$q\ln(n)-2\ell(\widehat{\Theta})$.

The best model is accepted with the least values of the Kolmogorov-Smirnov (K-S) statistics:

KS=$\max\{\frac{j}{k} - z_j, z_j - \frac{j}{k}\}$.

5.2. Applications

In what follows, the practical applicability of the GGGD is demonstrated using the two real-life datasets, Data Sets I and II, with model adequacy evaluated through multiple goodness-of-fit criteria, including the Anderson-Darling and Cramér–von Mises tests.

Data Set I: The following data is taken from Walpole et al. (2012, pp 13). According to the journal Chemical Engineering, an important property of a fiber is its water absorbency. A random sample of 20 pieces of cotton fiber is taken, and the absorbency of each piece is measured. The absorbency values are given as: 18.71,21.41,20.72,21.81,19.29,22.43,20.17, 23.71,19.44,20.50,18.92,20.33,23.00,22.85,19.25,21.77,22.11,19.77,18.04,21.12.

The following Table 5.1 provides the descriptive summary of the Data Set-I:

Table 5.1: Descriptive Summary (Data Set-I)

Mean	Median	Standard deviation	Quartile deviation	Skewness	Kurtosis	Min	Max
20.77	20.61	1.59	1.24	0.13	-1.00	18.04	23.71

Goodness-of-Fit Measures: The following Table 5.2 presents the fit and comparison of several well-known probability distributions based on the dataset (Data Set-I), using various model selection and goodness-of-fit criteria.

Table 5.2: Goodness-of-Fit Measures (Data Set-I) Based on Listed MLEs

Distribution	Parameters	LL	AIC	BIC	KS	AD	CvM
GGGD	P=1.66, α=0.425, ν=108.674	-37.12	80.23	83.22	0.10	-20.21	0.03
Pareto	[7.246, 18.04]	-41.00	86.01	88.00	0.18	-22.51	0.30
Dagum	[10., 26.649, 10.]	5414.69	-10823.38	-10820.39	1.00	-404.27	6.67
Lindley	[0.092]	-73.85	149.71	150.70	0.47	-26.72	1.43
Weibull	[14.52, 21.496]	-38.59	81.18	83.17	0.11	-20.29	0.04
Half-Normal	[20.825]	-75.24	152.48	153.47	0.56	-28.75	1.93
Lévy	[17.807, 1.612]	-50.00	104.00	105.99	0.40	-23.14	0.62
Rayleigh	[14.726]	-66.97	135.94	136.94	0.48	-26.67	1.43
Maxwell	[12.024]	-62.50	127.01	128.00	0.43	-25.70	1.19

It is observed from the above Table 5.2 that the GGGD (generalized generalized gamma distribution) emerges as the best-fitting model with a Log-Likelihood (LL) of -37.12, and the lowest AIC (80.23) and BIC (83.22) values among all realistic models. Additionally, it shows excellent goodness-of-fit with Kolmogorov-Smirnov (KS) = 0.10, Anderson-Darling (AD) = -20.21, and Cramer-von Mises (CvM) = 0.03, indicating a very close alignment with the observed data.The Weibull distribution also performs reasonably well, with LL = -38.59, AIC = 81.18, and BIC = 83.17, and slightly higher KS (0.11) and CvM (0.04) values than Gen Gamma, making it a second-best option. The Pareto distribution follows next but shows a higher KS = 0.18, indicating a weaker empirical fit despite acceptable AIC/BIC values.In contrast, distributions like Lindley, Half-Normal, Rayleigh, Lévy, and Maxwell show poorer performance with much higher AIC/BIC, and KS values ranging from 0.40 to 0.56, signifying significant deviation from the actual data pattern.Although the Dagum distribution shows extremely high Log-Likelihood (5414.69) and negative AIC/BIC values, its KS = 1.00 and very high CvM = 6.67 imply a completely unrealistic or misfitted model, likely due to numerical overfitting or data-scale mismatch.

Visual Fit: The following plot (Figure 7) shows the density comparison of the above-said probability distributions fitted in Table 4 to the dataset (Data Set-I), represented by the light blue histogram. The goal is to visually assess which distribution models the observed data most accurately.

