Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v9n1p169

The Dagum-Poisson distribution: model, properties and application

By Oluyede et al.

Published: 26 April 2016

Electronic Journal of Applied Statistical Analysis
Vol. 09, Issue 01, April 2016, 169-197
DOI: 10.1285/i20705948v9n1p169

The Dagum-Poisson distribution: model, properties and application

Broderick O. Oluyede*[a], Galelhakanelwe Motsewabagale[b], Shujiao Huang[c], Gayan Warahena-Liyanage[d], and Mavis Pararai[e]

[a] *Georgia Southern University, Department of Mathematical Sciences, Statesboro, GA 30460*
[b] *Botswana International University of Science and Technology, Department of Mathematical and Computational Sciences, Palapye, BW*
[c] *University of Houston, Department of Mathematics, Houston, TX*
[d] *Central Michigan University, Department of Mathematics, Mount Pleasant, MI*
[e] *Indiana University of Pennsylvania, Department of Mathematics, Indiana, PA*

Published: 26 April 2016

A new four parameter distribution called the Dagum-Poisson (DP) distribution is introduced and studied. This distribution is obtained by compounding Dagum and Poisson distributions. The structural properties of the new distribution are discussed, including explicit algebraic formulas for its survival and hazard functions, quantile function, moments, moment generating function, conditional moments, mean and median deviations, Bonferroni and Lorenz curves, distribution of order statistics and Rényi entropy. Method of maximum likelihood is used for estimating the model parameters. A Monte Carlo simulation study is conducted to examine the bias, mean square error of the maximum likelihood estimators and width of the confidence interval for each parameter. A real data set is used to illustrate the usefulness, applicability, importance and flexibility of the new distribution.

keywords: Dagum distribution, Dagum Poisson distribution, Poisson distribution, Moments, Maximum Likelihood Estimation.

1. Introduction

There are several new distributions in the literature for modeling lifetime data obtained by compounding distributions, including work by Barreto-Souza et al. (2011) on the

*Corresponding author: boluyede@georgiasouthern.edu.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

Weibull-geometric distribution, and Lu and Shi (2012) on the Weibull-Poisson distribution, among others. These new families of probability distributions that extend well-known families of distributions are very useful in modeling lifetime data and are of tremendous practical importance in several areas including economics, finance, reliability and medicine. Flexibility and applications have been key to the study of generalized distributions in recent years. Simplicity, extensions and applicability are are also important components in the derivations of these new distributions. Of particular importance are models that can be applied to various settings including size distribution of personal income and lifetime data, where Dagum distribution and its various extensions could be very useful. In a recent note, Huang and Oluyede (2014), developed a new class of distribution called the exponentiated Kumaraswamy Dagum distribution and applied the model to income and lifetime data. Oluyede et al. (2014) also developed and presented results on the gamma-Dagum distribution with applications to income and lifetime data.

The class of distributions generated by compounding well-known lifetime distributions such as exponential, Weibull, generalized exponential, exponentiated Weibull, inverse Weibull, logistic and log-logistic with some discrete distributions such as binomial, geometric, zero-truncated Poisson, logarithmic and the power series distributions in general are part of the recent developments in generating useful distributions. In this setting, the non-negative random variable X denoting the lifetime of such a system is defined by $X = \min_{1\leq i\leq N} X_i$ or $X = \max_{1\leq i\leq N} X_i$, where the distribution of X_i follows one of the lifetime distributions and the random variable N follows some discrete distribution mentioned above. This new class of distributions has received considerable attention over the past several years.

In the literature on statistical distributions, several authors have proposed new distributions that are far more flexible in modeling monotone or unimodal failure rates but they are not useful for modeling bathtub shaped or non-monotone failure rates. Adamidis and Loukas (1998) introduced a two-parameter exponential-geometric (EG) distribution by compounding an exponential distribution with a geometric distribution. Adamidis et al. (2005) proposed the extended exponential-geometric (EEG) distribution which generalizes the EG distribution and discussed its various statistical properties along with its reliability features. The exponential Poisson (EP) and exponential logarithmic (EL) distributions were introduced and studied by Kuş (2007), and Tahmasbi and Rezaei (2008), respectively. Recently, Chahkandi and Ganjali (2009) proposed the exponential power series (EPS) family of distributions, which contains as special cases these distributions. Barreto-Souza et al. (2011), and Lu and Shi (2012) introduced the Weibull-geometric (WG) and Weibull-Poisson (WP) distributions which are extensions of the EG and EP distributions, respectively. Barreto-Souza and Cribari-Neto (2009) presented a generalization of the exponential-Poisson distribution. Morais and Barreto-Souza (2011) developed and presented a compound class of Weibull and power series distributions.

The primary motivation for the development of the Dagum-Poisson distribution is the modeling of size distribution of personal income and lifetime data with a diverse model that takes into consideration not only shape, and scale but also skewness, kurtosis and tail variation. Also, motivated by various applications of Poisson and Dagum distributions in several areas including reliability, exponential tilting (weighting) in finance and actuarial sciences, as well as economics, where Dagum distribution plays an important role in size distribution of personal income, we construct and develop the statistical properties of this new class of generalized Dagum-type distribution called the Dagum-Poisson distribution and apply it to real lifetime data in order to demonstrate the usefulness of the proposed distribution. In this regard, we propose a new four-parameter distribution, called the Dagum-Poisson (DP) distribution.

1.1. Dagum Distribution

Dagum distribution was proposed by Camilo Dagum (see Dagum (1977) for details). Dagum distribution is used to fit empirical income and wealth data, that could accommodate both heavy tails in empirical income and wealth distributions, and also permit interior mode. Dagum distribution is a special case of generalized beta distribution of the second kind (GB2), McDonald (1984), McDonald and Xu (1995), when the parameter $q = 1$, where the probability density function (pdf) of the GB2 distribution is given by:

$$f_{GB2}(y; a, b, p, q) = \frac{ay^{ap-1}}{b^{ap}B(p,q)[1 + (\frac{y}{b})^a]^{p+q}} \quad \text{for } y > 0.$$

Note that $a > 0, p > 0, q > 0$, are shape parameters and $b > 0$ is a scale parameter, and $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ is the beta function. Kleiber (2008) traced the genesis of Dagum distribution and summarized several statistical properties of this distribution. See Kleiber and Kotz (2003) for additional results on income and size distributions. Domma et al. (2011) obtained the maximum likelihood estimates of the parameters of Dagum distribution for censored data. Domma and Condino (2013) presented the beta-Dagum distribution. Huang and Oluyede (2014) presented the exponentiated Kumaraswamy Dagum distribution and applied the model to income and lifetime data. Oluyede et al. (2014) developed a generalized gamma-Dagum distribution with applications to income and lifetime data.

The probability density function (pdf) and cumulative distribution function (cdf) of Dagum distribution are given by:

$$f_D(x; \lambda, \beta, \delta) = \beta\lambda\delta x^{-\delta-1}(1 + \lambda x^{-\delta})^{-\beta-1} \tag{1}$$

and

$$F_D(x; \lambda, \beta, \delta) = (1 + \lambda x^{-\delta})^{-\beta}, \tag{2}$$

for $x > 0$, where λ is a scale parameter, δ and β are shape parameters. Dagum (1977) refers to his model as the generalized logistic-Burr distribution. The survival and hazard rate functions of Dagum distribution are $\overline{G}_D(x; \lambda, \delta, \beta) = 1 - (1 + \lambda x^{-\delta})^{-\beta}$ and $h_D(x; \lambda, \delta, \beta) = \frac{\beta\lambda\delta x^{-\delta-1}(1+\lambda x^{-\delta})^{-\beta-1}}{1-(1+\lambda x^{-\delta})^{-\beta}}$, respectively. The r^{th} raw or non central moments of Dagum distribution are given by

$$E(X^r) = \beta\lambda^{\frac{r}{\delta}} B\left(\beta + \frac{r}{\delta}, 1 - \frac{r}{\delta}\right), \tag{3}$$

for $\delta > r$, and $\lambda, \beta > 0$, where $B(\cdot, \cdot)$ is the beta function. The q^{th} percentile of the Dagum distribution is

$$x_q = \lambda^{\frac{1}{\delta}} \left(q^{\frac{-1}{\beta}} - 1 \right)^{\frac{-1}{\delta}}. \quad (4)$$

Dagum distribution has positive asymmetry, and its hazard rate can be monotonically decreasing, upside-down bathtub and bathtub followed by upside-down bathtub (Domma, 2002). This behavior has led several authors to study the model in different fields. In fact, recently, the Dagum distribution has been studied from a reliability point of view and used to analyze survival data (see Domma et al. (2011)). For the proposed model in this paper, it is possible to verify that the hazard rate function is more flexible than that of Dagum distribution. Actually when $\beta = 1$; Dagum distribution is referred to as the log-logistic distribution. Burr type III distribution is obtained when $\lambda = 1$.

This paper is organized as follows. In section 2, we define the DP distribution, its probability density function (pdf) and cumulative distribution function (cdf). In section 3, some properties of the new distribution including the expansion of the density, quantile function, moments and moment generating function are presented. Mean and median deviations, Bonferroni and Lorenz curves are derived in section 4. The distribution of order statistics and Rényi entropy are given in section 5. Maximum likelihood estimates of the unknown parameters are presented in section 6. A simulation study is conducted in order to examine the bias, mean square error of the maximum likelihood estimators and width of the confidence interval for each parameter of the model in section 7. Section 8 contains an application of the model to real data, followed by concluding remarks.

