

Dagum Distribution: Properties and Different Methods of Estimation

Sanku Dey¹, Bander Al-Zahrani² & Samerah Basloom²

¹ Department of Statistics, St. Anthony's College Shillong-793001, Meghalaya, India

² Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia

Correspondence: Bander Al-Zahrani, Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia. E-mail: bmalzahrani@kau.edu.sa

Received: November 24, 2016 Accepted: December 29, 2016 Online Published: February 21, 2017

doi:10.5539/ijsp.v6n2p74

URL: <https://doi.org/10.5539/ijsp.v6n2p74>

Abstract

This article addresses the various properties and different methods of estimation of the unknown parameters of a three-parameter Dagum distribution from the frequentist point of view. Although, our main focus is on estimation from frequentist point of view, yet, various mathematical and statistical properties of the Dagum distribution (such as quantiles, moments, moment generating function, hazard rate, mean residual lifetime, mean past lifetime, mean deviation about mean and median, various entropies, Bonferroni and Lorenz curves and order statistics) are derived. We briefly describe different frequentist approaches, namely, maximum likelihood estimators, moments estimators, L-moment estimators, percentile based estimators, least squares estimators, maximum product of spacings estimators, minimum distances estimators, Cramér-von-Mises estimators, Anderson-Darling and right-tail Anderson-Darling estimators and compare them using extensive numerical simulations. Monte Carlo simulations are performed to compare the performances of the proposed methods of estimation for both small and large samples. Finally, a real data set have been analyzed for illustrative purposes.

Keywords: dagum distribution, maximum likelihood estimators; method of moment estimators; least squares estimators; weighted least squares estimators; percentile estimators; method of maximum product spacing.

1. Introduction

Dagum distribution was introduced by Dagum (Dagum, C., 1977) for modeling personal income data as an alternative to the Pareto and log-normal models. This distribution has been extensively used in various fields such as, income and wealth data, meteorological data, reliability and survival analysis. The Dagum distribution is also known as the inverse Burr XII distribution, especially in the actuarial literature. An important characteristic of Dagum distribution is that its hazard function can be monotonically decreasing, an upside-down bathtub, or bathtub and then upside-down bathtub shaped, for details see Domma (Domma, F., 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 (Domma, et al., 2011; Domma, et al., 2013). Kleiber and Kotz (Kotz, S., 2003) and Kleiber (Kleiber, C., 2008) provided an exhaustive review on the origin of the Dagum model and its applications. Domma et al. (Domma, et al., 2011) estimated the parameters of Dagum distribution with censored samples. Shahzad and Asghar (Shahzad, M. N., & Asghar, Z., 2013) used TL-moments to estimate the parameter of this distribution. Oluyede and Ye (Oluyede, B. O., & Ye, Y., 2014) presented the class of weighted Dagum and related distributions. Domma and Condino (Domma, F., & Condino, F., 2013) proposed the five parameter beta-Dagum distribution.

A continuous random variable T is said to have a three-parameter Dagum distribution, abbreviated as $T \sim \text{Dag}(\beta, \lambda, \delta)$, if its density probability function (pdf) is given as

$$f(t; \beta, \lambda, \delta) = \beta \lambda \delta t^{-\delta-1} (1 + \lambda t^{-\delta})^{-\beta-1}, \quad t > 0, \quad (1)$$

where $\lambda > 0$ is the scale parameter and its two shape parameters β and δ are both positive.

The corresponding distribution function of (1) is given by

$$F(t; \beta, \lambda, \delta) = (1 + \lambda t^{-\delta})^{-\beta}, \quad t > 0, \beta, \lambda, \delta > 0. \quad (2)$$

The main aim of this paper is to consider different estimation methods and study how the estimators behave for different sample sizes and for different parameter values. We mainly compare; maximum likelihood estimators, moments estimators, L-moment estimators, percentile based estimators, least squares estimators, maximum product of spacings estimators, minimum distances estimators, Cramér-von-Mises estimators, Anderson-Darling and right-tail Anderson-Darling estimators.

The maximum likelihood estimation (MLE) and the method of moments estimation (MME) are traditional methods of estimation. Although MLE is advantageous in terms of its efficiency and its nice theoretical properties, there is evidence that it does not perform well, especially, for small samples. The method of moments is easily applicable and often gives explicit forms for estimators of unknown parameters. There are, however, cases where the method of moments does not give explicit estimators e.g., for the parameters of the Weibull and Gompertz distributions. Therefore, other methods have been proposed in the literature as alternatives to the traditional methods of estimation. Among them, the L-moments estimator (LME), the least squares estimator (LSE), the generalized spacing estimator (GSE) and the percentile estimator (PE) are often suggested. Generally, these methods do not have good theoretical properties but in some cases they can provide better estimates of the unknown parameters than the MLE and the MME.

The appeal of the methods of estimation varies from user to user and according to the area of application. This paper considers ten different frequentist estimators for the Dagum distribution and evaluates their performance for different sample sizes and different parameter values. Simulations are used to compare the performance as it is not possible to compare all estimators theoretically, see Gupta and Kundu (Gupta, R. D. & Kundu, D., 2001; Gupta, R. D. & Kundu, D., 2007). Comparisons of estimation methods for other distributions have been performed in the literature, for example (Kundu, D., & Raqab, M. Z., 2005; Alkawasbeh, M. R., & Raqab, M. Z., 2009; Dey, S., Dey, T., & Kundu, D., 2014; Teimouri, M., et al., 2013).

The paper is organized as follows. In Section 2, we study the mathematical and statistical properties of the distribution. Section 3 deals with parameter estimation; simulation and real data application are presented in Section 4. The paper ends with a brief conclusion in Section 5.

2. Some Mathematical and Statistical Properties

In this section, we provide some important mathematical and statistical properties of Dagum Distribution like quantiles, moments, moment generating function, hazard rate and mean residual life functions, conditional moments, mean deviation, Bonferroni and Lorenz curves, Rényi and Shannon entropy.

2.1 Shape of Pdf

The study of shapes is useful to determine if a data set can be modeled by the $Dag(\beta, \lambda, \delta)$. The limit of Dagum density as $x \rightarrow \infty$ is 0 and the limit as $x \rightarrow 0$ is ∞ . The following theorem gives simple conditions under which the pdf (1) is decreasing or unimodal.

Theorem 1. The pdf, $f(x)$, of $X \sim Dag(\beta, \lambda, \delta)$ is decreasing (unimodal) function if $\beta\delta \leq 1$ ($\beta\delta > 1$).

Proof. The first derivative of $\ln f(t)$ is

$$\frac{d \ln f(t)}{dt} = -\frac{(\delta + 1)}{t} + \frac{\lambda\delta(\beta + 1)t^{-(\delta+1)}}{(1 + \lambda t^{-\delta})}$$

Now, it is easy to see that for $\beta\delta \leq 1$, the function $(\ln f(x))'$ is negative, which implies that the pdf $f(x)$ is decreasing function with $f(0) = \infty$ and $f(\infty) = 0$. We consider now the case $\beta\delta > 1$, the function $(\ln f(x))'$ is increasing for $\lambda > 1/\beta\delta$ and decreasing for $\lambda \leq 1/\beta\delta$. \square

Figure 1 shows different shapes of the pdf of the Dagum distribution for various parameter specifications.