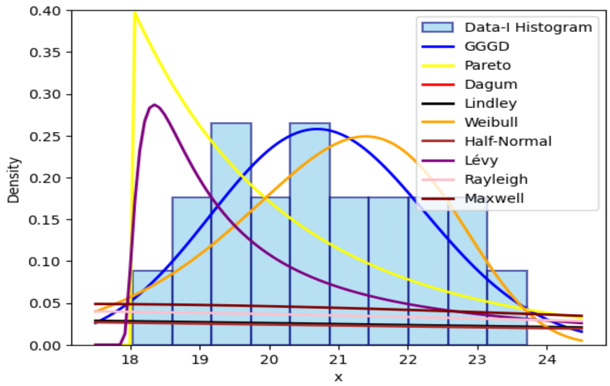

Fig. 7: Histogram of Data Set-I.

From Figure 7, it is obvious that the GGGD (Generalized Gamma-Gamma Distribution), shown in blue, closely follows the shape of the histogram, especially around the peak and both tails, suggesting it fits the data quite well. The Weibull distribution (orange) also provides a reasonably good fit, capturing the peak and general spread of the data, though it slightly underestimates the left tail. The Pareto distribution (yellow) starts very steeply and quickly diverges, indicating a poor fit for the middle and right parts of the data. The Lévy distribution (purple) shows a sharp initial peak but does not align well with the overall shape of the histogram, especially in the right tail. The Lindley (black), Half-Normal (brown), Rayleigh (pink), and Maxwell (dark brown) distributions appear to provide very flat density lines, indicating that they are not capturing the variability and structure of this dataset properly. Their curves remain almost constant and do not mimic the histogram's distribution. Overall, GGGD and Weibull emerge as the most suitable models for this dataset, effectively capturing both the central tendency and dispersion of the data, whereas the other distributions either oversimplify or poorly fit the observed pattern.

Data Set-II: The following data of McGilchrist and Aisbett (1991) is taken from Almetwally et al. (2020) for 30 patients, showing the first recurrence times to infection at point of insertion of the catheter for kidney patients using portable dialysis equipment:

8, 23, 22, 447, 30, 24, 7, 511, 53, 15, 7, 141, 96, 149, 536, 17, 185, 292, 22, 15, 152, 402, 13, 39, 12, 113, 132, 34, 2, 130.

The descriptive summary of the Data Set-II is provided in the following Table 5.3:

Table 5.3: Descriptive Summary (Data Set-II)

Mean	Median	Standard deviation	Quartile deviation	Skewness	Kurtosis	Min	Max
117.9	36.5	159.86	67	1.53	3.52	2	536

Goodness-of-Fit Measures: The following Table 5.4 presents a comparative analysis of several probability distributions fitted to the given Data Set-II, evaluated using various statistical metrics: Log-Likelihood, AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), KS (Kolmogorov-Smirnov statistic), AD (Anderson-Darling statistic), and CvM (Cramer-von Mises statistic).

Table 5.4: Goodness-of-Fit Measures (Data Set-II) Based on Listed MLEs

Distribution	Parameters	Log-Likelihood	AIC	BIC	KS	AD	CvM
GGGD	p; 0.073 α=65.386 ν=6.426	-170.43	346.87	351.07	0.12	-30.53	0.10
Pareto	[0.313 2.]	-181.38	366.75	369.55	0.26	-34.40	0.75
Dagum	[10. 50. 10.]	8436.39	-16867	-16863	0.55	-250.74	2.61
Lindley	[0.016]	-188.60	379.19	380.59	0.39	-41.25	1.16
Weibull	[0.751 100.582]	-171.99	346.98	351.08	0.15	-30.72	0.12
Half-Normal	[196.675]	-180.22	362.44	363.84	0.38	-38.04	1.22
Lévy	[-0.997 21.919]	-173.31	350.61	353.42	0.16	-30.77	0.10
Rayleigh	[139.073]	-209.54	421.09	422.49	0.50	-57.51	2.43
Maxwell	[113.553]	-244.57	491.13	492.53	0.54	-77.30	3.00

Among all distributions, the GGGD stands out with a high Log-Likelihood (-170.43) and lowest AIC (346.87) and BIC (351.07) among the remaining models, indicating the best balance between fit and model complexity. It also exhibits the lowest KS (0.12), AD (-30.53), and CvM (0.10) values, confirming an excellent goodness-of-fit across multiple criteria. The Weibull and Lévy distributions also perform reasonably well with relatively low KS, AD, and CvM values, but their AIC and BIC values are slightly higher than those of GGGD, making them less optimal choices. In contrast, the Dagum distribution also shows an extremely high Log-Likelihood (8436.39) and negative AIC and BIC values, suggesting a mathematically strong fit; however, its KS (0.55), AD (-250.74), and CvM (2.61) values indicate poor alignment with the empirical distribution, likely due to overfitting or numerical issues in parameter estimation. Furthermore, the Weibull and Lévy distributions also perform reasonably well with relatively low KS, AD, and CvM values, but their AIC and BIC values are slightly higher than those of GGGD, making them less optimal choices. Moreover, distributions like Pareto, Lindley, Half-Normal,