2. Dagum-Poisson Distribution, Sub-models and Properties

Suppose that the random variable X has the Dagum distribution where its cdf and pdf are given in equations (1) and (2), respectively. Given N, let $X_1, ..., X_N$ be independent and identically distributed random variables from Dagum distribution. Suppose N is a discrete random variable with a power series distribution (truncated at zero) and probability mass function (pmf) given by

$$P(N = n) = \frac{a_n \theta^n}{C(\theta)}, \quad n = 1, 2, ...,$$

where $a_n \geq 0$ depends only on n, $C(\theta) = \sum_{n=1}^{\infty} a_n \theta^n$ and $\theta \in (0, b)$ (b can be ∞) is chosen such that $C(\theta)$ is finite and its three derivatives with respect to θ are defined and given by $C'(.), C''(.)$ and $C'''(.)$, respectively. The power series family of distributions includes binomial, Poisson, geometric and logarithmic distributions Johnson et al. (1994). See Table 1 for some useful quantities including a_n, $C(\theta)$, $C(\theta)^{-1}$, $C'(\theta)$, and $C''(\theta)$ for the Poisson, geometric, logarithmic and binomial distributions.

The general form of the cdf and pdf of the Dagum-power series distribution are given by

$$F_{DPS}(x; \lambda, \delta, \beta, \theta) = \frac{C(\theta[1 + \lambda x^{-\delta}]^{-\beta})}{C(\theta)}, \quad (5)$$

Table 1: Useful Quantities for Some Power Series Distributions

Distribution	$C(\theta)$	$C'(\theta)$	$C''(\theta)$	$C(\theta)^{-1}$	a_n	Parameter Space
Poisson	$e^{\theta}-1$	e^{θ}	e^{θ}	$\log(1+\theta)$	$(n!)^{-1}$	$(0,\infty)$
Geometric	$\theta(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$2(1-\theta)^{-3}$	$\theta(1+\theta)^{-1}$	1	$(0,1)$
Logarithmic	$-\log(1-\theta)$	$(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$1-e^{-\theta}$	n^{-1}	$(0,1)$
Binomial	$(1+\theta)^m-1$	$m(1+\theta)^{m-1}$	$m(m-1)(1+\theta)^{m-2}$	$(\theta-1)^{1/m}-1$	$\binom{m}{n}$	$(0,1)$

and

$$f_{DPS}(x;\lambda,\delta,\beta,\theta)=\frac{\beta\lambda\delta\theta x^{-\delta-1}(1+\lambda x^{-\delta})^{-\beta-1}C'(\theta[1+\lambda x^{-\delta}]^{-\beta})}{C(\theta)}, \tag{6}$$

where $C(y)=\sum_{n=1}^{\infty}a_n y^n$ with $a_n>0$ depends only on n. Using Table 1, we can obtain Dagum-Poisson, Dagum-geometric, Dagum-logarithmic and Dagum-binomial distributions. In this note, we consider and study Dagum-Poisson distribution in detail. Let N be distributed according to the zero truncated Poisson distribution with pdf

$$P(N=n)=\frac{\theta^n e^{-\theta}}{n!(1-e^{-\theta})},\quad n=1,2,...,\theta>0.$$

Let X=max$(X_1,...,X_N)$, then the cdf of $X|N=n$ is given by

$$G_{X|N=n}(x) \;=\; \left[1+\lambda x^{-\delta}\right]^{-n\beta},\quad x>0,\ \lambda,\delta,\beta>0,$$

which is the Dagum distribution with parameters λ,δ and $n\beta$. The Dagum-Poisson (DP) distribution denoted by DP$(\lambda,\delta,\beta,\theta)$ is defined by the marginal cdf of X, that is,

$$F_{DP}(x;\lambda,\delta,\beta,\theta) \;=\; \frac{1-\exp(\theta[1+\lambda x^{-\delta}]^{-\beta})}{1-e^{\theta}}, \tag{7}$$

for $x>0,\lambda,\beta,\delta,\theta>0$. The DP density function is given by

$$f_{DP}(x;\lambda,\delta,\beta,\theta)=\frac{\beta\lambda\delta\theta x^{-\delta-1}(1+\lambda x^{-\delta})^{-\beta-1}\exp(\theta[1+\lambda x^{-\delta}]^{-\beta})}{e^{\theta}-1}. \tag{8}$$

A series representation of the DP cdf is given by

$$\begin{aligned}
F_{DP}(x;\lambda,\delta,\beta,\theta) &= \frac{1-\exp(\theta[1+\lambda x^{-\delta}]^{-\beta})}{1-e^{\theta}}\\
&= \frac{1}{e^{\theta}-1}\left[\sum_{k=0}^{\infty}\frac{\theta^k(1+\lambda x^{-\delta})^{-k\beta}}{k!}-1\right]\\
&= \frac{1}{e^{\theta}-1}\left[\sum_{k=1}^{\infty}\frac{\theta^k(1+\lambda x^{-\delta})^{-k\beta}}{k!}\right]\\
&= \sum_{k=0}^{\infty}\frac{\theta^{k+1}}{(e^{\theta}-1)(k+1)!}(1+\lambda x^{-\delta})^{-\beta(k+1)}\\
&= \sum_{k=0}^{\infty}\omega(k,\theta)F_D(x;\lambda,\delta,\beta(k+1)),
\end{aligned} \tag{9}$$

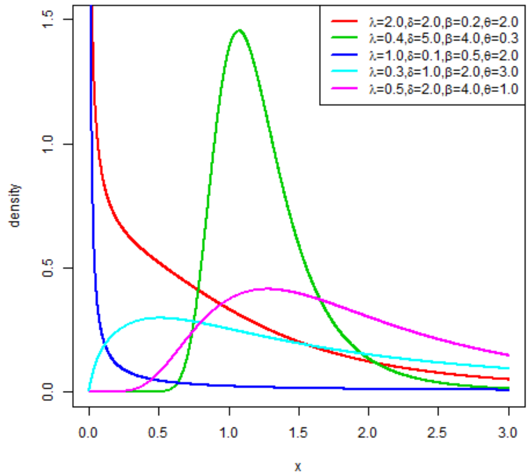

Figure 1: DP Density Functions

where $\omega(k,\theta) = \frac{\theta^{k+1}}{(k+1)!(e^{\theta}-1)}$, and $F_D(x;\lambda,\beta(k+1),\delta)$ is the Dagum cdf with parameters $\lambda, \beta(k+1), \delta > 0$.

2.1. Hazard and Reverse Hazard Functions

In this subsection, hazard and reverse hazard functions of the DP distribution are presented. The plots of DP pdf and hazard rate function for selected values of the model parameters λ, δ, β and θ are given in Figure 1 and Figure 2, respectively. The hazard and reverse hazard functions are given by

$$\begin{aligned} h_{DP}(x;\lambda,\delta,\beta,\theta) &= \frac{f_{DP}(x;\lambda,\delta,\beta,\theta)}{1-F_{DP}(x;\lambda,\delta,\beta,\theta)} \\ &= \frac{\lambda\delta\beta\theta(1+\lambda x^{-\delta})^{-\beta-1}e^{\theta(1+\lambda x^{-\delta})^{-\beta}}}{e^{\theta}-e^{\theta(1+\lambda x^{-\delta})^{-\beta}}} \end{aligned} \tag{10}$$

and

$$\begin{aligned} \tau_{DP}(x;\lambda,\delta,\beta,\theta) &= \frac{f_{DP}(x;\lambda,\delta,\beta,\theta)}{F_{DP}(x;\lambda,\delta,\beta,\theta)} \\ &= \frac{\lambda\delta\beta\theta(1+\lambda x^{-\delta})^{-\beta-1}e^{\theta(1+\lambda x^{-\delta})^{-\beta}}}{e^{\theta(1+\lambda x^{-\delta})^{-\beta}}-1}, \end{aligned} \tag{11}$$

respectively. The plots of the hazard rate function show various shapes including monotonically decreasing, monotonically increasing, unimodal, upside down bathtub and bathtub followed by upside down bathtub shapes for the combinations of the values of the parameters. This flexibility makes the DP hazard rate function suitable for both

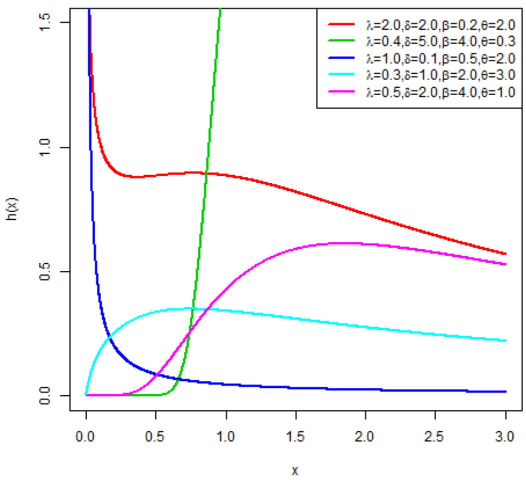

Figure 2: Plots of the DP Hazard Function

monotonic and non-monotonic empirical hazard behaviors that are likely to be encountered in real life situations. Unfortunately, the analytical analysis of the shape of both the density (except for zero modal when $\beta\delta \leq 1$, and unimodal if $\beta\delta > 1$, both for $\theta \to 0^+$) and hazard rate function seems to be very complicated.

2.2. Some Sub-models of the DP Distribution

DP distribution is a very flexible model that has different sub-models when its parameters are changed. DP distribution contains several sub-models including the following distributions.

- If $\beta = 1$, then DP distribution reduces to a new distribution called log-logistic Poisson (LLoGP) or Fisk-Poisson (FP) distribution with pdf given by

$$f_{LLOGP}(x; \lambda, \delta, \theta) = \lambda\theta\delta x^{-\delta-1} \frac{(1+\lambda x^{-\delta})^{-2} \exp(\theta(1+\lambda x^{-\delta})^{-1})}{\exp(\theta) - 1}, \quad x > 0.$$

  If in addition to $\beta = 1$, we have $\theta \downarrow 0$, then the resulting distribution is the log-logistic or Fisk distribution.