2.2 Quantile Function

Let X denote a random variable with the pdf given by (1). The quantile function, denoted by Q , is

$$Q(p) = \inf\{t \in R : F(t) \geq p\}, \text{ where } 0 < p < 1.$$

From (2), it follows that the quantile function

$$Q(p; \beta, \lambda, \delta) = \left\{ \frac{1}{\lambda} \left[p^{-\frac{1}{\beta}} - 1 \right] \right\}^{-\frac{1}{\delta}}. \quad (3)$$

The first quartile, the median and the third quartile can be obtained simply by applying (3). In particular, for $p = 0.5$ we have the median of the Dagum distribution as follows.

$$\text{Median}(T) = \left\{ \frac{1}{\lambda} \left[(0.5)^{-\frac{1}{\beta}} - 1 \right] \right\}^{-\frac{1}{\delta}}. \quad (4)$$

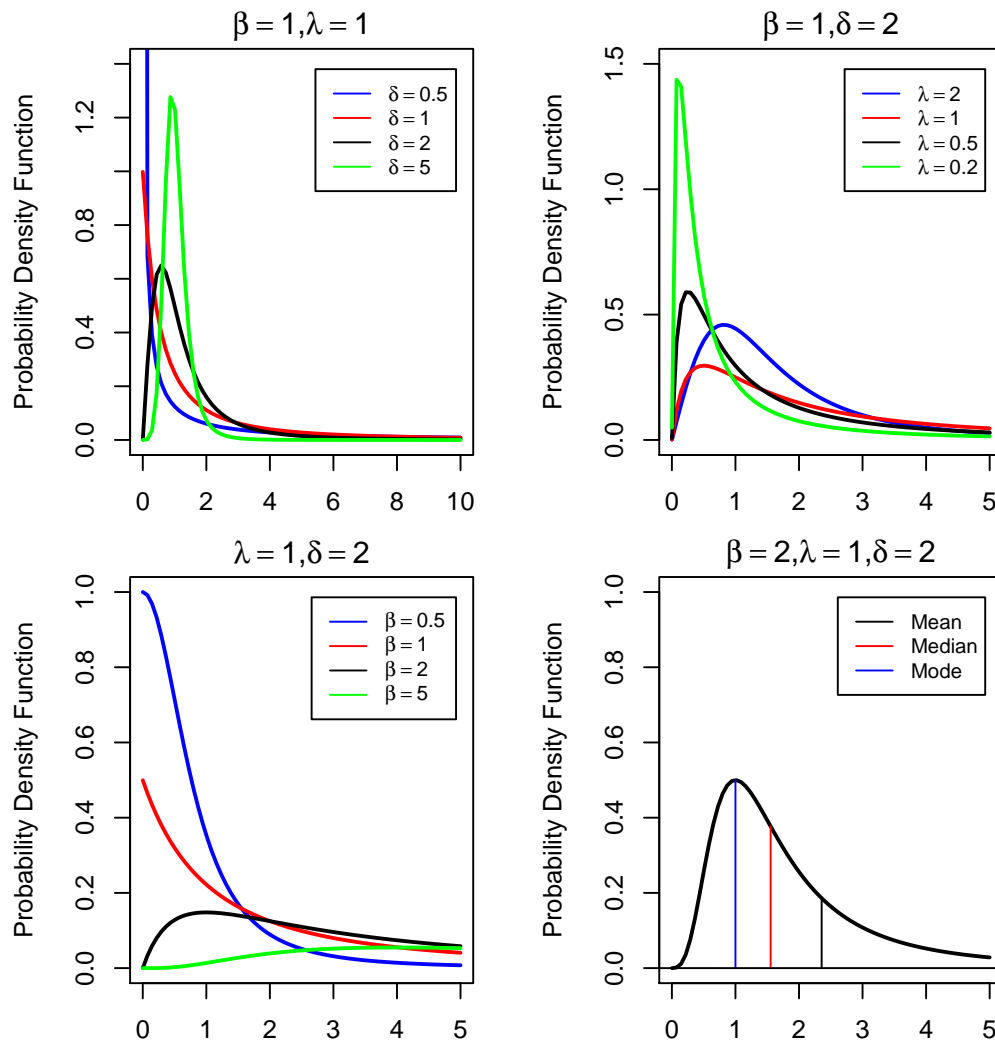


Figure 1. The pdf of the Dagum distribution for various parameter specifications

2.3 Moments

We hardly need to emphasize the necessity and importance of the moments in any statistical analysis especially in applied work. Some of the most important features and characteristics of a distribution can be studied through moments, e.g. tendency, dispersion, skewness, and kurtosis.

If the random variable T is distributed as $Dag(\beta, \lambda, \delta)$, then its k th moment around zero can be expressed as

$$E[T^k] = \beta\lambda\delta \int_0^{\infty} t^{k-\delta-1} (1 + \lambda t^{-\delta})^{-\beta-1} dt = \beta\lambda^{\frac{k}{\delta}} B\left(1 - \frac{k}{\delta}, \beta + \frac{k}{\delta}\right), \quad (5)$$

where $B(\cdot, \cdot)$ is the complete beta function. From relation (5) we can observe the mean, $\mu = E[T]$, and the variance, $\sigma^2 = V(T)$, of T as follows.

$$E[T] = \beta\lambda^{\frac{1}{\delta}} B\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right).$$

$$Var[T] = \beta\lambda^{\frac{2}{\delta}} B\left(1 - \frac{2}{\delta}, \beta + \frac{2}{\delta}\right) - \left\{\beta\lambda^{\frac{1}{\delta}} B\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right)\right\}^2.$$

2.4 Moment Generating Function

Many of the interesting characteristics and features of a distribution can be obtained via its moment generating function (mgf) and moments. Let T denote a random variable with the probability density function (1). By definition of moment generating function of T and using (1), we have

$$\begin{aligned} M_T(u) &= E(e^{ut}) = \beta\lambda\delta \int_0^\infty e^{ut} t^{-\delta-1} (1 + \lambda t^{-\delta})^{-\beta-1} dt \\ &= \beta\lambda\delta \sum_{j=0}^{\infty} \frac{t^j}{j!} \int_0^\infty t^{j-\delta-1} (1 + \lambda t^{-\delta})^{-\beta-1} dt \\ &= \lambda^{\frac{j}{\delta}} \sum_{j=0}^{\infty} \frac{t^j}{j!} \frac{\Gamma(1 - \frac{j}{\delta})\Gamma(\beta + \frac{j}{\delta})}{\Gamma(\beta)}. \end{aligned} \quad (6)$$

Consequently, the r th moment of T is

$$\mu'_r = E[T^r] = \lambda^{\frac{r}{\delta}} \frac{\Gamma(1 - \frac{r}{\delta})\Gamma(\beta + \frac{r}{\delta})}{\Gamma(\beta)}$$

The coefficient of variation(CV), Skewness(CS) and Kurtosis(CK) are, respectively, given by

$$\begin{aligned} CV &= \frac{\sigma}{\mu} = \frac{\left[c(\Gamma_2 - c\Gamma_1^2) \right]^{\frac{1}{2}}}{\Gamma_1} \\ CS &= \frac{\mu_3}{\sigma^3} = \frac{E(x^3) - 3\mu E(x^2) + 2\mu^3}{\sigma^3} = \frac{\Gamma_3 - 3c\Gamma_1\Gamma_2 + 2c^2\Gamma_1^3}{\sqrt{c(\Gamma_2 - c\Gamma_1^2)}} \\ CK &= \frac{\mu_4}{\sigma^4} = \frac{E(x^4) - 4\mu E(x^3) + 6\mu^2 E(x^2) - 3\mu^4}{\sigma^4} = \frac{\Gamma_4 - c\Gamma_1\Gamma_3 + 6c^2\Gamma_1\Gamma_2 - 3c^3\Gamma_1^4}{c^3[\Gamma_2 - c\Gamma_1^2]^4}, \end{aligned}$$

where $c = \frac{1}{\Gamma(\beta)}$ and $\Gamma_j = \lambda^{\frac{j}{\delta}} \Gamma(1 - \frac{j}{\delta}) \Gamma(\beta + \frac{j}{\delta})$.

2.5 Hazard Function

Among the basic tools for studying the ageing and reliability characteristics of a system is the hazard rate (HR) function. The HR gives the rate of failure of the system immediately after time t . Thus the hazard rate function of the Dagum distribution is given by

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{\beta\lambda\delta t^{-(\delta+1)}(1 + \lambda t^{-\delta})^{-(\beta+1)}}{1 - (1 + \lambda t^{-\delta})^{-\beta}}. \quad (7)$$

Using the Glaser's lemma (Glaser, 1980), it can be proved that if $F_X \sim \text{Dag}(\beta, \lambda, \delta)$ then F can have upside-down bathtub, decreasing, and upside-down bathtub and then bathtub shaped, see Domma (Domma, 2002).

Figure 1 gives different shapes of the hrf of the Dagum distribution for various parameter specifications.

