International Journal of Advanced Statistics and Probability, 12 (1) (2025) 1-11

International Journal of Advanced Statistics and Probability

Website: www.sciencepubco.com/index.php/IJASP
https://doi.org/10.14419/gxf6wa90
Research paper

A new log dagum singh maddala TX family of distribution

Aneeqa Khadim [1] *, Aamir Saghir [1], Tassadaq Hussain [1], M.Shakil [2]

[1] *Department of Mathematics Mirpur University of Science and Technology (MUST) Mirpur-10250 AJK, Pakistan*
[2] *Department of Mathematics Miami Dade College Hialeah FL33024, USA*
**Corresponding author E-mail: aneeqa89@gmail.com*

Abstract

This study introduces the Log-Dagum Singh Maddala (4P) distribution, a significant contribution to the field of continuous distributions. We investigate its statistical properties including probability density, hazard rate, survival functions, quantiles, and order statistics. The distribution is thoroughly characterized using multiple approaches and maximum likelihood estimation is used to determine the parameters. The model's effectiveness is evaluated using three real-world datasets and a comparative analysis with established distributions highlights its advantages. The results confirm the excellence of the proposed model demonstrating its practical significance in real data analysis.

Keywords: *Log-Dagum Distribution; Probability Distributions; Characterization; Simulation Study; Maximum Likelihood; Parameter Estimation; Singh Maddala Distribution.*

1. Introduction

Statistical distributions have been generalized to enhance their flexibility by introducing additional shape parameters. This allows for a better fit to real-world data which often exhibits complexities that traditional distributions cannot capture. These generalizations enable researchers to model various phenomena more accurately such as skewness and heavy-tailedness. The generalized beta distribution, proposed by Singh and Maddala (1976) is a notable example. The Dagum distribution, introduced by Camilo Dagum in (1977) is a continuous probability distribution used to model non-negative random variables. By applying a logarithmic transformation to this distribution we can derive the key characteristics of the resulting log-Dagum random variable offering benefits like variance stabilization and skewness normalization.

The Dagum distribution has attracted significant research interest due to its competitive performance in modeling various phenomena. To enhance its flexibility several extensions have been proposed, including the log-Dagum, Mc-Dagum, beta-Dagum and extended Dagum distributions. These models offer improved structural properties and parameter estimates making them valuable tools for data analysis. Additionally, weighted and Poisson variants of the Dagum distribution have been introduced further expanding its applicability. Recent advancements include the exponentiated generalized exponential Dagum distribution, power log-Dagum distribution and the odd Dagum-G family which provide even more versatility. The Dagum distribution's hazard rate function is a crucial aspect exhibiting diverse shapes that make it suitable for various fields. The Log-Dagum distribution introduced by Domma (2004) has been extensively studied for its kurtosis properties by Polisicchio and Zenga (1997). Further research by Afify and Alizadeh (2020) explored its properties and parameter estimation methods. The Log-Dagum distribution has since been widely applied in various fields including economics, web traffic analysis, insurance, seismology, finance and telecommunications showcasing its versatility. Recent advancements in distribution development involve adding parameters to baseline distributions generating new families of distributions that have been successfully employed in modeling diverse data sets across multiple areas. Some of the familiar generators are beta-G are studied in Eugene et al., (2002). Kumaraswamy-G are discussed Cordeiro and de Castro (2011). T-normal are studied in Alzaatreh et al., (2014a) gamma-G (type II) Risti´c and Balakrishnan (2012) transformed-transformer {T-X} family are discussed in Alzaatreh et al., (2013) new Weibull-G are studied in Tahir et al., (2016) some new members of the {T-X} family of distributions are studied in Jamal and Nasir (2016). Exponentiated {T-X} family is examined in Alzaghal et al. (2013) T-X{Y} family (a quantile based approach) Aljarrah et al., (2014). A modified {T-X} family of distributions is discussed in Aslam et al., (2020). Applications and properties of a new member of the {T-X} family of distributions are discussed Handique et al., (2021).

Recent research has explored the Log Dagum Weibull distribution, with Khadim et al., (2021) examining its properties and characterization. Additionally, Shakil et al., (2021) investigated the properties of the Burr (4P) distribution, while Shakil et al., (2021) conducted inference studies on the Dagum (4P) distribution. A new (T-X^{θ})family of distributions properties, discretization and estimation with applications is discussed by Mandouh et al., (2024).

A new distribution, the Log-Dagum Singh Maddala distribution is proposed by combining the Singh Maddala distribution and the Log-Dagum distribution. This compounding model offers a more flexible and robust framework for data analysis leveraging the strengths of both distributions. Unlike a straightforward extension this composition creates a new distribution that harnesses the benefits of both models providing a valuable tool for complex data sets. The Transformed-Transformer (T-X) method introduced by Cordeiro et al., (2013) has

Copyright © Aneeqa Khadim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

emerged as a key technique for generalizing distributions widely adopted by researchers. This approach has led to the creation of diverse continuous distribution families significantly expanding statistical modeling capabilities.
Consider a probability density function $s(t)$ defined on $[a, b]$. Let $W[G(x)]$ be a function of the cumulative distribution function. The {T-X} family introduced by Alzaatreh et al., (2013) has a cumulative distribution function of the form:

$$F(x) = \int_0^{W[G(x)]} s(t)dt, \tag{1}$$

Where $W[G(x)]$ satisfies the following conditions

i) $W[G(x)] \in [u; v]$;
ii) $W[G(x)]$ Is differentiable and monotonically non-decreasing, and
iii) $W[G(x)] \to u$ as $x \to -\infty$ and $W[G(x)] \in v$ as $x \to \infty$

The pdf can be obtained by differentiating equation (1).

$$f(x) = \left\{\frac{d}{dx} W[G(x)]\right\} s\{W[G(x)]\}.$$

2. The log dagum singh maddala distribution

The main purpose of this article is to introduce a new family of distributions called the log Dagum Singh Maddala family of distributions that are more adaptable to data in a wide range of applications.
The log Dagum distribution has random variable T its cumulative and probability density function is define as

$$\pi(t) = (1 + e^{-\lambda x})^{-\beta}, t \in R, \beta > 0, \lambda > 0, \tag{2}$$

$$s(t) = \beta\lambda\, e^{-\lambda x}(1 + e^{-\lambda x})^{-\beta-1}, t \in R, \beta > 0, \lambda > 0, \tag{3}$$