- When $\theta \downarrow 0$ in the DP distribution, we obtain Dagum (D) distribution.
- The Burr-III Poisson (BIIIP) distribution is obtained when $\lambda = 1$. If in addition, $\theta \downarrow 0$, Burr-III distribution with parameter $\delta, \beta > 0$ is obtained.

3. Some Statistical Properties

In this section, some statistical properties of DP distribution including expansion of density function, quantile function, moments, conditional moments, mean and median deviations, Lorenz and Bonferroni curves are presented.

3.1. Expansion of DP Density

In this subsection, we provide an expansion of the DP distribution. Note that, applying the fact that Maclaurin series expansion of $e^x = \sum_{k=0}^{\infty} x^k/k!$, the DP pdf can be written as follows:

$$
\begin{aligned}
f_{DP}(x;\lambda,\delta,\beta,\theta) &= \frac{\beta\lambda\delta\theta x^{-\delta-1}(1+\lambda x^{-\delta})^{-\beta-1}\exp(\theta(1+\lambda x^{-\delta})^{-\beta})}{\exp(\theta)-1} \\
&= \sum_{k=0}^{\infty} \frac{\theta^k}{k!(e^\theta-1)}\lambda\beta\delta\theta x^{-\delta-1}(1+\lambda x^{-\delta})^{-\beta(k+1)-1} \\
&= \sum_{k=0}^{\infty} \omega(k,\theta) f_D(x;\lambda,\delta,\beta(k+1)), \qquad (12)
\end{aligned}
$$

where $\omega(k,\theta) = \frac{\theta^{k+1}}{(k+1)!(e^\theta-1)}$, and $f_D(x;\lambda,\beta(k+1),\delta)$ is the Dagum pdf with parameters $\lambda,\beta(k+1),\delta,> 0$. Equation (12) also follows directly from equation (9). The above equation shows that the DP density is indeed a linear combination of Dagum densities. Hence, most of its mathematical properties can be immediately obtained from those of the Dagum distribution.

3.2. Quantile Function

In this subsection, we present the q^{th} quantile of the DP distribution. The q^{th} quantile of the DP distribution is obtained by solving the nonlinear equation

$$
\frac{e^{\theta(1+\lambda x^{-\delta})^{-\beta}}-1}{e^\theta-1} = U,
$$

where U is a uniform variate on the unit interval $[0,1]$. It follows that the q^{th} quantile of the DP distribution is given by

$$
X_q = \left(\frac{1}{\lambda}\left(\left[\frac{\ln(U(e^\theta-1)+1)}{\theta}\right]^{-1/\beta}-1\right)\right)^{-1/\delta}. \qquad (13)
$$

Consequently, random number can be generated based on equation (13).

3.3. Moments

In this subsection, we present the r^{th} moment of DP distribution. Moments are necessary and crucial in any statistical analysis, especially in applications. Moments can be used to study the most important features and characteristics of a distribution (e.g., tendency, dispersion, skewness and kurtosis). If the random variable X has a DP distribution, with parameter vector $\boldsymbol{\Theta} = (\lambda, \delta, \beta, \theta)$, then the r^{th} moment of X is given by

$$\begin{aligned} \mu_r' &= \int_0^\infty x^r f_{DP}(x; \lambda, \beta, \delta, \theta) dx \\ &= \sum_{k=0}^{\infty} \omega(k, \theta) \int_0^\infty x^r f_D(x; \lambda, \beta(k+1), \delta) dx \\ &= \sum_{k=0}^{\infty} \omega(k, \theta) \lambda^{\frac{r}{\delta}} B\left(\frac{r}{\delta} + \beta(k+1), 1 - \frac{r}{\delta}\right), \quad \delta > r. \end{aligned} \tag{14}$$

Note that the r^{th} non-central moment follows readily from the fact that DP pdf can be written as a linear combination of Dagum densities with parameters $\lambda, \beta(k+1), \delta > 0$. Based on the first four moments of the DP distribution, the measures of skewness CS and kurtosis CK of the DP distribution can obtained from

$$CS = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{\left[\mu_2 - \mu_1^2\right]^{\frac{3}{2}}} \quad \text{and} \quad CK = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{\left[\mu_2 - \mu_1^2\right]^2},$$

respectively. Plots of the skewness and kurtosis for selected choices of the parameter β as a function of δ, as well as for some selected choices of δ as a function of β are displayed in Figures 3, 4, 5, and 6. These plots clearly indicate that the skewness and kurtosis depend on the shape parameters δ and β. Table 2 lists the first six moments of the DP distribution for selected values of the parameters, by fixing $\delta = 6.5$. Table 3 lists the first six moments of the DP distribution for selected values of the parameters, by fixing $\lambda = 0.3$ and $\beta = 0.8$, and Table 4 lists the first six moments of the DP distribution for selected values of the parameters, by fixing $\delta = 8.5$ and $\beta = 1.5$. These values can be determined numerically using R and MATLAB.

Table 2: Moments of the DP distribution for some parameter values; $\delta = 6.5$.

μ_s'	$\lambda = 0.5, \beta = 2.5, \theta = 2.5$	$\lambda = 2.0, \beta = 0.5, \theta = 0.5$	$\lambda = 3.0, \beta = 1.5, \theta = 3.5$	$\lambda = 0.1, \beta = 0.2, \theta = 0.3$
μ_1'	1.30548	1.01330	1.65702	0.45015
μ_2'	1.81639	1.15863	2.93049	0.26502
μ_3'	2.74198	1.49281	5.62721	0.18674
μ_4'	4.65550	2.22301	12.15746	0.15704
μ_5'	9.74269	4.16566	32.37752	0.16950
μ_6'	38.13338	14.84150	161.26470	0.35254
SD	0.33484	0.36311	0.42985	0.24977
CV	0.25649	0.35835	0.25941	0.55485
CS	2.07731	1.07708	2.00161	0.72369
CK	15.69083	8.48089	15.25227	5.09487

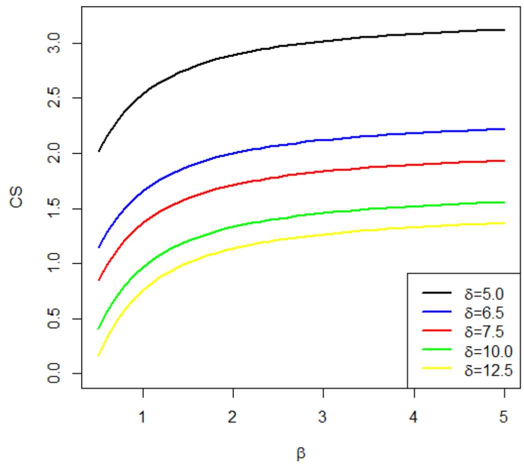

Figure 3: Skewness of DP distribution as a function of β for some values of δ with $\lambda = 0.5$ and $\theta = 1.5$.

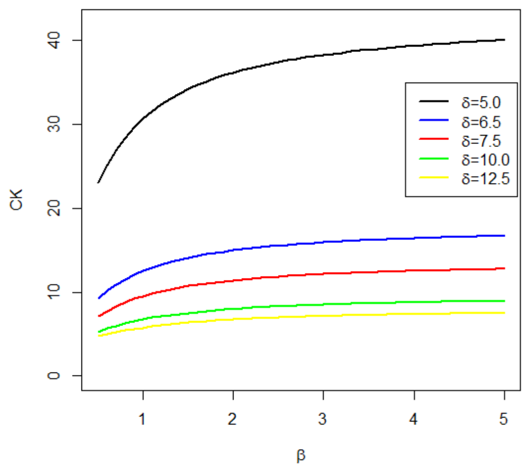

Figure 4: Kurtosis of DP distribution as a function of β for some values of δ with $\lambda = 0.5$ and $\theta = 1.5$.

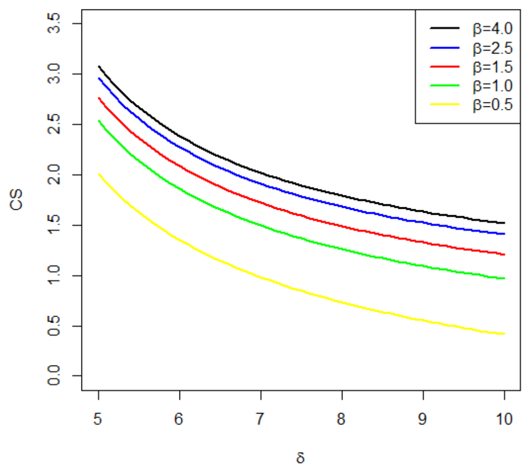

Figure 5: Skewness of DP distribution as a function of δ for some values of β with $\lambda = 0.5$ and $\theta = 1.5$.

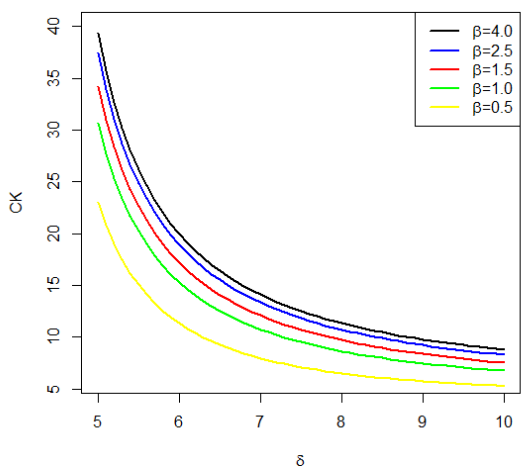

Figure 6: Kurtosis of DP distribution as a function of δ for some values of β with $\lambda = 0.5$ and $\theta = 1.5$.

Table 3: Moments of the DP distribution for some parameter values; $\lambda = 0.3$ and $\beta = 0.8$.