2.6 Mean Residual Lifetime

The mean residual life (MRL) is the expected remaining life, $T - t$, given that the item has survived to time t . Thus, in life testing situations, the expected additional lifetime given that a component has survived until time t is called the MRL. Since the MRL function is the expected remaining life, t must be subtracted, yielding

$$\mu(t) = E[T - t | T > t] = \frac{1}{\bar{F}(t)} \int_t^\infty \bar{F}(u) du, \quad t \geq 0.$$

Clearly, $\mu(0) = \mu = E[X]$. The mean residual lifetime function of $T \sim \text{Dag}(\beta, \lambda, \delta)$ is given by

$$\mu(t) = \frac{\beta\lambda\delta \int_t^\infty y^{-\delta}(1 + \lambda y^{-\delta})^{-(\beta+1)} dy}{1 - (1 + \lambda t^{-\delta})^{-\beta}} - t \quad (8)$$

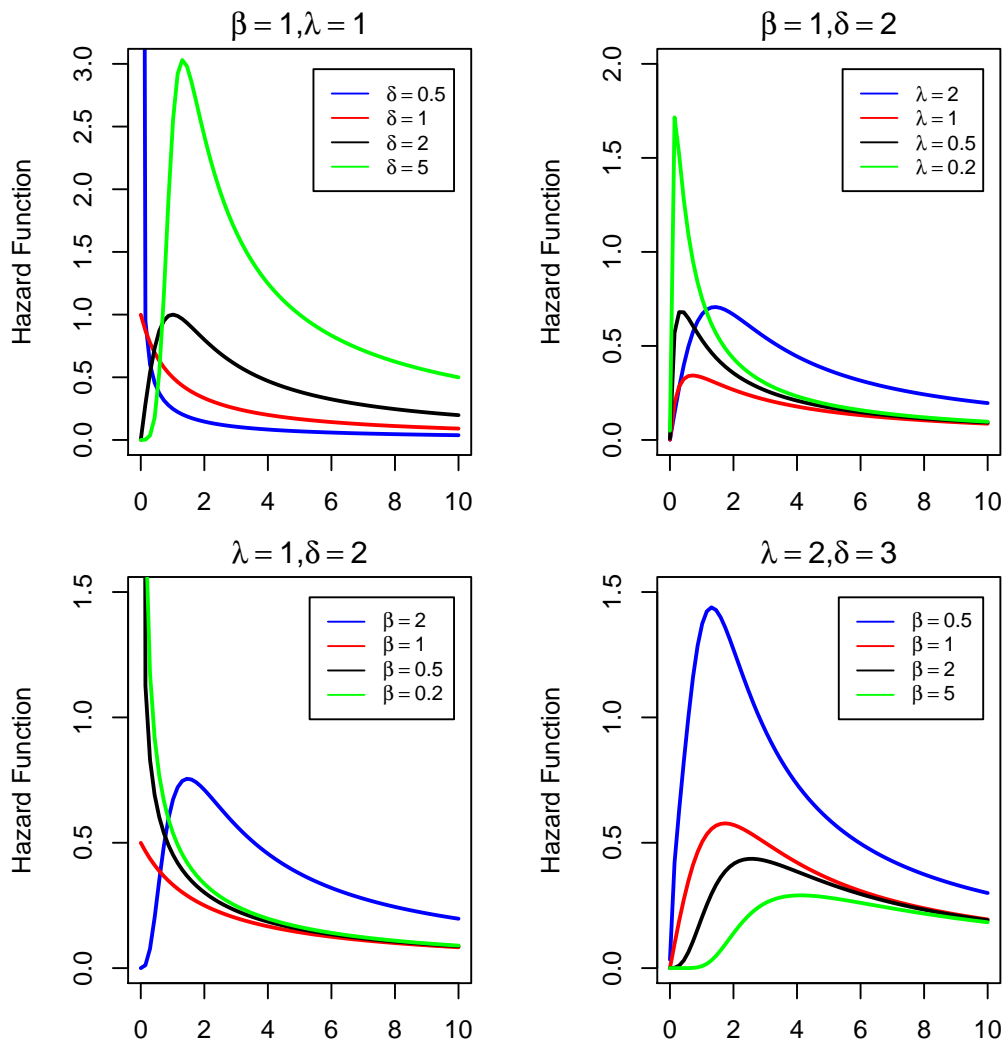


Figure 2. The hrf of the Dagum distribution for various parameter specifications

2.7 Mean Past Lifetime (MPL)

In a real life situation, where systems often are not monitored continuously, one might be interested in getting inference more about the history of the system e.g., when the individual components have failed. Assume now that a component with lifetime X has failed at or some time before t , $t \geq 0$. Consider the conditional random variable $t - T|T \leq t$. This conditional random variable shows, in fact, the time elapsed from the failure of the component given that its lifetime is less than or equal to t . Hence, the mean past lifetime (MPL) of the component can be defined as

$$\begin{aligned}
 k(t) &= E[t - T|T \leq t] = t - \frac{\int_0^t F(x)dx}{F(t)} = t - \frac{\int_0^t xf(x)dx}{F(t)} \\
 &= t - \frac{\beta\lambda^{\frac{1}{\delta}} \left[t^{-\delta}\lambda \right]^{-\frac{1}{\delta}} \sum_{n=0}^{\infty} (-1)^n \binom{\beta+n+1}{n} \frac{[t\lambda^{-\frac{1}{\delta}}]^{n+1}}{(n-\frac{1}{\delta}+1)}}{(1 + \lambda t^{-\delta})^{-\beta}}.
 \end{aligned}$$

One can easily show that $k(t) \rightarrow \infty$ as $t \rightarrow 0$.

2.8 Conditional Moments

For lifetime models, it is also of interest to know what $E(T^n|T > t)$ is. It can be easily seen that

$$E(T^n|T > t) = \frac{1}{\bar{F}(t)} \int_t^{\infty} t^n f(t)dt$$

In particular

$$\begin{aligned}
 E(T|T > t) &= \frac{1}{\bar{F}(t)} \beta \lambda^{\frac{1}{\delta}} (t^{-\delta} \lambda)^{-\frac{1}{\delta}} \sum_{n=0}^{\infty} (-1)^n \binom{\beta+n+1}{n} \frac{(t^{-\delta} \lambda)^{n+1}}{(n - \frac{1}{\delta} + 1)}. \\
 E(T^2|T > t) &= \frac{1}{\bar{F}(t)} \beta \lambda^{\frac{2}{\delta}} (t^{-\delta} \lambda)^{-\frac{2}{\delta}} \sum_{n=0}^{\infty} (-1)^n \binom{\beta+n+1}{n} \frac{(t^{-\delta} \lambda)^{n+1}}{(n - \frac{2}{\delta} + 1)}. \\
 E(T^3|T > t) &= \frac{1}{\bar{F}(t)} \beta \lambda^{\frac{3}{\delta}} (t^{-\delta} \lambda)^{-\frac{3}{\delta}} \sum_{n=0}^{\infty} (-1)^n \binom{\beta+n+1}{n} \frac{(t^{-\delta} \lambda)^{n+1}}{(n - \frac{3}{\delta} + 1)}. \\
 E(T^4|T > t) &= \frac{1}{\bar{F}(t)} \beta \lambda^{\frac{4}{\delta}} (t^{-\delta} \lambda)^{-\frac{4}{\delta}} \sum_{n=0}^{\infty} (-1)^n \binom{\beta+n+1}{n} \frac{(t^{-\delta} \lambda)^{n+1}}{(n - \frac{4}{\delta} + 1)}.
 \end{aligned}$$

The mean residual lifetime function is $E(T|T > t) - t$.