Let $G(x)$ be the baseline cdf by replacing $W[G(x)]$ by $\log(\frac{G(x)}{1-G(x)})$ and $s(t)$ with (3) in (1).
The cdf and pdf of the log Dagum-x family is identifying as

$$F(x) = \left[1 + \left(\frac{G(x)}{1-G(x)}\right)^{-\lambda}\right]^{-\beta}, \tag{4}$$

And

$$f(x) = \left[1 + \left(\frac{G(x)}{1-G(x)}\right)^{-\lambda}\right]^{-\beta-1} \left(\frac{G(x)}{1-G(x)}\right)^{-\lambda-1} \frac{\lambda\,\beta\, g(x)}{[1-G(x)]^2}, \tag{5}$$

The cumulative and probability density function are obtained by using (4) and (5) as

$$F(x) = \left[1 + \left[(1 + x^c)^k - 1\right]^{-\lambda}\right]^{-\beta}, \tag{6}$$

And

$$f(x) = \left[1 + \left[(1 + x^c)^k - 1\right]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1 + x^c)^{-1+k}(-1 + (1 + x^c)^k)^{-1-\lambda}\beta\lambda ck. \tag{7}$$

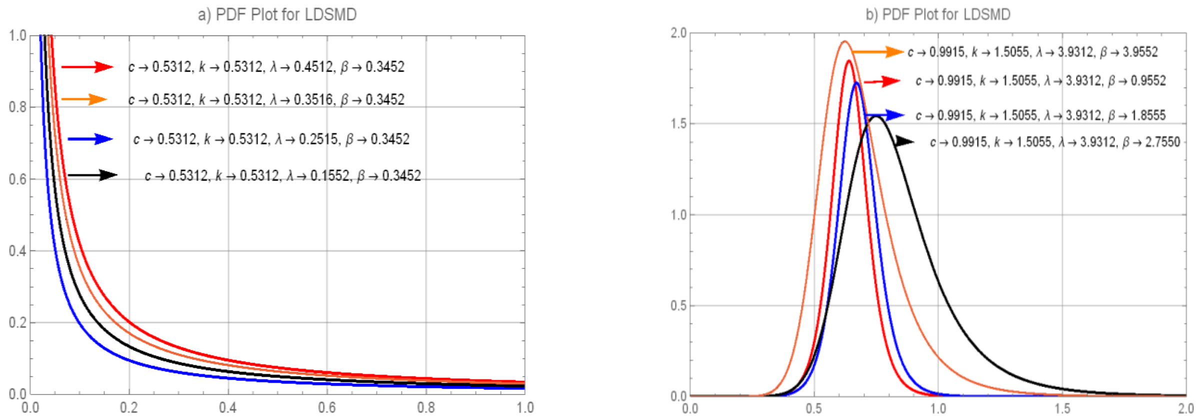

Fig. 1: Density Plot of LDSMD.

The density functions of the Log-Dagum Singh Maddala (LDSM) distribution are illustrated in figure 1 which reveals that the density function decreases as λ increases while holdingβ, C and K constant. However, when β is varied while keeping λ fixed the shape of the curve changes with the density function exhibiting a right-skewed nature as C and K decrease.
Sub-models

1) When $c = -\delta, k = -\alpha, \lambda = -\varphi, \beta = -\theta$ we get the cdf of exponentiated

Kumaraswamy III

$$F(x) = \left[1 - \left[1 - \left(1 + x^{-\delta}\right)^{-\alpha}\right]^{\varphi}\right]^{\theta} \text{ for } \alpha, \delta > 0 \text{ and } x > 0,$$

2) When $c = -\delta, k = -\alpha, \lambda = -\varphi, \beta = -1$ we get the cdf of Kumaraswamy-Burr III

$$F(x) = 1 - \left[1 - \left(1 + x^{-\delta}\right)^{-\alpha}\right]^{\varphi} \text{ for } \alpha, \delta > 0 \text{ and } x > 0,$$

3) When $c = -\delta, k = -\alpha, \lambda = \beta = -1$ we get the cdf of Burr III distribution.

$$F(x) = \left(1 + x^{-\delta}\right)^{-\alpha} \text{ for } \alpha, \delta > 0 \text{ and } x > 0.$$

3. Mathematical properties of the LDSM distribution

The properties of the LDSM distribution are observed and considered. These properties are essential when the distribution is applying to real life data.

3.1. Survival function

The survival function associated with LDSMD is define as

$$S(x) = 1 - G(x),$$

$$S(x) = 1 - \left[1 + \left[(1 + x^{c})^{k} - 1\right]^{-\lambda}\right]^{-\beta},$$

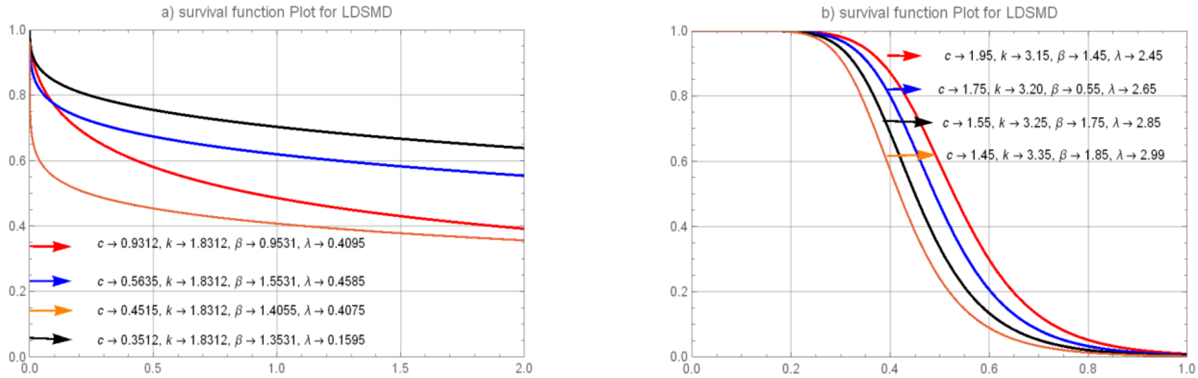

Fig. 2: Survival Plots of LDSM.

For different values of parameter the graph of survival function increases then abruptly starts steadily decreases and converges to zero.