μ'_s	$\delta = 7.0, \theta = 1.0$	$\delta = 10.5, \theta = 0.3$	$\delta = 7.5, \theta = 0.1$	$\delta = 6.5, \theta = 3.0$
μ'_1	0.89281	0.88817	0.84321	1.01339
μ'_2	0.85993	0.81665	0.76169	1.10774
μ'_3	0.90057	0.77797	0.74088	1.32478
μ'_4	1.04747	0.76984	0.78724	1.79292
μ'_5	1.42600	0.79537	0.94477	3.00491
μ'_6	2.69499	0.86581	1.40007	9.45480
SD	0.25065	0.16674	0.22515	0.28424
CV	0.28074	0.18773	0.26702	0.28049
CS	1.31050	0.70110	1.15070	1.67466
CK	9.60878	5.71218	8.24167	12.95428

Table 4: Moments of the DP distribution for some parameter values; $\delta = 8.5$ and $\beta = 1.5$.

μ'_s	$\lambda = 1.5, \theta = 1.5$	$\lambda = 1.5, \theta = 0.5$	$\lambda = 0.5, \theta = 1.5$	$\lambda = 0.5, \theta = 0.5$
μ'_1	1.24022	1.17897	1.08985	1.03602
μ'_2	1.60034	1.44651	1.23580	1.11701
μ'_3	2.16080	1.85768	1.46628	1.26059
μ'_4	3.08248	2.52160	1.83811	1.50366
μ'_5	4.72733	3.68135	2.47717	1.92907
μ'_6	8.07349	5.98802	3.71766	2.75735
SD	0.24939	0.23780	0.21915	0.20897
CV	0.20108	0.20170	0.20108	0.20170
CS	1.40204	1.40918	1.40197	1.40905
CK	8.97403	8.95440	8.97488	8.95504

Similarly, the moment generating function (mgf) of X is given by

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{\infty} \omega(k,\theta) M_Y(t) \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \omega(k,\theta) \frac{t^j}{j!} \lambda^{\frac{j}{\delta}} B\left(\frac{j}{\delta} + \beta(k+1), 1 - \frac{j}{\delta}\right), \quad \delta > j. \end{aligned}$$

This follows from the well known definition of the moment generating function given by $M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} f(x; \mathbf{\Theta}) dx$. Since $\sum_{j=0}^{\infty} \frac{t^j}{j!} x^j f(x)$ converges and each term is integrable for all t close to 0, then we can rewrite the moment generating function as

$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} E(X^j)$ by replacing $E(X^j)$ by the right side of Equation (14) to obtain the desired result.

3.4. Conditional Moments

For income and lifetime distributions, it is of interest to obtain the conditional moments and mean residual life function. The r^{th} conditional moment for DP distribution is given by

$$\begin{aligned} E(X^r|X > t) &= \frac{1}{\overline{F}_{DP}(t)} \int_t^{\infty} x^r f_{DP}(x) dx \\ &= \frac{1}{\overline{F}_{DP}(t)} \sum_{k=0}^{\infty} \omega(k,\theta) \bigg[B\bigg(\beta(k+1) + \frac{r}{\delta}, 1 - \frac{r}{\delta}\bigg) \\ &- B_{t(a)}\bigg(\beta(j+1) + \frac{r}{\delta}, 1 - \frac{r}{\delta}\bigg)\bigg], \end{aligned} \quad (15)$$

where $t(a) = (1 + \lambda a^{-\delta})^{-1}$, $\delta > r$, and $B_{t(a)}(c,d) = \int_0^{t(a)} u^{c-1}(1-u)^{d-1} du$. The mean residual life function is $E(X|X > t) - t$.

4. Mean and Median Deviations, Bonferroni and Lorenz Curves

In this section, mean and median deviations, as well as Bonferroni and Lorenz curves of the DP distribution are presented.

4.1. Mean and Median Deviations

The amount of dispersion in a population can be measured to some extent by the totality of deviations from the mean and the median. If X has the DP distribution, we can derive the mean deviation about the mean $\mu = E(X)$ and the mean deviation about the median M from

$$\delta_1 = \int_0^{\infty} |x - \mu| f_{DP}(x) dx \quad \text{and} \quad \delta_2 = \int_0^{\infty} |x - M| f_{DP}(x) dx,$$

respectively. The mean μ is obtained from equation (14) with $r = 1$, and the median M is given by equation (13) when $q = \frac{1}{2}$. The measure δ_1 and δ_2 can be calculated by the following relationships:

$$\delta_1 = 2\mu F_{DP}(\mu) - 2\mu + 2T(\mu) \quad \text{and} \quad \delta_2 = 2T(M) - \mu,$$

where $T(a) = \int_a^{\infty} x \cdot f_{DP}(x) dx$ follows from equation (15), that is,

$$T(a) = \sum_{k=0}^{\infty} \omega(k,\theta) \bigg[B\bigg(\beta(k+1) + \frac{1}{\delta}, 1 - \frac{1}{\delta}\bigg) - B_{t(a)}\bigg(\beta(k+1) + \frac{1}{\delta}, 1 - \frac{1}{\delta}\bigg)\bigg].$$

4.2. Bonferroni and Lorenz Curves

Bonferroni and Lorenz curves are widely used tools for analyzing and visualizing income inequality. Lorenz curve, $L(p)$ can be regarded as the proportion of total income volume accumulated by those units with income lower than or equal to the volume q, and Bonferroni curve, $B(p)$ is the scaled conditional mean curve, that is, ratio of group mean income of the population.

Let $I(a) = \int_0^a x \cdot f_{DP}(x)dx$ and $\mu = E(X)$, then Bonferroni and Lorenz curves are given by

$$B(p) = \frac{I(q)}{p\mu} \quad \text{and} \quad L(p) = \frac{I(q)}{\mu},$$

respectively, for $0 \le p \le 1$, and $q = F_{DP}^{-1}(p)$. The mean of the DP distribution is obtained from equation (14) with $r = 1$ and the quantile function is given in equation (13). Consequently,

$$I(a) = \sum_{k=0}^{\infty} \omega(k,\theta) B_{t(a)}\left(\beta(k+1) + \frac{1}{\delta}, 1 - \frac{1}{\delta}\right), \tag{16}$$

for $\delta > 1$, where $t(a) = (1 + \lambda a^{-\delta})^{-1}$, and $B_{F(x)}(c,d) = \int_0^{F(x)} t^{c-1}(1-t)^{d-1}dt$ for $0 < F(x) < 1$ is incomplete beta function.

5. Order Statistics and Rényi Entropy

In this section, the distribution of the k^{th} order statistic and Rényi entropy for the DP distribution are presented. The entropy of a random variable is a measure of variation of the uncertainty.

5.1. Order Statistics

The pdf of the k^{th} order statistics from a pdf $f(x)$ is

$$\begin{aligned} f_{k:n}(x) &= \frac{f(x)}{B(k, n-k+1)} F^{k-1}(x)[1 - F(x)]^{n-k} \\ &= k\binom{n}{k} f(x) F^{k-1}(x)[1 - F(x)]^{n-k}. \end{aligned} \tag{17}$$

We apply the series expansion

$$(1 - z)^{b-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j) j!} z^j, \tag{18}$$

for $b > 0$ and $|z| < 1$, to obtain the series expansion of the distribution of order statistics from DP distribution. Using equations (18), the pdf of the k^{th} order statistic from DP distribution is given by

$$\begin{aligned} f_{k:n}(x) &= \sum_{s=0}^{\infty} k\binom{n}{k}\frac{(-1)^s f(x)\Gamma(n-k+1)}{s!\Gamma(n-k+1-s)}[F(x)]^{s+k-1} \\ &= \sum_{s,t,w=0}^{\infty} k\binom{n}{k}\frac{(-1)^{s+t+1}\theta^w (t+1)^w \Gamma(n-k+1)\Gamma(s+k)}{(1-e^{\theta})^{s+k} s!t!(w+1)!\Gamma(n-k+1-s)\Gamma(s+k-t)} \\ &\times \lambda\beta(w+1)\delta x^{-\delta-1}(1+\lambda x^{-\delta})^{-\beta(w+1)-1}. \end{aligned}$$

That is,

$$f_{k:n}(x) = \sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\sum_{w=0}^{\infty} H(s,t,w,k)\cdot f_{D}(x;\lambda,\beta(w+1),\delta),$$

where $H(s,t,w,k) = k\binom{n}{k}\frac{(-1)^{s+t+1}\theta^w (t+1)^w \Gamma(n-k+1)\Gamma(s+k)}{(1-e^{\theta})^{s+k} s!t!(w+1)!\Gamma(n-k+1-s)\Gamma(s+k-t)}$. Thus, the pdf of the k^{th} order statistic from the DP distribution is a linear combination of Dagum pdfs with parameters $\lambda, \beta(w+1)$ and $\delta > 0$. The r^{th} moment of the distribution of the k^{th} order statistic is given by

$$E(X_{k:n}^r) = \sum_{s,t,w=0}^{\infty} H(s,t,w,k)\beta(w+1)\lambda^{\frac{r}{\delta}} B\left(\beta(w+1)+\frac{r}{\delta}, 1-\frac{r}{\delta}\right), \quad \delta > r.$$

5.2. Rényi Entropy

Rényi entropy of a distribution with pdf $f(x)$ is defined as

$$I_R(\tau) = (1-\tau)^{-1}\log\left\{\int_{\mathbb{R}} f^{\tau}(x)dx\right\}, \tau > 0, \tau \neq 1.$$

Note that by using equation (18), we have

$$f_{DP}^{\tau}(x) = \sum_{k=0}^{\infty}\frac{(\tau\theta)^k(\lambda\beta\delta\theta)^{\tau}}{k!(e^{\theta}-1)^{\tau}} x^{-\delta\tau-\tau}(1+\lambda x^{-\delta})^{-\beta\tau-\beta k-\tau}.$$

Consequently, Rényi entropy of DP distribution is given by

$$I_R(\tau) = \frac{1}{1-\tau}\log\left[\sum_{k=0}^{\infty}\frac{(\tau\theta)^k(\lambda\beta\delta\theta)^{\tau}}{k!(e^{\theta}-1)^{\tau}}\frac{\lambda^{1/\delta}}{\delta}B\left(\beta(k+\tau)-\frac{\tau-1}{\delta}, \tau+\frac{\tau-1}{\delta}\right)\right].$$

for $\beta(k+\tau)-\frac{\tau-1}{\delta} > 0$ and $\tau+\frac{\tau-1}{\delta} > 0$. Rényi entropy for the sub-models can be readily obtained.