2.9 Mean Deviation

The mean deviations about the mean and the median can be used as measures of spread in a population. Let $\mu = E(T)$ and M be the mean and the median of the Dagum distribution, respectively. The mean deviations about the mean and about the median can be calculated as

$$\delta_1(T) = \int_0^{\infty} |t - \mu| f(t) dt \quad \text{and} \quad \delta_2(T) = \int_0^{\infty} |t - M| f(t) dt, \quad (9)$$

respectively.

$$\begin{aligned}
 E(|T - m|) &= \int_0^{\infty} |t - m| f(t) dt \\
 &= \int_0^m (m - t) f(t) dt + \int_m^{\infty} (t - m) f(t) dt \\
 &= 2m \int_0^m f(t) dt - m - E(T) + 2 \int_m^{\infty} t f(t) dt \\
 &= 2mF(m) + 2 \int_m^{\infty} t f(t) dt - E(T) - m
 \end{aligned}$$

Therefore, since we have $\mu = \beta \lambda^{-\frac{1}{\delta}} B(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta})$ and $M = \left[\frac{1}{\lambda} \left[(0.5)^{-\frac{1}{\beta}} - 1 \right] \right]^{-\frac{1}{\delta}}$,

$$\begin{aligned}
 \delta_1(T) &= 2\mu F(\mu) - 2\mu + 2 \int_{\mu}^{\infty} t f(t) dt \\
 &= 2\mu \left((1 + \lambda \mu^{-\delta})^{-\beta} - 1 \right) + 2\beta \lambda^{\frac{1}{\delta}} (\mu^{-\delta} \lambda)^{-\frac{1}{\delta}} \sum_{n=0}^{\infty} (-1)^n \binom{\beta+n+1}{n} \frac{(\mu^{-\delta} \lambda)^{n+1}}{(n - \frac{1}{\delta} + 1)}.
 \end{aligned}$$

$$\begin{aligned}
 \delta_2(T) &= 2MF(M) - M + 2 \int_M^{\infty} t f(t) dt - \mu \\
 &= 2M(1 + \lambda M^{-\delta})^{-\beta} - M + 2\beta \lambda^{\frac{1}{\delta}} (M^{-\delta} \lambda)^{-\frac{1}{\delta}} \sum_{n=0}^{\infty} (-1)^n \binom{\beta+n+1}{n} \frac{(M^{-\delta} \lambda)^{n+1}}{(n - \frac{1}{\delta} + 1)} - \mu.
 \end{aligned}$$

2.10 Entropies

The concept of entropy is important in different areas such as physics, probability and statistics, communication theory, economics, etc. Several measures of entropy have been studied and compared in the literature. Entropy of a random variable T is a measure of variation of the uncertainty. If T has the probability density function $f(t)$ then the Shannon entropy (see Shannon (1951)) is defined by

$$\begin{aligned}
H(t) &= -E[\ln f(t)] = -\int_0^\infty f(t) \ln f(t) dt \\
&= -\int_0^\infty f(t) \ln(\beta\lambda\delta) dt + (\delta+1) \int_0^\infty f(t) \ln(t) dt + (\beta+1) \int_0^\infty f(t) \ln(1+\lambda t^{-\delta}) dt \\
&= -\ln(\beta\lambda\delta) + (\delta+1) \int_0^\infty f(t) \ln(t) dt + (\beta+1) \int_0^\infty f(t) \ln(1+\lambda t^{-\delta}) dt \\
&= -\ln(\beta\lambda\delta) + \frac{\beta}{\delta} \left[\ln(\lambda) \sum_{k=0}^\infty (-1)^k \binom{n-k-1}{k} (\lambda t^{-\delta})^k + \left[\frac{\psi(\beta) + \gamma}{\beta} \right] \right] + \frac{2(\beta+1)\beta}{(\beta-1)^2},
\end{aligned}$$

where $\psi(\cdot)$ is the digamma function. Rényi entropy (Rényi (1961)) can be expressed as

$$\begin{aligned}
H_\theta(t) &= \frac{1}{1-\theta} \ln \left(\int_0^1 f^\theta(t) dt \right), \quad \theta > 0, \theta \neq 1. \\
&= \frac{1}{1-\theta} \ln \left[\left(\frac{(\beta\lambda\delta)^\theta \lambda^{\frac{1-\theta}{\delta}-\theta}}{\delta} \right) \int_0^1 z^{\beta\theta + \frac{1-\theta}{\delta}-1} (1-z)^{\theta - \frac{1-\theta}{\delta}-1} dz \right] \\
&= \frac{1}{1-\theta} \ln \left[\left(\frac{(\beta\lambda\delta)^\theta \lambda^{\frac{1-\theta}{\delta}-\theta}}{\delta} \right) \frac{\Gamma(\beta\theta + \frac{1-\theta}{\delta}) \Gamma(\theta - \frac{1-\theta}{\delta})}{\Gamma(\beta\theta + \theta)} \right], \tag{10}
\end{aligned}$$

where $z = (1 + \lambda t^{-\delta})^{-1}$. When $\theta \rightarrow 1$, the Rényi entropy converges to the Shannon entropy.

2.11 Order Statistics

Suppose T_1, T_2, \dots, T_n is a random sample from (2). Let $T_{1:n} \leq T_{2:n} \leq \dots \leq T_{n:n}$ denote the the corresponding order statistics. It is well known that the probability density function of the of r th order statistic, say $T_{r:n}$, $1 \leq r \leq n$, is given by

$$\begin{aligned}
f_{r:n}(t) &= \frac{n!}{(r-1)!(n-r)!} [F(t)]^{r-1} [1-F(t)]^{n-r} f(t), \quad t > 0 \\
&= \frac{n!}{(r-1)!(n-r)!} \sum_{u=0}^{n-r} (-1)^u \binom{n-r}{u} [F(t)]^{r-1+u} f(t), \tag{11}
\end{aligned}$$

and the cumulative distribution function

$$\begin{aligned}
F_{r:n}(t) &= \sum_{l=k}^n \binom{n}{l} [F(t)]^l [1-F(t)]^{n-l} \\
&= \sum_{l=k}^n \sum_{u=0}^{n-r} (-1)^u \binom{n}{l} \binom{n-r}{u} [F(t)]^{l+u} \tag{12}
\end{aligned}$$

for $k = 1, 2, \dots, n$. It follows from (1) and (2) that

$$f_{r:n}(t) = \frac{\beta\lambda\delta t^{-\delta-1} n!}{(r-1)!(n-r)!} \sum_{u=0}^{n-r} (-1)^u \binom{n-r}{u} (1 + \lambda t^{-\delta})^{-\beta(r+u+\frac{1}{\beta})} \tag{13}$$

and

$$F_{r:n}(t) = \sum_{l=k}^n \sum_{u=0}^{n-r} (-1)^u \binom{n}{l} \binom{n-r}{u} (1 + \lambda t^{-\delta})^{-\beta(l+u)} \tag{14}$$

The j th moments of $T_{r:n}$ can be expressed

$$E[T_{r:n}^j] = \frac{n!\beta\lambda^{1+\frac{j}{\delta}}}{(r-1)!(n-r)!} \sum_{u=0}^{n-r} (-1)^u \binom{n-r}{u} B\left(1 - \frac{j}{\delta}, \frac{j}{\delta} - \beta\left(r+u+\frac{1}{\beta}\right) - 1\right) \tag{15}$$

2.12 Bonferroni and Lorenz Curves

Bonferroni and Lorenz curves are proposed by Bonferroni (1930). These curves have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medicine. They are defined as

$$B(p) = \frac{1}{p\mu} \int_0^q t f(t) dt \quad (16)$$

$$L(p) = \frac{1}{\mu} \int_0^q t f(t) dt, \quad (17)$$

where $\mu = E[X]$ and $q = F^{-1}(p)$. By using (1), one can reduce (16) and (17) to

$$B(p) = \frac{\beta \lambda^{\frac{1}{\delta}} (q^{-\delta} \lambda)^{-\frac{1}{\delta}}}{p\mu} \sum_{n=0}^{\infty} (-1)^n \binom{\beta + n + 1}{n} \frac{(q \lambda^{-\frac{1}{\delta}})^{n+1}}{(n - \frac{1}{\delta} + 1)}$$

and

$$L(p) = \frac{\beta \lambda^{\frac{1}{\delta}} (q^{-\delta} \lambda)^{-\frac{1}{\delta}}}{\mu} \sum_{n=0}^{\infty} (-1)^n \binom{\beta + n + 1}{n} \frac{(q \lambda^{-\frac{1}{\delta}})^{n+1}}{(n - \frac{1}{\delta} + 1)}.$$

3. Methods of Estimation

In this section we describe ten estimation methods for estimating the parameters β , λ and δ of the $Dag(\beta, \lambda, \delta)$ distribution. For all methods, we consider the case when all the parameters β , λ and δ are unknown. This is also considered in the simulation study presented in Section 4.

3.1 Method of Maximum Likelihood

The method of maximum likelihood is the most frequently used method of parameter estimation (see Casella and Berger (1990)). The method's success stems no doubt from its many desirable properties including consistency, asymptotic efficiency, invariance and simply its intuitive appeal. Let t_1, \dots, t_n be a random sample of size n from the $Dag(\beta, \lambda, \delta)$ distribution with parameters β, λ and δ .