Rayleigh, and Maxwell exhibit higher KS, AD, and CvM values, signaling weaker fits and greater deviation from the observed data. Overall, the Table 6 clearly supports that GGGD is the best fitting distribution for the dataset, both in terms of model selection criteria (AIC, BIC) and goodness-of-fit statistics (KS, AD, CvM).
Visual Fit: The best fitting of GGGD to the Data Set-II also aligns with the visual evidence from the density plot as provided in the following Figure 8:

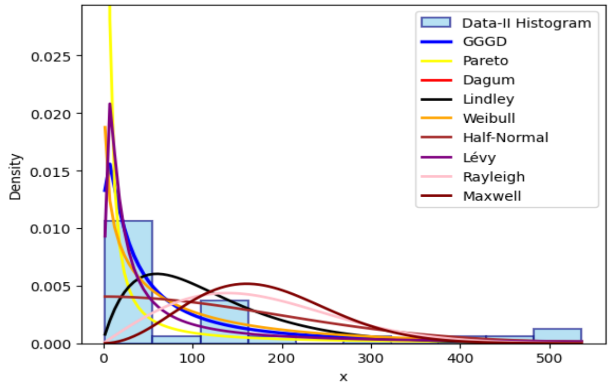

Fig. 8: Histogram of Data Set-II.

The provided plot in Figure 8 displays the comparison of several probability distributions fitted to a dataset, represented by the blue histogram labeled "Data-II Histogram". The purpose of the plot is to visually assess how well different distributions model the underlying data. The GGGD (Generalized Gamma-Gamma Distribution), shown in blue, aligns closely with the shape of the histogram in the lower range, suggesting it captures the concentration of values well near the left tail. The Pareto distribution (yellow) fits the initial steep decay but rapidly diverges, making it less suitable for the full range of data. The Dagum distribution (red) shows a smoother fit but starts diverging as x increases. The Lindley distribution (black) captures some of the peak and tail behavior but underestimates the central region. The Weibull distribution (orange) fails to model the peak and appears too flat across the range. The Half-Normal (brown), Lévy (purple), Rayleigh (pink), and Maxwell (dark brown) distributions generally show poorer fits, especially in the right tail, where the actual data shows some higher values extending out to $x > 400$, which these distributions fail to capture well. Overall, the GGGD appears to provide the best fit among the tested models, particularly in capturing the skewness and heavy tail observed in the data, followed by Lindley and Dagum to some extent. The comparison indicates that simple symmetric or light-tailed distributions (Weibull, Rayleigh, Maxwell, etc.) are not suitable for this dataset, which likely represents a positively skewed, heavy-tailed phenomenon.

6. Conclusion

This study proposed a differential equation-based construction of the Generalized Generalized Gamma Distribution (GGGD), demonstrating its theoretical foundation and statistical flexibility. The distribution's key properties and performance in modeling real-world data were validated through simulation and goodness-of-fit testing. While currently limited to the univariate case and classical estimation techniques, future extensions could include multivariate forms, Bayesian methods, and applications in reliability, survival analysis, and risk modeling. The GGGD thus offers promising avenues for both theoretical development and applied research in statistical distribution theory.

Declarations: Authors' Contributions

All the authors have equally made their contributions, and have read and approved the final manuscript.

Funding

The authors state that they have no funding source for this paper. Availability of Data and Materials: Not applicable.

Declarations Conflict of Interest

The authors declare that they have no competing interests. The authors state that no funding source or sponsor has participated in the realization of this work.