6. Maximum Likelihood Estimators

In this section, we consider the maximum likelihood estimators (MLE's) of the parameters of the DP distribution. Let $x_1, ..., x_n$ be a random sample of size n from DP distribution and $\mathbf{\Theta} = (\lambda, \delta, \beta, \theta)^T$ be the parameter vector. The log-likelihood function can be written as

$$\begin{aligned} L &= n\log\theta + n\log\beta + n\log\delta + n\log\lambda - (\delta+1)\sum_{i=1}^{n}\log x_i \\ &- (\beta+1)\sum_{i=1}^{n}\log(1+\lambda x_i^{-\delta}) + \theta\sum_{i=1}^{n}(1+\lambda x_i^{-\delta})^{-\beta}. \end{aligned} \tag{19}$$

The associated score function is given by

$$U_n(\mathbf{\Theta}) = \left(\frac{\partial L}{\partial\lambda}, \frac{\partial L}{\partial\delta}, \frac{\partial L}{\partial\beta}, \frac{\partial L}{\partial\theta}\right)^T,$$

where

$$\begin{aligned} \frac{\partial L}{\partial\lambda} &= \frac{n}{\lambda} - (\beta+1)\sum_{i=1}^{n}\frac{x_i^{-\delta}}{(1+\lambda x_i^{-\delta})} - \sum_{i=1}^{n}\beta\theta x_i^{-\delta}(1+\lambda x_i^{-\delta})^{-\beta-1}, \\ \frac{\partial L}{\partial\delta} &= \frac{n}{\delta} - \sum_{i=1}^{n}\log x_i + (\beta+1)\sum_{i=1}^{n}\frac{\lambda x_i^{\delta}\log x_i}{(1+\lambda x_i^{-\delta})} \\ &+ \sum_{i=1}^{n}\theta\lambda\beta(1+\lambda x_i^{-\delta})^{-\beta-1}x_i^{-\delta}\log x_i, \\ \frac{\partial L}{\partial\beta} &= \frac{n}{\beta} - \sum_{i=1}^{n}\log(1+\lambda x_i^{-\delta}) + \sum_{i=1}^{n}\theta(1+\lambda x_i^{-\delta})^{-\beta}\log(1+\lambda x_i^{-\delta}), \quad \text{and} \\ \frac{\partial L}{\partial\theta} &= \frac{n}{\theta} + \sum_{i=1}^{n}(1+\lambda x_i^{-\delta})^{-\beta}. \end{aligned}$$

The maximum likelihood estimate (MLE) of $\mathbf{\Theta}$, say $\widehat{\mathbf{\Theta}}$, is obtained by solving the nonlinear system $U_n(\mathbf{\Theta}) = 0$. The solution of this nonlinear system of equations is not in a closed form. These equations cannot be solved analytically, and statistical software can be used to solve them numerically via iterative methods. We can use iterative techniques such as a Newton-Raphson type algorithm to obtain the estimate $\widehat{\mathbf{\Theta}}$.

6.1. Fisher Information Matrix

In this subsection, we present a measure for the amount of information. This information can be used to obtain bounds on the variance of estimators, approximate the sampling distribution of an estimator and obtain an approximate confidence interval in the case of a large sample.

Let X be a random variable with the DP pdf $f_{DP}(x;\mathbf{\Theta})$, where $\mathbf{\Theta} = (\theta_1, \theta_2, \theta_3, \theta_4)^T = (\lambda, \delta, \beta, \theta)^T$. Then, Fisher information matrix (FIM) is the 4×4 symmetric matrix with elements:

$$\mathbf{I}_{ij}(\mathbf{\Theta}) = E_{\mathbf{\Theta}}\left[\frac{\partial \log(f_{DP}(X;\mathbf{\Theta}))}{\partial \theta_i}\frac{\partial \log(f_{DP}(X;\mathbf{\Theta}))}{\partial \theta_j}\right].$$

If the density $f_{DP}(x;\mathbf{\Theta})$ has a second derivative for all i and j, then an alternative expression for $\mathbf{I}_{ij}(\mathbf{\Theta})$ is

$$\mathbf{I}_{ij}(\mathbf{\Theta}) = -E_{\mathbf{\Theta}}\left[\frac{\partial^2 \log(f_{DP}(X;\mathbf{\Theta}))}{\partial \theta_i \partial \theta_j}\right].$$

For the DP distribution, all second derivatives exist; therefore, the formula above is appropriate and most importantly significantly simplifies the computations. Elements of the FIM can be numerically obtained by R or MATLAB. The total FIM $\mathbf{I}_n(\mathbf{\Theta})$ can be approximated by

$$\mathbf{J}_n(\hat{\mathbf{\Theta}}) \approx \left[-\left.\frac{\partial^2 \log L}{\partial \theta_i \partial \theta_j}\right|_{\mathbf{\Theta}=\hat{\mathbf{\Theta}}}\right]_{4\times 4}. \tag{20}$$

For real data, the matrix given in Equation (20) is obtained after the convergence of the Newton-Raphson procedure in MATLAB or R software.

6.2. Asymptotic Confidence Intervals

In this subsection, we present the asymptotic confidence intervals for the parameters of the DP distribution. The expectations in the Fisher Information Matrix (FIM) can be obtained numerically. Let $\hat{\mathbf{\Theta}} = (\hat{\lambda}, \hat{\delta}, \hat{\beta}, \hat{\theta})$ be the maximum likelihood estimate of $\mathbf{\Theta} = (\lambda, \delta, \beta, \theta)$. Under the usual regularity conditions and that the parameters are in the interior of the parameter space, but not on the boundary, we have: $\sqrt{n}(\hat{\mathbf{\Theta}} - \mathbf{\Theta}) \xrightarrow{d} N_4(\underline{\mathbf{0}}, \mathbf{I}^{-1}(\mathbf{\Theta}))$, where $\mathbf{I}(\mathbf{\Theta})$ is the expected Fisher information matrix. The asymptotic behavior is still valid if $\mathbf{I}(\mathbf{\Theta})$ is replaced by the observed information matrix evaluated at $\hat{\mathbf{\Theta}}$, that is $\mathbf{J}(\hat{\mathbf{\Theta}})$. The multivariate normal distribution with mean vector $\underline{0} = (0, 0, 0, 0)^T$ and covariance matrix $\mathbf{I^{-1}}(\mathbf{\Theta})$ can be used to construct confidence intervals for the model parameters. That is, the approximate $100(1-\eta)\%$ two-sided confidence intervals for λ, δ, β and θ are given by

$$\hat{\lambda} \pm Z_{\eta/2}\sqrt{\mathbf{I}^{-1}_{\lambda\lambda}(\hat{\mathbf{\Theta}})}, \quad \hat{\delta} \pm Z_{\eta/2}\sqrt{\mathbf{I}^{-1}_{\delta\delta}(\hat{\mathbf{\Theta}})}, \quad \hat{\beta} \pm Z_{\eta/2}\sqrt{\mathbf{I}^{-1}_{\beta\beta}(\hat{\mathbf{\Theta}})},$$

and $\hat{\theta} \pm Z_{\eta/2}\sqrt{\mathbf{I}^{-1}_{\theta\theta}(\hat{\mathbf{\Theta}})}$, respectively, where $\mathbf{I}^{-1}_{\lambda\lambda}(\hat{\mathbf{\Theta}}), \mathbf{I}^{-1}_{\delta\delta}(\hat{\mathbf{\Theta}}), \mathbf{I}^{-1}_{\beta\beta}(\hat{\mathbf{\Theta}})$, and $\mathbf{I}^{-1}_{\theta\theta}(\hat{\mathbf{\Theta}})$ are diagonal elements of $\mathbf{I}_n^{-1}(\hat{\mathbf{\Theta}}) = (n\mathbf{I}\hat{\mathbf{\Theta}}))^{-1}$ and $Z_{\eta/2}$ is the upper $(\eta/2)^{th}$ percentile of a standard normal distribution.

We can use the likelihood ratio (LR) test to compare the fit of the DP distribution with its sub-models for a given data set. For example, to test $\theta = 0$, the LR statistic is $\omega = 2[\ln(L(\hat{\lambda}, \hat{\delta}, \hat{\beta}, \hat{\theta})) - \ln(L(\tilde{\lambda}, \tilde{\delta}, \tilde{\beta}, 0))]$, where $\hat{\lambda}$, $\hat{\delta}$, $\hat{\beta}$, and $\hat{\theta}$, are the unrestricted estimates, and $\tilde{\lambda}$, $\tilde{\delta}$, and $\tilde{\beta}$ are the restricted estimates. The LR test rejects the null hypothesis if $\omega > \chi^2_\epsilon$, where χ^2_ϵ denote the upper $100\epsilon\%$ point of the χ^2 distribution with 1 degrees of freedom.