From (1) the likelihood is

$$L(\beta, \lambda, \delta) = (\alpha\beta\delta)^n \prod_{i=1}^n t_i^{-\delta-1} (1 + \lambda t_i^{-\delta})^{-\beta-1}, \quad (18)$$

and log-likelihood function is

$$\begin{aligned} l(\beta, \lambda, \delta) &= n \log \beta + n \log \lambda + n \log \delta - (\delta - 1) \sum_{i=1}^n \ln t_i \\ &\quad - (\beta + 1) \sum_{i=1}^n \ln(1 + \lambda t_i^{-\delta}). \end{aligned} \quad (19)$$

The maximum likelihood estimators of $\widehat{\beta}_{MLE}$, $\widehat{\lambda}_{MLE}$ and $\widehat{\delta}_{MLE}$ of the parameters β , λ and δ , can be obtained numerically by maximizing, with respect to β , λ and δ , the log-likelihood function (19). In this case, the log-likelihood function is maximized by solving in β , λ and δ , the non-linear equations are:

$$\frac{\partial}{\partial \beta} l(\beta, \lambda, \delta) = \frac{n}{\beta} - \sum_{i=1}^n \ln(1 + \lambda t_i^{-\delta}) = 0, \quad (20)$$

$$\frac{\partial}{\partial \lambda} l(\beta, \lambda, \delta) = \frac{n}{\lambda} - (\beta + 1) \sum_{i=1}^n \frac{t_i^{-\delta}}{(1 + \lambda t_i^{-\delta})} = 0, \quad (21)$$

$$\frac{\partial}{\partial \delta} l(\beta, \lambda, \delta) = \frac{n}{\delta} - \sum_{i=1}^n \log t_i + (\beta + 1) \sum_{i=1}^n \frac{t_i^{-\delta}}{(1 + \lambda t_i^{-\delta})} = 0. \quad (22)$$

3.2 Method of Moments

The MMEs of the three-parameter $Dag(\beta, \lambda, \delta)$ distribution can be obtained by equating the first three theoretical moments of (1) with the sample moments $\frac{1}{n} \sum_{i=1}^n t_i$, $\frac{1}{n} \sum_{i=1}^n t_i^2$ and $\frac{1}{n} \sum_{i=1}^n t_i^3$ respectively,

$$\frac{1}{n} \sum_{i=1}^n t_i = \beta \lambda^{\frac{1}{\delta}} B\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right), \quad (23)$$

$$\frac{1}{n} \sum_{i=1}^n t_i^2 = \beta \lambda^{\frac{2}{\delta}} B\left(1 - \frac{2}{\delta}, \beta + \frac{2}{\delta}\right) \quad (24)$$

$$\frac{1}{n} \sum_{i=1}^n t_i^3 = \beta \lambda^{\frac{3}{\delta}} B\left(1 - \frac{3}{\delta}, \beta + \frac{3}{\delta}\right) \quad (25)$$

3.3 Method of L-Moments

In this subsection we provide the L-moments estimators, which can be obtained as the linear combinations of order statistics. The L-moments estimators were originally proposed by Hosking (1990), and it is observed that the L-moments estimators are more robust than the usual moment estimators. The L-moment estimators are also obtained along the same way as the ordinary moment estimators, i.e., by equating the sample L-moments with the population L-moments. L-moment estimation provides an alternative method of estimation analogous to conventional moments and have the advantage that they exist whenever the mean of the distribution exists, even though some higher moments may not exist, and are relatively robust to the effects of outliers (Hosking, 1994).

Let $t_{1:n} < \dots < t_{n:n}$ be the order statistics of a random sample of size n from $DAG(\beta, \lambda, \delta)$ distribution. From Hosking(1990), the first, second and third sample L-moments, respectively, are

$$\begin{aligned} l_1 &= \bar{t} \\ l_2 &= \frac{2}{n(n-1)} \sum_{i=2}^n (i-1) t_{i:n} - l_1 \\ l_3 &= \frac{6}{n(n-1)(n-2)} \sum_{i=3}^n (i-1)(i-2) t_{i:n} - \frac{6}{n(n-1)} \sum_{i=2}^n (i-1) t_{i:n} + l_1 \end{aligned}$$

Since the quantile function of the $Dag(\beta, \lambda, \delta)$ distribution is as given in (3), then the first, second and third population L-moments of $\theta = (\beta, \lambda, \delta)$, respectively, are

$$\begin{aligned} \lambda_1(\theta) &= \int_0^1 Q(p|\theta) dp = \beta \lambda^{\frac{1}{\delta}} B\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right) \\ \lambda_2(\theta) &= \int_0^1 Q(p|\theta)(2p-1) dp = 2\beta \lambda^{\frac{1}{\delta}} B\left(1 - \frac{1}{\delta}, 2\beta + \frac{1}{\delta}\right) - \lambda_1(\theta), \end{aligned}$$

and

$$\begin{aligned} \lambda_3(\theta) &= \int_0^1 Q(p|\theta)(6p^2 - 6p + 1) dp \\ &= \beta \lambda^{\frac{1}{\delta}} \left[6B\left(1 - \frac{1}{\delta}, 3\beta + \frac{1}{\delta}\right) - 6B\left(1 - \frac{1}{\delta}, 2\beta + \frac{1}{\delta}\right) \right. \\ &\quad \left. + B\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right) \right]. \end{aligned}$$

The L-moments estimators $\widehat{\beta}_{LME}$, $\widehat{\lambda}_{LME}$ and $\widehat{\delta}_{LME}$ of the parameters β , λ and δ can be obtained by solving numerically the equations

$$\begin{aligned} \lambda_1(\widehat{\beta}_{LME}, \widehat{\lambda}_{LME}, \widehat{\delta}_{LME}) &= l_1, \\ \lambda_2(\widehat{\beta}_{LME}, \widehat{\lambda}_{LME}, \widehat{\delta}_{LME}) &= l_2 \\ \lambda_3(\widehat{\beta}_{LME}, \widehat{\lambda}_{LME}, \widehat{\delta}_{LME}) &= l_3 \end{aligned}$$

3.4 Method of Maximum Product of Spacings

Cheng and Amin (1979, 1983) introduced the maximum product of spacings (MPS) method as an alternative to MLE for the estimation of parameters of continuous univariate distributions. Ranneby (1984) independently developed the same method as an approximation to the Kullback-Leibler measure of information.

Let us define the uniform spacings of a random sample from the Dagum distribution distribution as

$$D_i(\beta, \lambda, \delta) = F(t_{i:n} | \beta, \lambda, \delta) - F(t_{i-1:n} | \beta, \lambda, \delta),$$

where $F(t_{0:n} | \beta, \lambda, \delta) = 0$, $F(t_{n+1:n} | \beta, \lambda, \delta) = 1$, and $\sum_{i=1}^{n+1} D_i(\beta, \lambda, \delta) = 1$.

The maximum product of spacings estimators $\widehat{\beta}_{MPS}$, $\widehat{\lambda}_{MPS}$ and $\widehat{\delta}_{MPS}$ of the parameters β , λ and δ are obtained by maximizing the geometric mean of the spacings with respect to β , λ and δ .

$$G(\beta, \lambda, \delta) = \left[\prod_{i=1}^{n+1} D_i(\beta, \lambda, \delta) \right]^{\frac{1}{n+1}}, \quad (26)$$

or, equivalently, by maximizing the function

$$H(\beta, \lambda, \delta) = \frac{1}{n+1} \sum_{i=1}^{n+1} \ln D_i(\beta, \lambda, \delta). \quad (27)$$

The estimators $\widehat{\theta}_{MPS} = (\widehat{\beta}_{MPS}, \widehat{\lambda}_{MPS}, \widehat{\delta}_{MPS})$ of the parameters $\theta = (\beta, \lambda, \delta)$ can also be obtained by solving the nonlinear equations

$$\frac{\partial}{\partial \beta} H(\theta) = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\theta)} [\Delta_1(t_{i:n} | \theta) - \Delta_1(t_{i-1:n} | \theta)] = 0, \quad (28)$$

$$\frac{\partial}{\partial \lambda} H(\theta) = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\theta)} [\Delta_2(t_{i:n} | \theta) - \Delta_2(t_{i-1:n} | \theta)] = 0, \quad (29)$$

$$\frac{\partial}{\partial \delta} H(\theta) = \frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\theta)} [\Delta_3(t_{i:n} | \theta) - \Delta_3(t_{i-1:n} | \theta)] = 0, \quad (30)$$

where

$$\Delta_1(t_{i:n} | \beta, \lambda, \delta) = -(1 + \lambda t_{i:n}^{-\delta})^{-\beta} \log(1 + \lambda t_{i:n}^{-\delta}). \quad (31)$$

$$\Delta_2(t_{i:n} | \beta, \lambda, \delta) = -\beta(1 + \lambda t_{i:n}^{-\delta})^{-\beta-1} t_{i:n}^{-\delta}. \quad (32)$$

$$\Delta_3(t_{i:n} | \beta, \lambda, \delta) = \beta \lambda (1 + \lambda t_{i:n}^{-\delta})^{-\beta-1} t_{i:n}^{-\delta} \log(t_{i:n}). \quad (33)$$

Cheng and Amin (1983) showed that maximizing H as a method of parameter estimation is as efficient as MLE estimation and the MPS estimators are consistent under more general conditions than the MLE estimators.