3.2. Hazard function

The hazard function of LDSMD is specified by

$$h(x) = \frac{\left[1+\left[(1+x^{c})^{k}-1\right]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1+x^{c})^{-1+k}\left(-1+(1+x^{c})^{k}\right)^{-1-\lambda}\beta\lambda c k}{1-\left[1+\left[(1+x^{c})^{k}-1\right]^{-\lambda}\right]^{-\beta}},$$

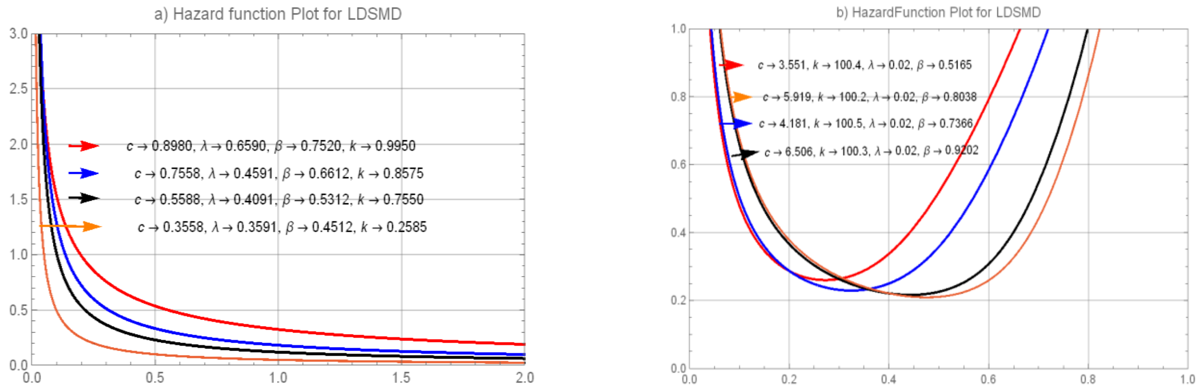

Fig. 3: Hazard Plot of LDSMD.

Figure 3 shows that for different values of parameters the hazard function is increasing, decreasing and bathtub shapes.

3.3. Nature of hazard function

We can describe analytically the shape of the density function as:

$$\frac{d}{dx}\left[\frac{\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\beta\lambda ck}{1-\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta}}\right]=0. \quad (12)$$

There might be more than one root

3.4. Concavity

The concavity of the (hazard rate function) is defined as $h''(x)=0$

$$\frac{d^2}{dx^2}\left[\frac{\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\beta\lambda ck}{1-\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta}}\right]=0$$

For the different values of parameter hazard function is concave down and concave up where the concavity is change that point is called point of inflection.

4. Quantile points

The quantile points of LDW distribution are computed by numerically for different sets of value of the parameters as provided in table below by solving the equation

$$x=Q=\left[\left(\left(u^{-\frac{1}{\beta}}-1\right)^{-\lambda}+1\right)^{\frac{1}{k}}-1\right]^{1/c},$$

Table 1: Quantile Points of the LDSM Distribution

Parameters		0.75	0.80	0.85	0.90	0.95	0.99
$C=1, k=1, \beta=0.2, \lambda=0.5$	Q_u	0.24428	0.34683	0.49511	0.731669	1.22153	3.07543
$C=2, k=0.2, \beta=0.2, \lambda=0.1$	Q_u	0.88619	1.45393	2.82083	8.47264	96.1237	6678.9
$C=2, k=0.1, \beta=0.2, \lambda=0.2$	Q_u	18.4398	22.6082	28.5979	38.5418	61.4492	175.066
$C=2, k=2, \beta=0.1, \lambda=2$	Q_u	0.042177	0.084905	0.172068	0.366441	0.892319	2.91736
$C=1, k=1, \beta=0.1, \lambda=1$	Q_u	0.059674	0.12029	0.245135	0.53534	1.49213	9.45829
$C=3, k=3, \beta=0.5, \lambda=2$	Q_u	0.727075	0.847329	0.993758	1.18797	1.50746	2.31696

5. Characterization

In this section for various method of characterization of probability distribution we present some characterization results of the LDSM (4P) distribution. We refer to Ahsanullah (2013a) and references therein

5.1. Characterization via truncated moments

Characterize the Log Dagum Singh Mandela (4P) distribution in following theorems.
Preposition 1.
Suppose that Y is absolutely continuous random variable and has cdf $F(y)$ with

$F(0)=0, F(y)>0, \forall y>0$, p.d.f $f(y)=F'(y)$,

Then

$E[Y|Y\geq y]=u(y)h(y), y>0,$

And where

$$h(y)=\frac{f(y)}{1-F(y)},$$

$$u(y)=\frac{y\left(1-\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta}+\int_0^y\left(1-\left[1+\left[(1+t^c)^k-1\right]^{-\lambda}\right]^{-\beta}dt\right.\right.}{\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\beta\lambda ck}.$$

Proof given in appendix

Proposition 2.

Suppose the random variable Y is absolutely continuous and has cdf $F(y)$ with

$F(0) = 0, F(y) > 0, \forall\, y > 0$, p.d.f $f(y) = F'(y)$,

Then

$E[Y|Y \geq y] = s(y)h(y), y > 0,$

Where, $h(y) = \frac{f(y)}{1-F(y)}$,

$$s(y) = \frac{y\left(1-\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta} + \int_y^\infty \left(1-\left[1+\left[(1+t^c)^k-1\right]^{-\lambda}\right]^{-\beta} dt\right.\right.}{\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}$$

Proof given in appendix

5.2. Characterization via order statistics

Proposition 3: If $\tau_1, \tau_2, \ldots, \tau_n$ be the n independent copies of the random variable τ with $F(\tau)$ absolutely continuous distribution function and $f(\tau)$ pdf and if the corresponding order statistics $\tau_{1,n} \leq \tau_{2,n} \leq \cdots \leq \tau_{n,n}$, then it is identified from Arnold et al., (2005) chapter 2, or Ahsanullah et al., (2013a), chapter 5, that $\tau_{j,n}/\tau_{k,n} = \tau$, for $1 \leq k < j \leq n$, is distributed as the $(j-k)th$ order statistics from random variable V from the independent observations having the pdf $f_V(v/\tau)$ where $f_V\left(\frac{v}{\tau}\right) = \frac{f(v)}{1-F(\tau)}$, $0 \leq v < x$, and $\tau_{i,n}/\tau_{k,n} = \tau$, for $1 \leq k < i \leq n$, is distributed as ith order statistics from k independent observations from the random variable W having the pdf $f_W(w/\tau)$ where $f_W(w/\tau) = \frac{f(w)}{F(\tau)}$, $w < \tau$. $f_W(w/\tau) = \frac{f(w)}{F(\tau)}$, $w < \tau$.