7. Monte Carlo Simulation Study

In this section, a simulation study is conducted to assess the performance and examine the mean estimate, average bias, root mean square error of the maximum likelihood estimators and width of the confidence interval for each parameter. We study the performance of the DP distribution by conducting various simulations for different sample sizes and different parameter values. Equation (13) is used to generate random data from the DP distribution. The simulation study is repeated for $N = 5,000$ times each with sample size $n = 25, 50, 75, 100, 200, 400, 800$ and parameter values $I : \lambda = 3.0, \delta = 1.2, \beta = 0.2, \theta = 0.3$ and $II : \lambda = 3.5, \delta = 1.0, \beta = 0.2, \theta = 0.4$. Simulation results for various other parameters sets including when the shape parameters $\delta < 1$ and $\beta > 1$ are available upon request from the authors. An R algorithm for the simulations is given in the appendix of this paper. Five quantities are computed in this simulation study.

(a) Mean estimate of the MLE $\hat{\vartheta}$ of the parameter $\vartheta = \lambda, \delta, \beta, \theta$:

$$\frac{1}{N}\sum_{i=1}^{N}\hat{\vartheta}.$$

(b) Average bias of the MLE $\hat{\vartheta}$ of the parameter $\vartheta = \lambda, \delta, \beta, \theta$:

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\vartheta} - \vartheta).$$

(c) Root mean squared error (RMSE) of the MLE $\hat{\vartheta}$ of the parameter $\vartheta = \lambda, \delta, \beta, \theta$:

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\vartheta} - \vartheta)^2}.$$

(d) Coverage probability (CP) of 95% confidence intervals of the parameter $\vartheta = \lambda, \delta, \beta, \theta$, i.e., the percentage of intervals that contain the true value of parameter ϑ.

(e) Average width (AW) of 95% confidence intervals of the parameter $\vartheta = \lambda, \delta, \beta, \theta$.

Table 5 presents the Average Bias, RMSE, CP and AW values of the parameters λ, δ, β and θ for different sample sizes. From the results, we can verify that as the sample size n increases, the RMSEs decay toward zero. The average biases for the parameter θ are all positive and slightly larger for small to moderate sample sizes but tend to get smaller as the sample size n increases. We also observe that for all the parametric values, the biases decrease as the sample size n increases. Also, the table shows that the coverage probabilities of the confidence intervals are quite close to the nominal level of 95% and that the average confidence widths decrease as the sample size increases. Consequently, the MLE's and their asymptotic results can be used for estimating and constructing confidence intervals even for reasonably small sample sizes.

Table 5: Monte Carlo Simulation Results: Average Bias, RMSE, CP and AW

		I					II				
Parameter	n	Mean	Average Bias	RMSE	CP	AW	Mean	Average Bias	RMSE	CP	AW
λ	25	4.8530	1.8530	7.2589	0.8132	44.1472	5.9531	2.4531	9.1470	0.8146	55.0576
	50	4.2227	1.2227	5.6644	0.8214	27.9452	5.4153	1.9153	7.4619	0.8358	37.5203
	75	3.9258	0.9258	4.8932	0.8286	21.7068	4.9140	1.4140	6.3977	0.8372	27.8844
	100	3.5974	0.5974	4.1241	0.8278	16.7771	4.5233	1.0233	5.2756	0.8378	22.3011
	200	3.0985	0.0985	2.5313	0.8330	10.0894	3.7666	0.2666	3.0828	0.8480	12.6877
	400	2.8628	-0.1372	1.4204	0.8570	6.5562	3.5132	0.0132	2.0123	0.8640	8.3644
	800	2.8558	-0.1442	1.0060	0.8882	4.6937	3.4562	-0.0438	1.2749	0.8954	5.8443
δ	25	1.4116	0.2116	0.5876	0.9606	2.5486	1.1663	0.1663	0.4842	0.9676	2.0609
	50	1.3148	0.1148	0.4152	0.9628	1.7464	1.1036	0.1036	0.3433	0.9642	1.4579
	75	1.2825	0.0825	0.3494	0.9628	1.4374	1.0681	0.0681	0.2795	0.9622	1.1698
	100	1.2580	0.0580	0.2989	0.9592	1.2210	1.0498	0.0498	0.2477	0.9594	1.0090
	200	1.2143	0.0143	0.2085	0.9456	0.8563	1.0140	0.0140	0.1718	0.9528	0.7063
	400	1.1944	-0.0056	0.1468	0.9468	0.6080	0.9973	-0.0027	0.1226	0.9468	0.5021
	800	1.1909	-0.0091	0.1052	0.9488	0.4356	0.9967	-0.0033	0.0878	0.9488	0.3607
β	25	0.1600	-0.0400	0.1253	0.8940	0.5016	0.1635	-0.0365	0.1425	0.9132	0.5677
	50	0.1558	-0.0442	0.0803	0.8822	0.3540	0.1565	-0.0435	0.0827	0.8914	0.3546
	75	0.1605	-0.0395	0.0702	0.8760	0.2937	0.1610	-0.0390	0.0706	0.8858	0.2981
	100	0.1646	-0.0354	0.0638	0.8836	0.2581	0.1640	-0.0360	0.0631	0.8884	0.2613
	200	0.1782	-0.0218	0.0470	0.9104	0.1858	0.1793	-0.0207	0.0464	0.9142	0.1876
	400	0.1895	-0.0105	0.0314	0.9392	0.1291	0.1900	-0.0100	0.0321	0.9414	0.1316
	800	0.1950	-0.0050	0.0219	0.9452	0.0902	0.1953	-0.0047	0.0224	0.9472	0.0914
θ	25	1.7827	1.4827	1.9319	0.9788	10.2716	1.8219	1.4219	1.8842	0.9794	10.4265
	50	1.6338	1.3338	1.8923	0.9788	8.4984	1.6645	1.2645	1.8594	0.9824	8.5670
	75	1.4764	1.1764	1.8148	0.9810	7.2930	1.5348	1.1348	1.7991	0.9778	7.3682
	100	1.3709	1.0709	1.7215	0.9772	6.3903	1.4247	1.0247	1.6870	0.9782	6.4487
	200	1.0276	0.7276	1.3219	0.9816	4.4662	1.0617	0.6617	1.2299	0.9854	4.4305
	400	0.7053	0.4053	0.7313	0.9794	2.9341	0.7702	0.3702	0.7428	0.9798	2.9895
	800	0.5335	0.2335	0.4320	0.9754	2.0481	0.5935	0.1935	0.4372	0.9488	2.0669

8. Application

In this section, we present an example to illustrate the flexibility of the DP distribution and its sub-models for data modeling. Estimates of the parameters of DP distribution (standard error in parentheses), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), sum of squares (SS) from the probability plots, Cramér-von Mises and Andersen Darling goodness-of-fit statistics W^* and A^* are presented for the data set. The command NLP in SAS and nlm in R are used here.

The table in the example contain the goodness-of-fit statistics W^* and A^*, and are described by Chen and Balakrishnan (1995). These statistics can be used to verify which distribution fits better to the data. In general, the smaller the values of W^* and A^*, the better the fit. Let $G(x;\mathbf{\Theta})$ be the cdf, where the form of G is known but the k-dimensional parameter vector, say $\mathbf{\Theta}$ is unknown. We can obtain the statistics W^* and A^* as follows: (i) Compute $u_i = G(x_i;\hat{\mathbf{\Theta}})$, where the x_i's are in ascending order; (ii) Compute $y_i = \Phi^{-1}(u_i)$, where $\Phi(.)$ is the standard normal cdf and $\Phi^{-1}(.)$ its inverse; (iii) Compute $v_i = \Phi((y_i - \overline{y})/s_y)$, where $\overline{y} = n^{-1}\sum_{i=1}^{n} y_i$ and $s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \overline{y})^2$; (iv) Calculate $W^2 = \sum_{i=1}^{n}\{v_i - (2i-1)/(2n)\}^2 + 1/(12n)$ and $A^2 = -n - n^{-1}\sum_{i=1}^{n}\{(2i-1)\log(v_i) + (2n+1-2i)\log(1-v_i)\}$; (v) Modify W^2 into $W^* = W^2(1+0.5/n)$ and A^2 into $A^* = A^2(1+0.75/n+2.25/n^2)$. In order to compare the models, we use the criteria stated above. Note that for the value of the log-likelihood function at its maximum (ℓ_n), larger value is good and preferred, for AIC, AICC and BIC, smaller values are preferred and for the Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics, smaller values are preferred.

We also compared the DP distribution with other distributions including exponentiated Weibull-Poisson (EWP) by Mahmoudi and Sepahdar (2013) and exponentiated power Lindley Poisson (EPLP) by Pararai et al. (2014) distributions. The cdf of the EWP distribution is given by

$$F_{EWP}(x;\delta,\beta,\gamma,\theta) = \left(\frac{e^{\theta(1-e^{-(\beta x)^{\gamma}})^{\delta}} - 1}{e^{\theta} - 1}\right), \quad x > 0. \tag{21}$$

The cdf and pdf of the EPLP distribution, denoted by EPLP$(\alpha,\beta,\omega,\theta)$ are given by

$$F_{EPLP}(y;\alpha,\beta,\omega,\theta) = \frac{\exp\left\{\theta\left[1-\left(1+\frac{\beta y^{\alpha}}{\beta+1}\right)e^{-\beta y^{\alpha}}\right]^{\omega}\right\} - 1}{e^{\theta}-1}, \tag{22}$$

and

$$\begin{aligned} f_{EPLP}(y;\alpha,\beta,\omega,\theta) &= \frac{\alpha\beta^2\omega\theta}{(\beta+1)(e^{\theta}-1)}(1+y^{\alpha})y^{\alpha-1}e^{-\beta y^{\alpha}}\left[1-\left(1+\frac{\beta y^{\alpha}}{\beta+1}\right)e^{-\beta y^{\alpha}}\right]^{\omega-1} \\ &\times \exp\left\{\theta\left[1-\left(1+\frac{\beta y^{\alpha}}{\beta+1}\right)e^{-\beta y^{\alpha}}\right]^{\omega}\right\} \end{aligned} \tag{23}$$

for $x > 0, \alpha > 0, \beta > 0, \omega > 0, \theta > 0$, respectively.