3.5 Methods of Ordinary and Weighted Least-Squares

The least square estimators and weighted least square estimators were proposed by Swain et al. (1988) to estimate the parameters of Beta distributions. Using the same notations in subsection (3.3), it is well known that

$$E[F(t_{i:n} | \beta, \lambda, \delta)] = \frac{i}{n+1},$$

$$V[F(t_{i:n} | \beta, \lambda, \delta)] = \frac{i(n-i+1)}{(n+1)^2(n+2)}.$$

The ordinary least square estimates $\widehat{\beta}_{OLSE}$, $\widehat{\lambda}_{OLSE}$ and $\widehat{\delta}_{OLSE}$ of the parameters β , λ and δ are obtained by minimizing the function:

$$S(\beta, \lambda, \delta) = \sum_{i=1}^n \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right]^2. \quad (34)$$

These estimates can also be obtained by solving the non-linear equations:

$$\begin{aligned}\sum_{i=1}^n \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right] \Delta_1(t_{i:n} | \beta, \lambda, \delta) &= 0, \\ \sum_{i=1}^n \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right] \Delta_2(t_{i:n} | \beta, \lambda, \delta) &= 0, \\ \sum_{i=1}^n \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right] \Delta_3(t_{i:n} | \beta, \lambda, \delta) &= 0,\end{aligned}$$

where $\Delta_1(\cdot | \beta, \lambda, \delta)$, $\Delta_2(\cdot | \beta, \lambda, \delta)$ and $\Delta_3(\cdot | \beta, \lambda, \delta)$ are given by (31), (32) and (33), respectively.

The weighted least-squares estimators $\widehat{\beta}_{WLS}$, $\widehat{\lambda}_{WLS}$ and $\widehat{\delta}_{WLS}$ of the parameters β , λ and δ are obtained by minimizing the function:

$$W(\beta, \lambda, \delta) = \sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right]^2. \quad (35)$$

These estimators can also be obtained by solving the non-linear equations:

$$\sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right] \Delta_1(t_{i:n} | \beta, \lambda, \delta) = 0, \quad (36)$$

$$\sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right] \Delta_2(t_{i:n} | \beta, \lambda, \delta) = 0, \quad (37)$$

$$\sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(t_{i:n} | \beta, \lambda, \delta) - \frac{i}{n+1} \right] \Delta_3(t_{i:n} | \beta, \lambda, \delta) = 0, \quad (38)$$

where $\Delta_1(\cdot | \beta, \lambda, \delta)$, $\Delta_2(\cdot | \beta, \lambda, \delta)$ and $\Delta_3(\cdot | \beta, \lambda, \delta)$ are given by (31), (32) and (33), respectively.

3.6 Method of Percentiles

The Dagum distribution has an explicit distribution function, therefore in this case the unknown parameters β , λ and δ can be estimated by equating the sample percentile points with the population percentile points and it is known as the percentile method. This method was originally suggested by Kao (1958, 1959) and it has been used for Weibull distribution and for generalized exponential distribution. In this paper, we apply the same technique for the Dagum distribution. If p_i denotes an estimate of $F(t_{i:n} | \beta, \lambda, \delta)$, then the percentile estimators $\widehat{\beta}_{PCE}$, $\widehat{\lambda}_{PCE}$ and $\widehat{\delta}_{PCE}$ of the parameters β , λ and δ can be obtained by minimizing, with respect to β , λ and δ the function:

$$P(\beta, \lambda, \delta) = \sum_{i=1}^n \left[t_{i:n} - \left[\frac{1}{\lambda} \left[p_i^{-\frac{1}{\beta}} - 1 \right] \right]^{-\frac{1}{\delta}} \right]^2,$$

where $p_i = \frac{i}{n+1}$ is the unbiased estimator of $F(t_{i:n} | \beta, \lambda, \delta)$ (see Mann et al. (1974)). The estimates of β , λ and δ can be obtained by solving the following nonlinear equations

$$\begin{aligned}\sum_{i=1}^n \frac{\lambda^{\frac{1}{\delta}} \left[t_{i:n} - \left(c1 \lambda [p_i^{-\frac{1}{\beta}} - 1] \right)^{-\frac{1}{\delta}} \right] \ln(p_i)}{\delta \beta^2 \left[(p_i^{-\frac{1}{\beta}} - 1)^{\frac{1}{\delta} + 1} \right] p_i^{\frac{1}{\beta}}} &= 0 \\ \sum_{i=1}^n \frac{\lambda^{\frac{1}{\delta} - 1} \left[t_{i:n} - \left(\frac{1}{\lambda} [p_i^{-\frac{1}{\beta}} - 1] \right)^{-\frac{1}{\delta}} \right]}{\delta \left[(p_i^{-\frac{1}{\beta}} - 1)^{\frac{1}{\delta}} \right]} &= 0 \\ \sum_{i=1}^n \frac{\lambda^{\frac{1}{\delta}} \ln \lambda \left[t_{i:n} - \left(\frac{1}{\lambda} [p_i^{-\frac{1}{\beta}} - 1] \right)^{-\frac{1}{\delta}} \right] \ln(p_i^{-\frac{1}{\beta}} - 1)}{\delta^2 \left[(p_i^{-\frac{1}{\beta}} - 1)^{\frac{1}{\delta}} \right]} &= 0\end{aligned}$$

respectively.

3.7 Methods of Minimum Distances

In this section we present three estimation methods for β , λ and δ based on the minimization of the goodness-of-fit statistics with respect to β , λ and δ . This class of statistics is based on the difference between the estimates of the cumulative distribution function and the empirical distribution function.

3.7.1 Method of Cramér-von-Mises

To motivate our choice of Cramér-von-Mise type minimum distance estimators, MacDonald (1971) provided empirical evidence that the bias of the estimator is smaller than the other minimum distance estimators. Thus, the Cramér-von Mises estimates $\widehat{\beta}_{CME}$, $\widehat{\lambda}_{CME}$ and $\widehat{\delta}_{CME}$ of the parameters β , λ and δ are obtained by minimizing $C(\beta, \lambda, \delta)$ with respect to β , λ and δ

$$C(\beta, \lambda, \delta) = \frac{1}{12n} + \sum_{i=1}^n \left(F(t_{i:n} | \beta, \lambda, \delta) - \frac{2i-1}{2n} \right)^2. \quad (39)$$

The estimators can be obtained by solving the following non-linear equations

$$\begin{aligned} \sum_{i=1}^n \left(F(t_{i:n} | \beta, \lambda, \delta) - \frac{2i-1}{2n} \right) \Delta_1(t_{i:n} | \beta, \lambda, \delta) &= 0, \\ \sum_{i=1}^n \left(F(t_{i:n} | \beta, \lambda, \delta) - \frac{2i-1}{2n} \right) \Delta_2(t_{i:n} | \beta, \lambda, \delta) &= 0, \\ \sum_{i=1}^n \left(F(t_{i:n} | \beta, \lambda, \delta) - \frac{2i-1}{2n} \right) \Delta_3(t_{i:n} | \beta, \lambda, \delta) &= 0, \end{aligned}$$

where $\Delta_1(\cdot | \beta, \lambda, \delta)$, $\Delta_2(\cdot | \beta, \lambda, \delta)$ and $\Delta_3(\cdot | \beta, \lambda, \delta)$ are given by (31), (32) and (33), respectively.