Let $S_{k-1} = \frac{1}{k-1}(\tau_{1,n} + \tau_{2,n} + \cdots + \tau_{k-1,n})$ and $T_{k,n} = \frac{1}{n-k}(\tau_{k+1,n} + \tau_{k+2,n} + \cdots + \tau_{n,n})$.

Theorem: *Let τ be the absolutely continuous the random variable with the cumulative and probability density function $F(\tau)$ and $f(\tau)$. we suppose that $\omega = inf\{\tau/F(\tau) > 0\}$ and $\delta = sup\{\tau/F(\tau) < 1\}$. We also suppose that $E(\tau)$ exists $f(\tau)$ is a differentiable for all τ. Taking $\omega = 0$ and $\delta = \infty$, we have $E(S_{k-1}/\tau_{k,n} = \tau) = h(\tau)m(\tau)$, where $h(\tau)$ and $h(\tau)$ are respectively given by the expressions in Proposition 1, if and only if X has the pdf (9).*

Proof: It is known that $E(S_{k-1}/\tau_{k,n} = \tau) = E(\tau/\tau \leq \tau)$. See David and Nagaraja (2003) and Ahsanullah et al. (1995). Hence, by proposition 1, the result follows.

Proposition 4. Suppose absolutely continuous random variable τ with the CDF $F(\tau)$ and the PDF $f(\tau)$ suppose that $\omega = 0$ and $\delta = \infty$. We also suppose that $E(\tau)$ exists and $f(\tau)$ is a differentiable for all τ. Then $E(T_{k,n}/\tau_{k,n} = \tau) = s(\tau)h(\tau)$, where $s(\tau)$ and $h(\tau)$ are respectively given by the expressions in Proposition 2, if and only if τ has the pdf (9).

Proof: Since $E(T_{k,n}/\tau_{k,n} = \tau) = E(\tau/\tau \geq \tau)$, see David and Nagaraja (2003) and Ahsanullah et al. (1995) the result follows from Proposition 2.

5.3. Characterization via upper record values

Proposition 5: Consider a sequence of absolutely continuous random variable of independent and identically distributed distribution function $F(\tau)$ and $f(\tau)$.
If $\tau_n = (\tau_1, \tau_2, \ldots \tau_n)$ for $n \geq 1$ and $\tau >, j > 1$, then τ_j is termed an upper record value of $\{\tau_n, n \geq 1\}$. The indices at which the upper records occur are given by the record times $\{U(n) > \min (j \backslash j > U(n+1), \tau_j > \tau_{U(n-1)}, n > 1)$ and $U(1) = 1$. Let $\tau(n) = \tau_{U(n)}$ is denoted by nth upper record value.

Theorem: *Now, we assume that the random variable Y is absolutely continuous with the CDF $F(\tau)$ and the PDF $f(\tau)$. We suppose that $\omega = 0$ and $\delta = \infty$. also we suppose that $E(\tau)$ exists and $f(\tau)$ is differentiable for all τ. Then $E(\tau(n+1) \backslash \tau(n) = \tau) = s(\tau)h(\tau)$, where $s(\tau)$ and $h(\tau)$ are respectively given by the expressions in Proposition 2, if and only if τ has the pdf (9).*

Proof: $E(\tau(n+1) \backslash \tau(n) = \tau) = E(\tau \backslash \tau \geq \tau)$ and it is recognized from Nevzorov (2001) and Ahsanullah et al., (1995).Then, from proposition 2 the result follows.

6. Order statistic

The probability density function of the jth order statistic is given by

$$f_{j,n}(x) = \frac{n!}{(j-1)(n-j)!}[F(x)]^{j-1}[1-F(x)]^{n-j}f(x), \quad (17)$$

The pdf of the jth order statistics for size n random sample from the LDSM distribution is, yet, specified as follows

$$f_{j,n}(x)=\frac{n!}{(n-j)(j-1)!}\left[\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta}\right]^{j-1}\times$$

$$\left[1-\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta}\right]^{n-j}\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta-1}\times$$

$$(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}x^{-1+c}\beta\lambda ck\,,$$

So, by substituting $j=1$ we obtained the pdf of minimum order statistics

$$f_{1,n}(x)=\frac{n!}{(n-1)!}\left[1-\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta}\right]^{n-1}\times$$

$$\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck,$$

Pdf of maximum order statistics is obtained by substituting $j=n$ in above equation

$$f_{n,n}(x)=\frac{n!}{(j-1)!}\left[\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta}\right]^{n-1}\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta-1}$$

$$\times\ x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck.$$

7. Application study

7.1. Maximum likelihood estimation

The maximum likelihood estimation method is widely used for parameter estimation. The total log-likelihood function for the LDSM distribution is

$$L(c,k,\lambda,\beta)=\beta^n\prod_{i=1}^n\left(1+((1+x^c)^k-1)^{-\lambda}\right)^{-\beta-1}\lambda^n\times$$

$$\prod_{i=1}^n((1+x^c)^k-1)^{-\lambda-1}k^n\prod_{i=1}^n(1+x^c)^{k-1}c^n\prod_{i=1}^n x^{c-1},$$

$$L(c,k,\lambda,\beta)=n\ln[c]+n\ln[k]+n\ln[\lambda]+n\ln[\beta]-(\beta+1)\times$$

$$\sum_{i=1}^n\ln\left[1+\left((1+x_i{}^c)^k-1\right)^{-\lambda}\right]-(\lambda+1)\sum_{i=1}^n\ln\left[(1+x_i{}^c)^k-1\right]$$

$$+\ (k-1)\sum_{i=1}^n\ln[1+x_i{}^c]+(c-1)\,,$$

Taking first derivative of the log-likelihood function with respect to parameters given as follow

$$\frac{\delta L}{\delta c}=\frac{n}{c}+\sum_{i=1}^n\mathrm{Log}[x_i]+(-1+k)\sum_{i=1}^n\frac{\mathrm{Log}[x_i]x_i^c}{1+x_i^c}-(1+\lambda)\sum_{i=1}^n\frac{k\mathrm{Log}[x_i]x_i^c(1+x_i^c)^{-1+k}}{-1+\left(1+x_i^c\right)^k}-(1+$$