Plots of the fitted densities and the histogram of the data are given in Figure 3. The probability plots (Chambers et al., 1983) are also presented in Figure 4. For the probability plot, we plotted $G_{DP}(x_{(j)};\hat{\lambda},\hat{\delta},\hat{\beta},\hat{\theta})$ against $\dfrac{j-0.375}{n+0.25}$, $j = 1, 2, \cdots, n$, where $x_{(j)}$ are the ordered values of the observed data. We also computed a measure of closeness of each plot to the diagonal line. This measure of closeness is given by the sum of squares $SS = \sum_{j=1}^{n}\left[G_{DP}(x_{(j)};\hat{\lambda},\hat{\delta},\hat{\beta},\hat{\theta}) - \left(\dfrac{j-0.375}{n+0.25}\right)\right]^2$.

Table 6: Failure times data of Kevlar 49/epoxy strands with pressure at 90%

0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.07
0.08	0.09	0.09	0.10	0.10	0.11	0.11	0.12	0.13	0.18	0.19	0.20
0.23	0.24	0.24	0.29	0.34	0.35	0.36	0.38	0.40	0.42	0.43	0.52
0.54	0.56	0.60	0.60	0.63	0.65	0.67	0.68	0.72	0.72	0.72	0.73
0.79	0.79	0.80	0.80	0.83	0.85	0.90	0.92	0.95	0.99	1.00	1.01
1.02	1.03	1.05	1.10	1.10	1.11	1.15	1.18	1.20	1.29	1.31	1.33
1.34	1.40	1.43	1.45	1.50	1.51	1.52	1.53	1.54	1.54	1.55	1.58
1.60	1.63	1.64	1.80	1.80	1.81	2.02	2.05	2.14	2.17	2.33	3.03
3.03	3.34	4.20	4.69	7.89							

8.1. Kevlar 49/Epoxy Strands Failure Times

This data set consists of 101 observations corresponding to the failure times of Kevlar 49/epoxy strands with pressure at 90%. The failure times in hours were originally given in Barlow et al. (1984), Andrews and Herzberg (2012) and analyzed by Cooray and Ananda (2008). The data is presented in Table 6.

Estimates of the parameters of the distributions, standard errors (in parentheses), −2log-likelihood statistic, Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC) and Bayesian Information Criterion (BIC) are given in the table below. The results obtained from fitting the DP distribution and its sub-models FP and BIIIP distributions and other alternatives including EWP and EPLP distributions are presented for the data set. We also use the LR test to compare the DP distribution and its sub-models.

The estimated variance-covariance matrix for the DP distribution is given by:

$$\begin{pmatrix} 46.2099 & 3.8087 & -0.1424 & -4.3109 \\ 3.8087 & 0.4739 & -0.0268 & -0.1673 \\ -0.1424 & -0.0268 & 0.0036 & -0.0296 \\ -4.3109 & -0.1673 & -0.0296 & 1.1016 \end{pmatrix}$$

and the 95% two-sided asymptotic confidence intervals for λ, δ, β and θ are given by $7.2834 \pm 13.3269, 3.3711 \pm 1.3493, 0.2015 \pm 0.11799$, and 0.3744 ± 2.0572, respectively.

Plots of the fitted densities and the histogram, observed probability vs predicted probability are given in Figures 3 and 4, respectively.

The LR test statistics of the hypothesis $H_0 : FP$ vs $H_a : DP$ and $H_0 : BIIIP$ vs $H_a : DP$ are 61.31 (p-value < 0.0001) and 7.40 (p-value $= 0.0065$). The DP distribution is significantly better than FP and BIIIP distributions. The DP distribution is also significantly better than Fisk and Burr-III distributions. There is no difference between DP and Dagum distribution based on the LR test, however the goodness of fit statistics W^*

Table 7: Estimates of Models for Kevlar Strands Failure Times Data

	Estimates				Statistics						
Model	λ	δ	β	θ	$-2\log L$	AIC	$AICC$	BIC	W^*	A^*	SS
DP	7.2834	3.3711	0.2015	0.3744	200.09	208.09	208.51	218.55	0.0657	0.4632	0.0699
	(6.7978)	(0.6884)	(0.0602)	(1.0496)							
FP	0.0108	0.6444	1.0000	39.4781	261.40	267.40	267.82	275.24	1.1126	6.0129	1.0084
	(0.0015)	(0.0455)	-	(4.09E-05)							
BIIIP	1.0000	2.1278	0.3085	1.7423	207.48	213.48	213.90	221.33	0.2039	1.1841	0.2319
	-	(0.2470)	(0.0920)	(0.8605)							
Dagum	9.0357	3.4267	0.2110		200.21	206.21	206.46	214.06	0.0742	0.4980	0.0809
	(6.9969)	(0.6904)	(0.0548)								
Fisk	0.6240	1.2705	1.0000		225.37	229.37	229.49	234.60	0.5654	3.0709	0.3893
	(0.0850)	(0.1069)	-								
	c	k									
Burr III	1.1737	1.6327			217.10	221.10	221.22	226.33	0.4401	2.3866	0.4741
	(0.0983)	(0.1637)									
	α	β	ω	θ							
EPLP	0.7894	1.7952	0.9385	1.1684	204.44	212.44	212.86	222.90	0.1349	0.8141	0.1292
	(0.2025)	(0.6112)	(0.3829)	(1.2585)							
	δ	β	γ	θ							
EWP	0.8588	1.3030	0.8717	1.2662	204.62	212.62	213.03	223.08	0.1408	0.8415	0.1347
	(0.3679)	(0.7394)	(0.2408)	(1.2007)							

and A^* and the SS statistic from the probability plots clearly show that DP distribution is better than Dagum distribution for the data. We also compared the DP distribution to non-nested EWP and EPLP distributions using the AIC, AICC, BIC, W^*, A^* and SS statistics. The model with the smallest value for each of the statistics will be the best one to be used in fitting the data. When the DP distribution is compared to the non-nested EWP and EPLP distributions, it is clear that it is superior based on the AIC, AICC and BIC values. The DP distribution has the smallest goodness of fit statistic W^* and A^* values as well as the smallest SS value among all the models that were fitted. Clearly, the DP model has points closer to the diagonal line corresponding to the smallest SS value for the probability plots when compared to the non-nested EWP and EPLP distributions. Hence, the DP distribution is the "best" fit for the data when compared to all the other models that were considered.

9. Conclusion

A new class of generalized Dagum distribution called the Dagum-Poisson (DP) distribution is proposed and studied. The DP distribution has the Fisk or log-logistic Poisson, Fisk or log-logistic, Burr III-Poisson, Burr III and Dagum distributions as special cases. The DP distribution is flexible for modeling various types of lifetime and reliability data. We also obtain closed form expressions for the moments, conditional moments, mean deviations, Lorenz and Bonferroni curves, distribution of order statistics and Rényi entropy. Maximum likelihood estimation technique was used to estimate the model parameters.

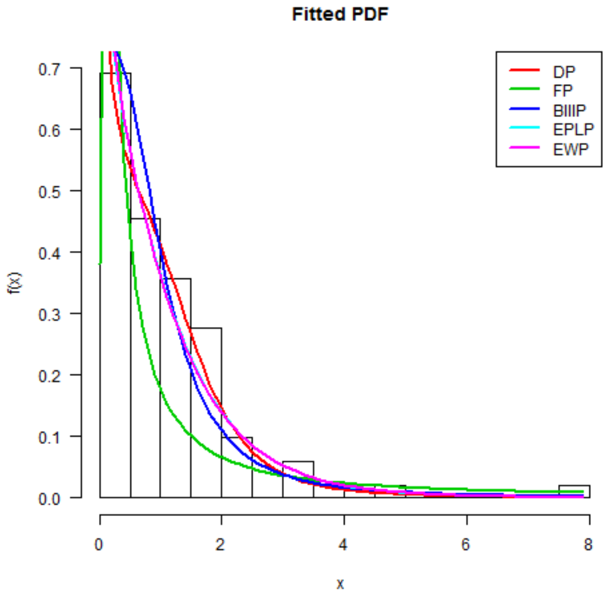

Figure 7: Fitted Densities for Kevlar Strands Failure Times Data

A simulation study was conducted to examine the bias and mean square error of the maximum likelihood estimators. Finally, the DP model was fitted to a real data set in order to illustrate the applicability and usefulness of the distribution.

References

Adamidis, K., Dimitrakopoulou, T., and Loukas, S. (2005). On an extension of the exponential-geometric distribution. *Statistics & probability letters*, 73(3):259–269.

Adamidis, K. and Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters*, 39(1):35–42.

Andrews, D. F. and Herzberg, A. M. (2012). *Data: a collection of problems from many fields for the student and research worker*. Springer Science & Business Media.

Barlow, R., Toland, R., and Freeman, T. (1984). A bayesian analysis of stress-rupture life of kevlar 49/epoxy spherical pressure vessels. In *Proc. Conference on Applications of Statistics, Marcel Dekker, New York*.

Barreto-Souza, W. and Cribari-Neto, F. (2009). A generalization of the exponential-poisson distribution. *Statistics & Probability Letters*, 79(24):2493–2500.

Barreto-Souza, W., de Morais, A. L., and Cordeiro, G. M. (2011). The weibull-geometric distribution. *Journal of Statistical Computation and Simulation*, 81(5):645–657.