3.7.2 Methods of Anderson-Darling and Right-tail Anderson-Darling

The Anderson-Darling test was developed by Anderson and Darling (1952, 1954) as an alternative to other statistical tests for detecting sample distributions departure from normality. Specifically, the AD test converge very quickly towards the asymptote.

The Anderson-Darling estimates $\widehat{\beta}_{ADE}$, $\widehat{\lambda}_{ADE}$ and $\widehat{\delta}_{ADE}$ of the parameters β , λ and δ are obtained by minimizing $A(\beta, \lambda, \delta)$ with respect to β , λ and δ .

$$\begin{aligned} A(\beta, \lambda, \delta) &= -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left\{ \log F(t_{i:n} | \beta, \lambda, \delta) \right. \\ &\quad \left. + \log \bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta) \right\}. \end{aligned} \quad (40)$$

The estimators can be obtained by solving following the non-linear equations:

$$\begin{aligned} \sum_{i=1}^n (2i-1) \left[\frac{\Delta_1(t_{i:n} | \beta, \lambda, \delta)}{F(t_{i:n} | \beta, \lambda, \delta)} - \frac{\Delta_1(t_{n+1-i:n} | \beta, \lambda, \delta)}{\bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta)} \right] &= 0, \\ \sum_{i=1}^n (2i-1) \left[\frac{\Delta_2(t_{i:n} | \beta, \lambda, \delta)}{F(t_{i:n} | \beta, \lambda, \delta)} - \frac{\Delta_2(t_{n+1-i:n} | \beta, \lambda, \delta)}{\bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta)} \right] &= 0, \\ \sum_{i=1}^n (2i-1) \left[\frac{\Delta_3(t_{i:n} | \beta, \lambda, \delta)}{F(t_{i:n} | \beta, \lambda, \delta)} - \frac{\Delta_3(t_{n+1-i:n} | \beta, \lambda, \delta)}{\bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta)} \right] &= 0, \end{aligned}$$

where $\Delta_1(\cdot | \beta, \lambda, \delta)$, $\Delta_2(\cdot | \beta, \lambda, \delta)$ and $\Delta_3(\cdot | \beta, \lambda, \delta)$ are given by (31), (32) and (32), respectively.

The right-tail Anderson-Darling estimates $\widehat{\alpha}_{RTADE}$ and $\widehat{\beta}_{RTADE}$ of the parameters α and β are obtained by minimizing $R(\beta, \lambda, \delta)$ with respect to α and β :

$$\begin{aligned} R(\beta, \lambda, \delta) &= \frac{n}{2} - 2 \sum_{i=1}^n F(t_{i:n} | \beta, \lambda, \delta) \\ &\quad - \frac{1}{n} \sum_{i=1}^n (2i-1) \log \bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta). \end{aligned} \quad (41)$$

The estimators can also be obtained by solving the following non-linear equations:

$$\begin{aligned} -2 \sum_{i=1}^n \frac{\Delta_1(t_{i:n} | \beta, \lambda, \delta)}{F(t_{i:n} | \beta, \lambda, \delta)} + \frac{1}{n} \sum_{i=1}^n (2i-1) \frac{\Delta_1(t_{n+1-i:n} | \beta, \lambda, \delta)}{\bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta)} &= 0, \\ -2 \sum_{i=1}^n \frac{\Delta_2(t_{i:n} | \beta, \lambda, \delta)}{F(t_{i:n} | \beta, \lambda, \delta)} + \frac{1}{n} \sum_{i=1}^n (2i-1) \frac{\Delta_2(t_{n+1-i:n} | \beta, \lambda, \delta)}{\bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta)} &= 0, \\ -2 \sum_{i=1}^n \frac{\Delta_3(t_{i:n} | \beta, \lambda, \delta)}{F(t_{i:n} | \beta, \lambda, \delta)} + \frac{1}{n} \sum_{i=1}^n (2i-1) \frac{\Delta_3(t_{n+1-i:n} | \beta, \lambda, \delta)}{\bar{F}(t_{n+1-i:n} | \beta, \lambda, \delta)} &= 0, \end{aligned}$$

where $\Delta_1(\cdot | \beta, \lambda, \delta)$, $\Delta_2(\cdot | \beta, \lambda, \delta)$ and $\Delta_3(\cdot | \beta, \lambda, \delta)$ are given by (31), (32) and (33), respectively.

4. Numerical computations

4.1 Simulation Study

In this section, we conduct numerical computations to compare the performances of the estimators proposed in the previous sections. The generation of the Dagum distribution can be easily obtained through the transformation $T = \left[\frac{1}{\lambda} (U^{-1/\beta} - 1) \right]^{-1/\delta}$, where U is a uniform on $(0, 1)$. We consider two choices of the shape parameter, $\beta = 1, 2$ and for sample sizes $n = 20, 60, 80$ and 100 . In both cases considered, we take $\lambda = 1$ and $\delta = 4$.

For each estimate $\hat{\theta} = (\hat{\beta}, \hat{\lambda}, \hat{\delta})$ we compute the bias and the root mean-squared error (RMSE), respectively, as

$$\begin{aligned} \text{Bias}(\hat{\theta}) &= \frac{1}{B} \sum_{i=1}^B (\hat{\theta}_i - \theta) \\ \text{RMSE}(\hat{\theta}) &= \sqrt{\frac{1}{B} \sum_{i=1}^B (\hat{\theta}_i - \theta)^2} \end{aligned}$$

The following observations can be drawn from the Tables 1 and 2.

1. All the estimators show the property of consistency i.e., the bias and RMSE decreases as sample size increases.
2. The bias of $\hat{\beta}$ generally increases with increasing β for any given λ and δ and n and for all methods of estimation.
3. The RMSE of $\hat{\beta}$ generally increases with increasing β for any given λ and δ and n and for all methods of estimation.
4. Some estimators of the parameters are positively biased and some are negatively biased.
5. In terms of performance of the methods of estimation, we found that maximum product spacing (MPS) estimators, ordinary least square estimators (OLSE), L-moment estimators (LME) and Cramér-von-Mises estimators (CME) are equally produces the least biases of the estimates with least RMSE for most of the configurations considered in our studies.

4.2 Real Application

Here we use one data set that will be used for the purpose of making comparisons between the estimators presented in this study. The data set is taken from Nichols and Padgett (Nichols Padgett, 2006) consisting of 100 observations on breaking stress of carbon fibers (in Gba). The data are : 3.7, 2.74, 2.73, 3.11, 3.27, 2.87, 4.42, 2.41, 3.19, 3.28, 3.09, 1.87, 3.75, 2.43, 2.95, 2.96, 2.3, 2.67, 3.39, 2.81, 4.2, 3.31, 3.31, 2.85, 3.15, 2.35, 2.55, 2.81, 2.77, 2.17, 1.41, 3.68, 2.97, 2.76, 4.91, 3.68, 3.19, 1.57, 0.81, 1.59, 2, 1.22, 2.17, 1.17, 5.08, 3.51, 2.17, 1.69, 1.84, 0.39, 3.68, 1.61, 2.79, 4.7, 1.57, 1.08, 2.03, 1.89, 2.88, 2.82, 2.5, 3.6, 1.47, 3.11, 3.22, 1.69, 3.15, 4.9, 2.97, 3.39, 2.93, 3.22, 3.33, 2.55, 2.56, 3.56, 2.59, 2.38, 2.83, 1.92, 1.36, 0.98, 1.84, 1.59, 5.56, 1.73, 1.12, 1.71, 2.48, 1.18, 1.25, 4.38, 2.48, 0.85, 2.03, 1.8, 1.61, 2.12, 2.05, 3.65.