$$\beta)\sum_{i=1}^n-\frac{k\lambda\mathrm{Log}[x_i]x_i^c(1+x_i^c)^{-1+k}\left(-1+(1+x_i^c)^k\right)^{-1-\lambda}}{1+\left(-1+\left(1+x_i^c\right)^k\right)^{-\lambda}}$$

$$\frac{\delta L}{\delta k}=\frac{n}{k}+\sum_{i=1}^n\mathrm{Log}[1+x_i^c]-(+\lambda)\sum_{i=1}^n\frac{\mathrm{Log}[1+x_i^c]\left(1+x_j^c\right)^k}{-1+\left(1+x_i^c\right)^k}-$$

$$(1+\beta)\sum_{i=1}^n-\frac{\lambda\mathrm{Log}[1+x_i^c](1+x_i^c)^k\left(-1+(1+x_i^c)^k\right)^{-1-\lambda}}{1+\left(-1+\left(1+x_i^c\right)^k\right)^{-\lambda}},$$

$$\frac{\delta L}{\delta\beta}=\frac{n}{\beta}-\sum_{i=1}^n\mathrm{Log}\left[1+\left(-1+(1+x_i^c)^k\right)^{-\lambda}\right]$$

$$\frac{\delta L}{\delta\lambda}=\frac{n}{\lambda}-\sum_{i=1}^n\mathrm{Log}[-1+(1+x_i^c)^k]-$$

$$(1+\beta)\sum_{i=1}^n-\frac{\mathrm{Log}[-1+(1+x_i^c)^k](-1+(1+x_i^c)^k)^{-\lambda}}{1+(-1+(1+x_i^c)^k)^{-\lambda}}.$$

We can get the estimates of the unknown parameters by equating equations to zero and solving them numerically.

8. Simulation study

In simulation study, we generated random variables in proposed model with value of four different parameters for sample size n= {25, 50, 100}.
By maximum likelihood method parameters are estimated with using each generated random variable, these estimated parameters are used for the MSE and Bias of LDSMD. Computational software "Mathematica" is used for all simulation. In table 2 and 3 results are reported.

Table 2: For the Data Leukaemia-Free Survival Times of 50 Patients MSE and Average Mean of Bias

n	Bias($\hat{c}$)	Bias($\hat{k}$)	Bias($\hat{\lambda}$)	Bias($\hat{\beta}$)	MSE($\hat{c}$)	MSE($\hat{k}$)	MSE($\hat{\lambda}$)	MSE($\hat{\beta}$)
25	0.5890	0.9572	−72.2842	2.5964	1.6950	0.2038	30665.88	2.8494
50	0.4387	0.6156	−52.0084	1.6717	0.5313	5.0934	21187.27	1.4248
100	0.4123	0.4116	−49.1837	1.4401	0.4005	0.1907	21448.53	1.4035

Table 3: For the Data of Failure of Eighteen Electronic Devices MSE Values and Average Mean of Bias

n	Bias($\hat{c}$)	Bias($\hat{k}$)	Bias($\hat{\lambda}$)	Bias($\hat{\beta}$)	MSE($\hat{c}$)	MSE($\hat{k}$)	MSE($\hat{\lambda}$)	MSE($\hat{\beta}$)
25	-1.4559	0.50777	−1.38894	7.70205	2.24302	0.29178	17252.54	6.60165
50	−1.3468	0.43116	−1.38893	2.20285	2.14764	0.21054	17252.53	4.72285
100	-1.21219	0.33369	−1.29788	1.07593	1.64885	0.13929	774.67745	3.18139

9. Comparative analysis of the log-dagum singh maddala distribution

We evaluated the effectiveness of the Log-Dagum Singh Maddala (LDSM) distribution by comparing it with seven other distributions using three data sets. Our analysis employed four goodness-of-fit test statistics: Cramér-von Mises (CVM), Anderson-Darling (AD), Kolmogorov-Smirnov (K-S), and p-values. All calculations are performed using the computational software Mathematica, ensuring precise and efficient computations. The results demonstrate the superior fit of the LDSM distribution in all three data sets, confirming its versatility and reliability in modeling real-world phenomena. The LDSM distribution outperformed the other distributions in terms of goodness-of-fit test statistics, highlighting its effectiveness in capturing the underlying data patterns.

Data 1: This Data Set Consists of 50 Patient's Leukaemia-Free Survival Time with Autologous Transplant Presented Below

0.030	0.493	0.855	1.184	1.283	1.480	1.776	2.138	2.500	2.763
2.993	3.224	3.421	4.178	4.441	5.691	5.855	6.941	6.941	7.993
8.882	8.882	9.145	11.480	11.513	12.105	12.796	12.993	13.849	16.612
17.138	20.066	20.329	22.368	26.776	28.717	28.717	32.928	33.783	34.211
34.770	39.539	41.118	45.033	46.053	46.941	48.289	57.401	58.322	60.625

Table 4: For the Data Set 1 (AD), (CVM) the (K-S) Statistics and P-Values

Distributions	A^*	W^*	K-S	p-value
LDSMD	0.15444	0.01792	0.05671	0.99709
SMD	4.8874	0.93761	0.22769	0.01120
LDWD	0.40399	0.06517	0.07695	0.94357
GD	0.36998	0.04963	0.08476	0.86513
WD	0.41154	0.056242	0.08685	0.84501
LD	2.50484	0.37995	0.19666	0.04182
NEED	0.66609	0.09625	0.09064	0.80595
EED	0.36283	0.04839	0.08444	0.86817

Table 5: For Data Set 1, Information Criteria of Different Distributions

Distributions	AIC	AICC	BIC	HQIC	CAIC
LDSMD	392.112	393.001	399.760	395.025	393.001
SMD	435.348	435.602	439.171	436.803	435.602
LDWD	394.682	395.140	398.235	396.37	395.140
GD	395.057	395.312	398.881	396.51	395.312
WD	395.433	395.689	399.257	396.89	395.689
LD	394.783	395.304	400.519	396.97	395.304
NEED	396.045	396.301	399.869	397.50	396.301
EED	394.954	395.209	398.778	396.41	395.209