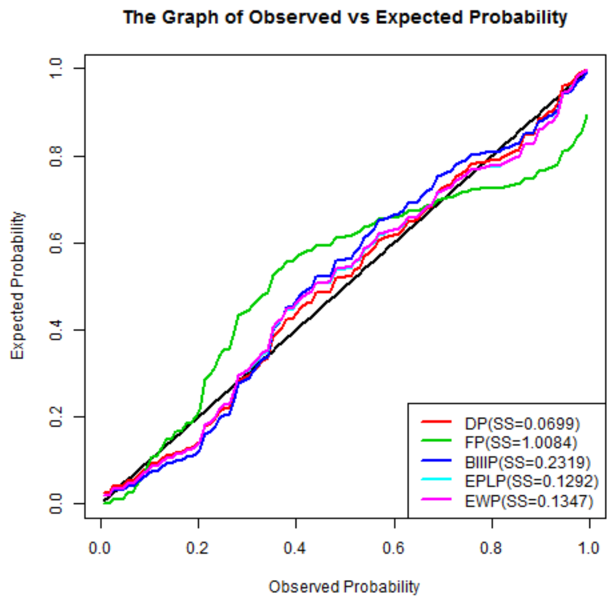

Figure 8: Probability Plots for Kevlar Strands Failure Times Data

Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics & Data Analysis*, 53(12):4433–4440.

Chambers, J., Cleveland, W., and Kleiner, B. (1983). Graphical methods for data analysis. *Wadsworth statistics/probability series.*

Chen, G. and Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, 27(2):154–161.

Cooray, K. and Ananda, M. M. (2008). A generalization of the half-normal distribution with applications to lifetime data. *Communications in StatisticsTheory and Methods*, 37(9):1323–1337.

Dagum, C. (1977). New model of personal income-distribution-specification and estimation. *Economie appliquée*, 30(3):413–437.

Domma, F. (2002). Landamento della hazard function nel modello di dagum a tre parametri. *Quaderni di Statistica*, 4:1–12.

Domma, F. and Condino, F. (2013). The beta-dagum distribution: definition and properties. *Communications in Statistics-Theory and Methods*, 42(22):4070–4090.

Domma, F., Giordano, S., and Zenga, M. (2011). Maximum likelihood estimation in dagum distribution with censored samples. *Journal of Applied Statistics*, 38(12):2971–2985.

Huang, S. and Oluyede, B. O. (2014). Exponentiated kumaraswamy-dagum distribution with applications to income and lifetime data. *Journal of Statistical Distributions and*

Applications, 1(1):1–20.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). Continuous univariate distributions, vol. 1-2.

Kleiber, C. (2008). *A Guide to the Dagum Distribution. In: Duangkamon,C. (ed.) Modeling Income Distributions and Lorenz Curves Series: Economic Studies in Inequality, Social Exclusion and Well-Being.* Springer.

Kleiber, C. and Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*, volume 470. John Wiley & Sons.

Kuş, C. (2007). A new lifetime distribution. *Computational Statistics & Data Analysis*, 51(9):4497–4509.

Lu, W. and Shi, D. (2012). A new compounding life distribution: the weibull–poisson distribution. *Journal of applied statistics*, 39(1):21–38.

Mahmoudi, E. and Sepahdar, A. (2013). Exponentiated weibull–poisson distribution: Model, properties and applications. *Mathematics and Computers in Simulation*, 92:76–97.

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica: journal of the Econometric Society*, pages 647–663.

McDonald, J. B. and Xu, Y. J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics*, 66(1):133–152.

Morais, A. L. and Barreto-Souza, W. (2011). A compound class of weibull and power series distributions. *Computational Statistics & Data Analysis*, 55(3):1410–1425.

Oluyede, B. O., Huang, S., and Pararai, M. (2014). A new class of generalized dagum distribution with applications to income and lifetime data. *Journal of Statistical and Econometric Methods*, 3(2):125–151.

Pararai, M., Warahena-Liyanage, G., and Oluyede, B. (2014). Exponentiated power lindley poisson distribution: properties and application. *Submitted to Communications in Statistics-Theory and Methods.*

Tahmasbi, R. and Rezaei, S. (2008). A two-parameter lifetime distribution with decreasing failure rate. *Computational Statistics & Data Analysis*, 52(8):3889–3901.

A. R Code: Simulations

```
library(numDeriv)
library(Matrix)

delta=1.0
beta=0.2
lambda=3.5
theta=0.4

#Define the quantile of DP
quantile=function(delta,beta,lambda,theta,u){
result <-((((log(u*(exp(theta)-1)+1))/theta)^(-1/beta)-1)/lambda)^(-1/delta)
}

DP_LL<-function(par){-sum(log(par[2]*par[3]*par[1]*par[4]
*(x^(-par[1]-1))*((1+par[3]*x^(-par[1]))^(-par[2]-1))*exp(par[4]
*(1+par[3]*x^(-par[1]))^(-par[2]))/(exp(par[4])-1)))}

n1=c(25,50,75,100,200,400,800)
# If you want to check one sample at a time then use n1<-c(sample size)

for (j in 1:length(n1)){
n=n1[j]
N=5000
mle_lambda<-c(rep(0,N))
mle_beta<-c(rep(0,N))
mle_delta<-c(rep(0,N))
mle_theta<-c(rep(0,N))

LC_lambda<-c(rep(0,N))
UC_lambda<-c(rep(0,N))
LC_beta<-c(rep(0,N))
UC_beta<-c(rep(0,N))
LC_delta<-c(rep(0,N))
UC_delta<-c(rep(0,N))
LC_theta<-c(rep(0,N))
UC_theta<-c(rep(0,N))

count_lambda=0
count_beta=0
count_delta=0
count_theta=0
```

```
temp=1
HH1<-matrix(c(rep(2,16)),nrow=4,ncol=4)
HH2<-matrix(c(rep(2,16)),nrow=4,ncol=4)
for (i in 1:N)
{
#print(i)
#flush.console()
repeat{
x<-c(rep(0,n))

#Generate a random variable from uniform distribution
u<-0
u<-runif(n,min=0,max=1)


for (k in 1:n){
x[k]<-quantile(delta,beta,lambda,theta,u[k])
}

#Maximum likelihood estimation
mle.result<-nlminb(c(delta,beta,lambda,theta),DP_LL,lower=0,upper=Inf)

temp=mle.result$convergence
if(temp==0){
temp_lambda<-mle.result$par[3]
temp_beta<-mle.result$par[2]
temp_delta<-mle.result$par[1]
temp_theta<-mle.result$par[4]

HH1<-hessian(DP_LL,c(temp_delta,temp_beta,temp_lambda,temp_theta))
if(( rcond(HH1)>1e-8) & sum(is.nan(HH1))==0 & (diag(HH1)[1]>0)\\
 & (diag(HH1)[2]>0) & (diag(HH1)[3]>0) & (diag(HH1)[4]>0) ){
HH2<-solve(HH1)
#print(det(HH1))
}
else{
temp=1}
}


if ((temp==0) & (diag(HH2)[1]>0) & (diag(HH2)[2]>0)\\
 & (diag(HH2)[3]>0) & (diag(HH2)[4]>0)  & (sum(is.nan(HH2))==0)){
break
```

```
}
}
#print(temp)
temp=1
mle_lambda[i]<-mle.result$par[3]
mle_beta[i]<-mle.result$par[2]
mle_delta[i]<-mle.result$par[1]
mle_theta[i]<-mle.result$par[4]

HH<-hessian(DP_LL,c(mle_delta[i],mle_beta[i],mle_lambda[i],mle_theta[i]))
H<-solve(HH)
LC_lambda[i]<-mle_lambda[i]-1.96*sqrt(diag(H)[3])
UC_lambda[i]<-mle_lambda[i]+1.96*sqrt(diag(H)[3])
if( (LC_lambda[i]<=lambda) & (lambda<=UC_lambda[i])){
count_lambda=count_lambda+1
}

LC_beta[i]<-mle_beta[i]-1.96*sqrt(diag(H)[2])
UC_beta[i]<-mle_beta[i]+1.96*sqrt(diag(H)[2])
if( (LC_beta[i]<=beta) & (beta<=UC_beta[i])){
count_beta=count_beta+1
}

LC_delta[i]<-mle_delta[i]-1.96*sqrt(diag(H)[1])
UC_delta[i]<-mle_delta[i]+1.96*sqrt(diag(H)[1])
if( (LC_delta[i]<=delta) & (delta<=UC_delta[i])){
count_delta=count_delta+1
}

LC_theta[i]<-mle_theta[i]-1.96*sqrt(diag(H)[4])
UC_theta[i]<-mle_theta[i]+1.96*sqrt(diag(H)[4])
if( (LC_theta[i]<=theta) & (theta<=UC_theta[i])){
count_theta=count_theta+1
}
}

#Calculate Average Bias
Bias_lambda<-sum(mle_lambda-lambda)/N
Bias_beta<-sum(mle_beta-beta)/N
Bias_delta<-sum(mle_delta-delta)/N
Bias_theta<-sum(mle_theta-theta)/N

print(cbind(Bias_lambda,Bias_beta,Bias_delta,Bias_theta))
```

```
#Calculate RMSE
RMSE_lambda<-sqrt(sum((lambda-mle_lambda)^2)/N)
RMSE_beta<-sqrt(sum((beta-mle_beta)^2)/N)
RMSE_delta<-sqrt(sum((delta-mle_delta)^2)/N)
RMSE_theta<-sqrt(sum((theta-mle_theta)^2)/N)
print(cbind(RMSE_lambda,RMSE_beta,RMSE_delta,RMSE_theta))

#Converge Probability
CP_lambda<-count_lambda/N
CP_beta<-count_beta/N
CP_delta<-count_delta/N
CP_theta<-count_theta/N
print(cbind(CP_lambda,CP_beta,CP_delta,CP_theta))

#Average Width
AW_lambda<-sum(abs(UC_lambda-LC_lambda))/N
AW_beta<-sum(abs(UC_beta-LC_beta))/N
AW_delta<-sum(abs(UC_delta-LC_delta))/N
AW_theta<-sum(abs(UC_theta-LC_theta))/N
print(cbind(AW_lambda,AW_beta,AW_delta,AW_theta))
}
```