References

[1] M. Ahsanullah, Characterizations of Univariate Continuous Distributions, Atlantis-Press, Paris, France, 2017. https://doi.org/10.2991/978-94-6239-139-0
[2] H. Akaike, Information Theory as an Extension of the Maximum Likelihood Principle. In: Petrov, B.N., and Csaki, F., Eds., Second International Symposium on Information Theory, Akademiai Kiado, Budapest, 267-281 (1973).
[3] E. M. Almetwally, H. Z. Muhammed, and E. S. A. El-Sherpieny, Bivariate Weibull distribution: properties and different methods of estimation. Annals of Data Science, 7(1), 163-193 (2020). https://doi.org/10.1007/s40745-019-00197-5
[4] J. Galambos, and S. Kotz, Characterizations of probability distributions: A unified approach with an emphasis on exponential and related 18 models, Lecture Notes in Mathematics, 675, Springer, Berlin, Germany, 1978.
[5] W. Glanzel, A. Telcs, and A. Schubert, Characterization by truncated moments and its application to Pearson-type distributions, Z. Wahrsch. Verw. Gebiete, 66, 173 – 183 (1984). https://doi.org/10.1007/BF00531527
[6] S. Gradshteyn, and I. M. Ryzhik, Table of integrals, series, and products, Academic Press, Inc., San Diego, California, USA, 2000.

[7] E.T. Jaynes, Information theory and statistical mechanics. Phys. Rev. 106(4), 620–630 (1957). https://doi.org/10.1103/PhysRev.106.620
[8] E. Koudou, and C. Ley, Characterizations of GIG laws: A survey, Probability surveys, 11: 161 – 176 (2014). https://doi.org/10.1214/13-PS227.
[9] P. Langevin, Sur la théorie du mouvement brownien [On the Theory of Brownian Motion]. C. R. Acad. Sci. Paris. 146: 530–533 (1908).
[10] C. A. McGilchrist and C. W. Aisbett, Regression with frailty in survival analysis. Biometrics 47: 461–466 (1991). https://doi.org/10.2307/2532138.
[11] H. Nagaraja, Characterizations of Probability Distributions. In Springer Handbook of Engineering Statistics (pp. 79 - 95), Springer, London, UK, 2006. https://doi.org/10.1007/978-1-84628-288-1_4.
[12] K. Pearson, Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material, Philo- sophical Transactions of the Royal Society of London, A186 (1895), 343-414. https://doi.org/10.1098/rsta.1895.0010.
[13] K. Pearson, Mathematical contributions to the theory of evolution, X: Supplement to a memoir on skew of variation, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 197 (1901), 343-414. https://doi.org/10.1098/rsta.1901.0023.
[14] K. Pearson, Mathematical contributions to the theory of evolution, XIX: Second supplement to a memoir on skew of variation, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 216 (1916), 429-457. https://doi.org/10.1098/rsta.1916.0009.
[15] V. D. Rao, S. V. S. Girija, and Y. Phani, Differential Approach to Cardioid Distribution, Computer Engineering and Intelligent Systems, Vol 2, No.8 (2011), 1-6.
[16] G. Schwarz, Estimating the dimension of a model. The annals of statistics, 6 (2), 461 – 464 (1978). https://doi.org/10.1214/aos/1176344136.
[17] M. Shakil, M. Ahsanullah, and B. M. G. Kibria, On the Characterizations of Chen's Two-Parameter Exponential Power Life-Testing Distribution, Journal of Statistical Theory and Applications, 17(3) (2018), 393-407. https://doi.org/10.2991/jsta.2018.17.3.1.
[18] M. Shakil, M. Ahsanullah, and B. M. G. Kibria, Some characterizations and applications of a size-biased weighted distribution useful in lifetime modelling, Journal of Statistics Applications & Probability, 10(3) (2021), 607-624. https://doi.org/10.18576/jsap/100301.
[19] M. Shakil, J. N. Singh, L. Tomy, R. C. S. Chandel, T. Hussain, A. Khadim, M. Ahsanullah, and B. M. G. Kibria. A First Order Linear Differential Equation Approach of Generating Continuous Probability Distributions Review, Analysis and Characterizations. Jnanabha, Vol. 54(2) (2024), 233-250. https://doi.org/10.58250/jnanabha.2024.54223
[20] C. E. Shannon, A mathematical theory of communication. Bell System Tech. J., Vol. 27 (1948), 379 – 423; 623 – 656. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.
[21] V. P. Singh, Systems of frequency distributions for water and environmental engineering, Physica A: Statistical Mechanics and its Applications, 506 (2018), 50-74. https://doi.org/10.1016/j.physa.2018.03.038.
[22] V. P. Singh, and L. Zhang, Pearson System of Frequency Distributions, Systems of Frequency Distributions for Water and Environmental Engineering, Chapter 2 (2020), 11 – 39, Cambridge University Press, USA. https://doi.org/10.1017/9781108859530.003.
[23] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, Probability & Statistics for Engineers & Scientists, 9th Edition, Page 13, Prentice Hall, New York, USA (2012).