Data 2: This Data Set Consists Eighteen Electronic Devices Times to Failure

5	11	21	31	46	75	98	122	145
165	196	224	245	293	321	330	350	420

Table 6: For the Data Set 2 (AD), (CVM), the (K-S) Statistics and P-Values

Distributions	A^*	W^*	K-S	$p-$value
LDSMD	0.2622	0.0405	0.1015	0.9925
SMD	4.6864	0.9375	0.9121	0.0143
LDWD	0.1725	0.0236	0.0840	0.9996
GD	0.4487	0.0699	0.1206	0.9561
WD	0.4609	0.0644	0.1132	0.9752
LD	28.2328	5.0981	0.9157	1.54817×10^{-13}
NEED	2.4695	0.4826	0.2814	0.1155
EED	0.4456	0.0708	0.1214	0.9535

Table 7: Information Criteria for Data 2

Model	AIC	BIC	HQIC	AICC	CAIC
LDSMD	225.340	228.902	225.832	228.418	228.418
SMD	266.104	269.665	266.595	269.181	269.181
LDWD	208.29	210.963	208.659	210.006	210.006
WD	395.43	397.214	395.679	396.233	396.233
GD	226.1	227.9	226.9	229.9	229.9
NEED	237.86	239.640	238.105	238.660	238.660
LD	341.41	343.196	341.661	342.215	342.215
EED	225.253	227.034	225.49	226.053	226.053

Data 3: This Data Consist Breaking Stress of Carbon Fibres for 100 Uncensored Data

0.39	0.81	0.85	0.98	1.08	1.12	1.17	1.18	1.22	1.25
1.36	1.41	1.47	1.57	1.57	1.59	1.59	1.61	1.61	1.69
1.69	1.71	1.73	1.8	1.84	1.84	1.87	1.89	1.92	2
2.03	2.03	2.05	2.12	2.17	2.17	2.17	2.35	2.38	2.41
2.43	2.48	2.48	2.5	2.53	2.55	2.55	2.56	2.59	2.67
2.73	2.74	2.76	2.77	2.79	2.81	2.81	2.82	2.83	2.85
2.87	2.88	2.93	2.95	2.96	2.97	2.97	2.9	3.09	3.11
3.11	3.15	3.15	3.19	3.19	3.22	3.22	3.27	3.28	3.31
3.31	3.33	3.39	3.39	3.51	3.56	3.6	3.65	3.68	3.68
3.7	3.75	4.2	4.38	4.42	4.7	4.9	4.91	5.08	5.56

Table 8: For the Data Set 3 (AD), (CVM), the (K-S) Statistics and P-Values

Distributions	A ∗	W ∗	K-S	p-value
LDSMD	0.2799	0.0438	0.0553	0.9195
SMD	4.8874	0.9376	0.2277	0.0112
LDWD	0.3967	0.0650	0.0618	0.839
GD	200.5016	32.9885	0.9996	2.2204×10^{-16}
WD	18.9521	3.7772	0.3341	4.0284×10^{-10}
LD	79.3018	17.3623	0.8210	-2.2204×10^{-16}
NEED	16.9307	3.3516	0.3170	3.7314×10^{-9}
EED	1.2341	0.2303	0.1077	0.1962

Table 9: Information Criteria for Data Set 3

Model	AIC	BIC	AICC	HQIC	CAIC
LDSMD	289.708	300.129	290.129	293.925	290.129
SMD	382.967	386.791	383.222	384.423	383.222
LDWD	288.62	296.43	288.869	296.488	296.4880
WD	289.06	296.87	289.309	292.222	289.309
GD	290.467	295.678	290.591	292.576	290.591
NEED	393.847	399.058	393.971	395.956	393.971
LD	474.087	477.911	474.343	475.544	474.343
EED	296.365	301.57	296.488	298.473	296.488

To evaluate the goodness of fit, various criteria are employed to compare the fitted models. Generally, smaller values indicate a better fit to the data. In this study, we utilized Mathematica 11.0 for all computational tasks, ensuring precise and efficient calculations.

Tables 4-9 show that the Log-Dagum Singh Maddala (LDSM) model yields higher p-values, indicating a better fit compared to other distributions (SM, LDWD, WD, GD, LD, EED and NEED) for the estimated parameters. To visualize the fit, the probability density functions (pdfs) of the distributions are superimposed on the histograms of the three data sets, as shown in figures 4-6. These plots provide a visual confirmation of the LDSM model's superior fit, demonstrating its ability to accurately capture the characteristics of the data.

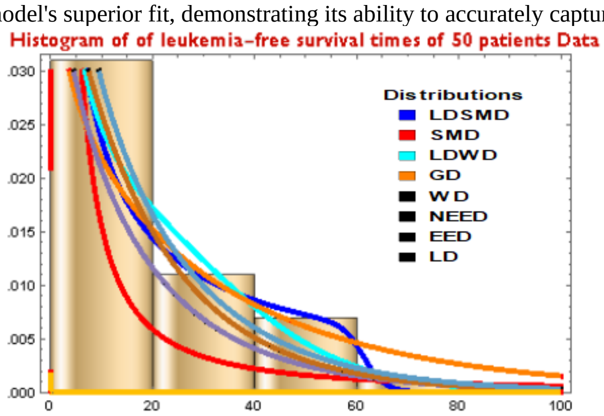

Fig. 4: Fitted Densities for Data 1.

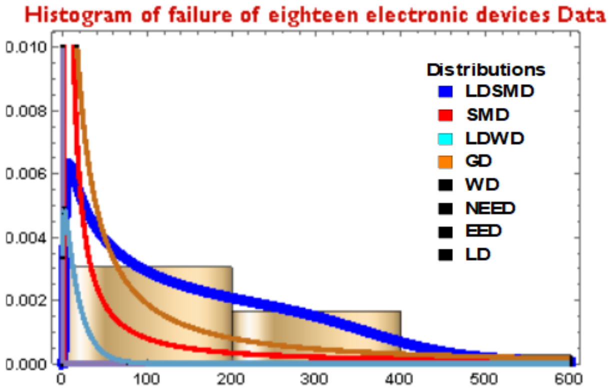

Fig. 5: Fitted Densities for Data 2.

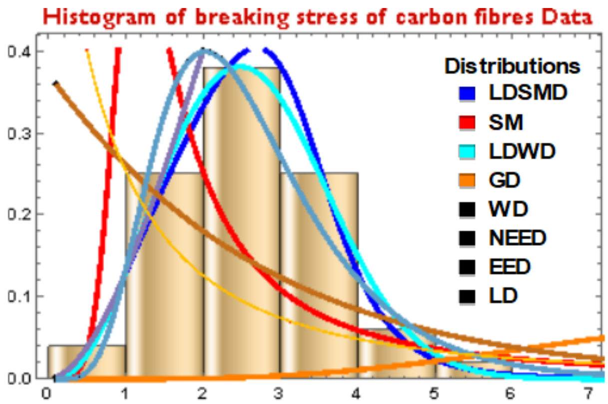

Fig. 6: Fitted Densities for Data 3.

10. Conclusion

This paper explores the Log-Dagum Singh Maddala (LDSM) distribution a 4-parameter extension of the traditional distributions. We thoroughly investigate the LDSM distribution deriving its statistical properties, parameter estimation methods, quantile points and characterizations. The practical applicability of the LDSM distribution is demonstrated through three real-life datasets, showcasing its superior fit compared to established distributions like Singh Maddala, Weibull, Log-Dagum Weibull, Lomax, Gamma, Exponentiated Exponential, and Nadarajah Exponentiated Exponential. Our findings indicate that the LDSM distribution is a versatile and reliable model for various fields including theoretical and applied sciences, engineering, hydrology, survival analysis, lifetime data and economics. We anticipate that our results will be valuable to researchers and practitioners seeking accurate and robust statistical modeling solutions.

References

[1] Afify A. Z., M. Alizadeh. “The Odd Dagum Family of Distributions: Properties and Applications”, journal of Applied Probability and Statistics, Vol. 15, 2020, 45–72.

[2] Ahsanullah M., Nevzorov V. B., and Shakil M. “An Introduction to Order Statistics”, Atlantis Press, Paris, France, 2013a, 1879-6893. https://doi.org/10.2991/978-94-91216-83-1.

[3] Ahsanullah M. “Record Statistics”, Nova Science Publishers, New York, USA, 1995.

[4] Alzaatreh A., Lee C., and Famoye F. “A new method for generating families of continuous distributions”, Metron, Vol. 71, No. 1, 2013, 63-79. https://doi.org/10.1007/s40300-013-0007-y.

[5] Alzaatreh A., Famoye F. and Lee C. “T-normal family of distributions: A new approach to generalize the normal distribution”, Journal of Statistical Distributions and Applications, Vol. 16, No.1. 2014b. https://doi.org/10.1186/2195-5832-1-16.

[6] Alzaghal A., Famoye F., and Lee C. “Exponentiated T-X family of distributions with some applications”, International Journal of Probability and Statistics, Vol. 2, No. 3, 2013, 31–49. https://doi.org/10.5539/ijsp.v2n3p31.

[7] Aljarrah M. A., Lee C., and Famoye F. “On generating T-X family of distributions using quantile functions”, Journal of Statistical Distributions and Applications, Vol. 2, No.1.2014. https://doi.org/10.1186/2195-5832-1-2.

[8] Arnold B. C., Balakrishnan Nagaraja H. N. “First Course in Order Statistics”, Wiley, New York, USA, 2005.

[9] Aslam M., Asghar Z., Hussain Z., and Farooq S. S. “A modified T-X family of distributions classical and Bayesian analysis”, Journal of Taibah University for Science, Vol. 14, No. 1, 2020, 254–264. https://doi.org/10.1080/16583655.2020.1732642.

[10] Cordeiro G.M., Ortega E.M.M.and da Cunha D.C.C. “The exponentiated generalized class of distributions”, Journal of Data Science, Vol. 11, 2013.1–27. https://doi.org/10.6339/JDS.201301_11(1).0001.

[11] Cordeiro G.M., de Castro, M. “A new family of generalized distributions”, Journal of Statistical Computation and Simulation, 81, 7, 2011. 883–898 https://doi.org/10.1080/00949650903530745.

[12] David H. A., Nagaraja H. N. Order Statistics, Third Edition, Wiley, New York, USA, 2003. https://doi.org/10.1002/0471722162.

[13] Domma F. “Kurtosis diagram for the log-Dagum distribution”, Statistica and Applicatzioni, Vol. 2, 2004, 3–23.

[14] Eugene N., Lee C., and Famoye F. “Beta-normal distribution and its applications”, Communication Statistical Theory & Methods, Vol.31, No.4, 2002. 497 – 512. https://doi.org/10.1081/STA-120003130.

[15] Handique L., Akbar M., Mohsin M., and Jamal F. Properties and applications of a new member of the T-X family of distributions, Thailand statistician, Vol.19, 2021, 248-260.

[16] Jamal F., Nasir M. A. (2019). "Some new members of the T-X family of distributions", Proceedings of the 17th International Conference on Statistical Sciences, Lahore, Pakistan.2019.

[17] Khadim A., Saghir A., Hussain T., Shakil M., and Ahsanullah M. "Log Dagum Weibull distribution", Applied and Computational Mathematics, Vol.10, No.5, 2021, 100-113. https://doi.org/10.11648/j.acm.20211005.11.

[18] Nevzorov V. B. "Records: Mathematical Theory-Translation of Mathematical Monograph", American Mathematical Society, Rhode Island, USA. 2001. https://doi.org/10.1090/mmono/194.

[19] Polisicchio M., and Zenga M. "Kurtosis diagram for continuous variables", Metron, Vol.55, 1997, 21–41.

[20] Rasha M. Mandouh., Mahmoud R., and Rasha E. Abdelatty. "A new (T X^θ) family of distributions", Scientific Reports, vol.14, 2024. https://doi.org/10.1038/s41598-023-49425-2.

[21] Risti´c M. M., and Balakrishnan N. "The gamma-exponentiated exponential Distribution", Journal of Statistical Computation and Simulation, Vol. 82, 2012, 1191–1206. https://doi.org/10.1080/00949655.2011.574633.

[22] Shakil M., Kibria B.M., and Ahsanullah M. "Some Inferences on Dagum (4P) Distribution", Statistical Distribution and their Applications, Vol.154, 2021, 1-33.

[23] Shakil M., Kibria B. M., Singh J.N., and Ahsanullah M. "On Burr (4P) Distribution Application of Breaking Stress Data" , Statistical Distributions and their Applications, Vol. 2, 2021, 190-199. https://doi.org/10.1080/00949655.2011.574633.

[24] Singh S.K., and Mandela G.S. "A function for size distribution of Incomes", Econometric, Vol. 44, 1976, 963–970. https://doi.org/10.2307/1911538.

[25] Tahir M.H., Zubair M., Mansoor M., Gauss M., Cordeiro Morad Alizadeh., and Hamedani G. G. "A new Weibull-G family of distributions", Hacettepe Journal of Mathematics and Statistics, Vol. 45, No. 2, 2016, 629-647. https://doi.org/10.15672/HJMS.2015579686.

Appendix

Proof preposition 1:

$E[Y \mid Y \geq y] = \frac{1}{1-F(y)}\int_0^y sf(s)ds \, , u(y)f(y) = \int_0^y sf(s)ds$

$u(y) = \frac{\int_0^y sf(s)ds}{f(y)} = \frac{y(1-F(y)+\int_0^y(1-F(s))ds}{f(y)}$

Putting (2) and (1) then it is simply seen that

$$u(y) = \frac{y(1-\left[1+[(1+x^c)^k-1\,]^{-\lambda}\right]^{-\beta} + \int_0^y(1-\left[1+[(1+t^c)^k-1\,]^{-\lambda}\right]^{-\beta}\mathrm{ds}}{\left[1+[(1+x^c)^k-1\,]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}$$

Simple differentiation and simplification gives $u^0(y) = -y - A(y)u(y)$,
Where

$$A(y) = \frac{f`(y)}{f(y)} = \frac{\beta\lambda \mathrm{ck}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\left(1+\left(-1+(1+x^c)^k\right)^{-\lambda}\right)^{-1-\beta}[c(-1+k)\mathrm{kx}^{-2+2c}(1+x^c)^{-1}+(-1+c)x^{-2+c}+ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1}(-1-\lambda)-ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}(1+(-1+(1+x^c)^k)^{-\lambda})^{-1}(-1-\beta)\lambda]}{\left[1+[(1+x^c)^k-1\,]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}$$

From which we obtain

$$\frac{-y+u`(y)}{u(y)} = \left\{\frac{\beta\lambda \mathrm{ck}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\left(1+\left(-1+(1+x^c)^k\right)^{-\lambda}\right)^{-1-\beta}[c(-1+k)\mathrm{kx}^{-2+2c}(1+x^c)^{-1}+(-1+c)x^{-2+c}+ ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1}(-1-\lambda)-ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}(1+(-1+(1+x^c)^k)^{-\lambda})^{-1}(-1-\beta)\lambda]}{\left[1+[(1+x^c)^k-1\,]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}\right\}$$

We have

$\frac{f`(y)}{f(y)} = \frac{-y+u`(y)}{u(y)}$

It follows that

$$\frac{f`(y)}{f(y)} = \left\{\frac{\beta\lambda \mathrm{ck}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\left(1+\left(-1+(1+x^c)^k\right)^{-\lambda}\right)^{-1-\beta}[c(-1+k)\mathrm{kx}^{-2+2c}(1+x^c)^{-1}+(-1+c)x^{-2+c}+ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1}(-1-\lambda)-ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}(1+(-1+(1+x^c)^k)^{-\lambda})^{-1}(-1-\beta)\lambda]}{\left[1+[(1+x^c)^k-1\,]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}\right\}$$

Integrating with respect to 'y' and after simplifying we obtain

$$lnf(y) = \ln\,[C\left[1+[(1+x^c)^k-1\,]^{-\lambda}\right]^{-\beta-1}x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck$$

Since C is determined by $\int_0^y f(y)$) = 1, we get the pdf.

Proof preposition 2:

$E[Y|Y \geq y] = \frac{1}{1-F(y)}\int_y^\infty tf(t)dt$, $f(y)s(y) = \int_y^\infty tf(t)dt$,

$$s(y) = \frac{\int_y^\infty tf(t)dt}{f(y)} = \frac{y(1-F(y)+\int_y^\infty (1-F(t))dt}{f(y)},$$

Substituting (8) and (9). Then it is easily seen that

$$s(y) = \frac{y(1-\left[1+\left[(1+x^c)^k-1\right]^{-\lambda}\right]^{-\beta} + \int_y^\infty (1-\left[1+\left[(1+t^c)^k-1\right]^{-\lambda}\right]^{-\beta} \mathrm{dt}}{\left[1+[(1+x^c)^k-1]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}$$

Simple differentiation and simplification gives $s'(y) = -y - s(y)A(y)$, where

$$A(y) = \frac{f`(y)}{f(y)} = \frac{\beta\lambda \mathrm{ck}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\left(1+\left(-1+(1+x^c)^k\right)^{-\lambda}\right)^{-1-\beta}[c(-1+k)\mathrm{kx}^{-2+2c}(1+x^c)^{-1} + (-1-\lambda)-ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}(1+(-1+(1+x^c)^k)^{-\lambda})^{-1}(-1-\beta)\lambda]}{\left[1+[(1+x^c)^k-1]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}$$

Then we
From which we obtain

$$\frac{-y+s`(y)}{s(y)} = \left\{\frac{\beta\lambda \mathrm{ck}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\left(1+\left(-1+(1+x^c)^k\right)^{-\lambda}\right)^{-1-\beta}[c(-1+k)\mathrm{kx}^{-2+2c}(1+x^c)^{-1} + (-1+c)x^{-2+c}+ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1}(-1-\lambda)-ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}(1+(-1+(1+x^c)^k)^{-\lambda})^{-1}(-1-\beta)\lambda]}{\left[1+[(1+x^c)^k-1]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}\right\}$$

We have

$$\frac{f`(y)}{f(y)} = \frac{-y+s`(y)}{s(y)}$$

It follows that

$$\frac{f`(y)}{f(y)} = \left\{\frac{\beta\lambda \mathrm{ck}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}\left(1+\left(-1+(1+x^c)^k\right)^{-\lambda}\right)^{-1-\beta}[c(-1+k)\mathrm{kx}^{-2+2c}(1+x^c)^{-1}+(-1+c)x^{-2+c}+ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1}(-1-\lambda)- ckx^{-2+2c}(1+x^c)^{-1+k}\left(-1+(1+x^c)^k\right)^{-1-\lambda}(1+(-1+(1+x^c)^k)^{-\lambda})^{-1}(-1-\beta)\lambda]}{\left[1+[(1+x^c)^k-1]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck}\right\}$$

Integrating with respect to 'y' after simplification we get

$$lnf(y) = \ln\left[C\left[1+[(1+x^c)^k-1]^{-\lambda}\right]^{-\beta-1} x^{-1+c}(1+x^c)^{-1+k}(-1+(1+x^c)^k)^{-1-\lambda}\beta\lambda ck\right]$$

Since C is determined by $\int_0^\infty f(y)$) = 1, we obtain the pdf