We have fitted the Dagum distribution to the data set using MLE (standard error in parentheses), and compared the proposed distribution with Marshall-Olkin extended Lomax (MOEL) and Fisk models, see Table 3. The model selection is carried out using the AIC (Akaike information criterion), the BIC (Bayesian information criterion), the CAIC (consistent

Table 1. Average bias and RMSE of β , δ and λ for $\beta = 1$, $\lambda = 1$ and $\delta = 4$

n	Method	β		λ		δ	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
20	MLE	0.0763	0.4290	0.1048	0.6226	0.0356	0.3186
	MME	-0.0166	0.0807	-0.0078	0.0924	0.0036	0.1227
	LME	-0.0016	0.0188	-0.0073	0.0501	-0.0004	0.0189
	MPS	0.0080	0.0873	0.0076	0.0820	0.0066	0.0706
	OLES	0.0219	0.1225	0.0042	0.0526	0.0174	0.0951
	WLES	0.0435	0.2185	0.0103	0.0789	0.0290	0.1574
	PCE	-0.0137	0.1041	0.0072	0.0616	0.0231	0.1047
	CME	0.0172	0.1123	0.0012	0.0486	0.0154	0.0845
	ADE	0.0045	0.0999	0.0074	0.0598	-0.0008	0.0910
	RTADE	0.0014	0.1374	0.0617	0.2761	-0.1263	0.5648
60	MLE	0.0076	0.0937	0.0068	0.1186	0.0030	0.0960
	MME	-0.0050	0.0403	-0.0021	0.0432	0.0003	0.0581
	LME	-0.0004	0.0048	-0.0030	0.0237	-0.0007	0.0058
	MPS	0.0023	0.0479	0.0022	0.0448	0.0019	0.0385
	OLES	0.0019	0.0245	-0.0003	0.0173	0.0043	0.0345
	WLES	0.0065	0.0615	0.0001	0.0251	0.0064	0.0535
	PCE	-0.0058	0.0521	0.0022	0.0257	0.0050	0.0384
	CME	0.0012	0.0202	-0.0007	0.0158	0.0036	0.0294
	ADE	0.0012	0.0587	0.0020	0.0353	0.0000	0.0551
	RTADE	0.0044	0.0610	0.0185	0.1431	-0.0386	0.2992
80	MLE	0.0035	0.0597	0.0036	0.0812	0.0019	0.0729
	MME	-0.0036	0.0334	-0.0012	0.0365	0.0006	0.0511
	LME	-0.0003	0.0033	-0.0023	0.0206	-0.0005	0.0050
	MPS	0.0018	0.0391	0.0017	0.0366	0.0015	0.0315
	OLES	0.0009	0.0150	-0.0002	0.0144	0.0033	0.0298
	WLES	0.0040	0.0446	0.0001	0.0209	0.0047	0.0436
	PCE	-0.0047	0.0472	0.0019	0.0215	0.0034	0.0302
	CME	0.0005	0.0111	-0.0005	0.0126	0.0023	0.0219
	ADE	0.0011	0.0511	0.0016	0.0294	-0.0001	0.0467
	RTADE	0.0041	0.0486	0.0139	0.1244	-0.0282	0.2526
100	MLE	0.0021	0.0443	0.0020	0.0601	0.0011	0.0555
	MME	-0.0028	0.0284	-0.0008	0.0308	0.0007	0.0449
	LME	-0.0002	0.0025	-0.0018	0.0184	-0.0004	0.0045
	MPS	0.0014	0.0351	0.0013	0.0329	0.0011	0.0283
	OLES	0.0007	0.0124	-0.0002	0.0120	0.0026	0.0267
	WLES	0.0048	0.0546	0.0001	0.0204	0.0046	0.0480
	PCE	-0.0042	0.0454	0.0016	0.0191	0.0020	0.0202
	CME	0.0004	0.0096	-0.0003	0.0108	0.0018	0.0187
	ADE	0.0007	0.0467	0.0014	0.0282	-0.0002	0.0427
	RTADE	0.0034	0.0425	0.0111	0.1114	-0.0225	0.2248

Akaike information criteria) and the HQIC (Hannan-Quinn information criterion), where

$$AIC = -2l(\hat{\theta}) + 2q$$

$$BIC = -2l(\hat{\theta}) + q \log(n)$$

$$HQIC = -2l(\hat{\theta}) + 2q \log(\log(n))$$

$$CAIC = -2l(\hat{\theta}) + \frac{2qn}{n - q - 1}$$

where $l(\hat{\theta})$ denotes the log-likelihood function evaluated at the maximum likelihood estimates, q is the number of parameters, and n is the sample size. The model with minimum AIC or (BIC, CAIC and HQIC) value is chosen as the best model to fit the data. Estimates of the parameters of Dagum distribution, MOEL and Fisk models (standard error in parentheses),

Table 2. Average bias and RMSE of β , δ and λ for $\beta = 2$, $\lambda = 1$ and $\delta = 4$

n	Method	β		λ		δ	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
20	MLE	0.1475	0.8432	0.1389	0.7943	0.0318	0.2538
	MME	-0.0431	0.1928	-0.0090	0.1031	-0.0033	0.1168
	LME	0.0343	0.2368	0.0673	0.3721	0.0122	0.0751
	MPS	0.0148	0.1284	0.0154	0.1344	0.0130	0.1107
	OLES	0.0111	0.1017	0.0155	0.0844	0.0217	0.1123
	WLES	0.0317	0.2019	0.0339	0.1635	0.0256	0.1472
	PCE	0.0147	0.1327	0.0059	0.0341	0.0415	0.1879
	CME	0.0072	0.1065	0.0124	0.0801	0.0196	0.0983
	ADE	-0.0077	0.1101	0.0128	0.0842	0.0019	0.1206
	RTADE	-0.0309	0.1384	0.0565	0.2529	-0.1133	0.5067
60	MLE	0.0250	0.2701	0.0080	0.1441	0.0025	0.0786
	MME	-0.0142	0.1098	-0.0028	0.0579	-0.0024	0.0608
	LME	0.0208	0.1914	0.0100	0.1176	0.0054	0.0525
	MPS	0.0034	0.0605	0.0036	0.0640	0.0029	0.0516
	OLES	0.0005	0.0199	0.0009	0.0151	0.0038	0.0313
	WLES	0.0041	0.0608	0.0047	0.0442	0.0073	0.0631
	PCE	0.0003	0.0544	0.0007	0.0112	0.0081	0.0629
	CME	-0.0002	0.0195	0.0004	0.0124	0.0032	0.0273
	ADE	-0.0033	0.0677	0.0041	0.0481	0.0010	0.0726
	RTADE	-0.0102	0.0789	0.0188	0.1460	-0.0378	0.2925
80	MLE	0.0141	0.1901	0.0045	0.1032	0.0015	0.0608
	MME	-0.0105	0.0940	-0.0018	0.0498	-0.0016	0.0522
	LME	0.0018	0.0451	0.0050	0.0770	0.0020	0.0265
	MPS	0.0025	0.0495	0.0027	0.0522	0.0022	0.0424
	OLES	0.0004	0.0150	0.0006	0.0113	0.0029	0.0268
	WLES	0.0028	0.0479	0.0029	0.0329	0.0049	0.0472
	PCE	-0.0002	0.0373	0.0002	0.0101	0.0060	0.0538
	CME	-0.0002	0.0124	0.0002	0.0084	0.0018	0.0179
	ADE	-0.0024	0.0574	0.0033	0.0412	0.0006	0.0628
	RTADE	-0.0076	0.0683	0.0141	0.1265	-0.0283	0.2533
100	MLE	0.0096	0.1486	0.0023	0.0754	0.0008	0.0470
	MME	-0.0084	0.0837	-0.0013	0.0445	-0.0012	0.0468
	LME	0.0103	0.0272	0.0052	0.0809	0.0022	0.0300
	MPS	0.0020	0.0438	0.0021	0.0463	0.0017	0.0375
	OLES	0.0002	0.0116	0.0004	0.0084	0.0020	0.0208
	WLES	0.0027	0.0477	0.0030	0.0359	0.0044	0.0465
	PCE	-0.0008	0.0302	0.0000	0.0081	0.0041	0.0418
	CME	-0.0001	0.0109	0.0001	0.0069	0.0014	0.0151
	ADE	-0.0020	0.0531	0.0026	0.0382	0.0004	0.0570
	RTADE	-0.0061	0.0611	0.0113	0.1131	-0.0227	0.2266

AIC, BIC, CAIC and HQIC for remission times data are presented in Table 4.2. We conclude that the Dagum distribution is comparable to the MOEL and Fisk models.

For an ordered random sample, X_1, X_2, \dots, X_n , from $Dag(\beta, \sigma)$, where the parameters β , and σ are unknown to us, the Kolmogorov-Smirnov D_n , Cramér-von Mises W_n^2 , Anderson and Darling A_n^2 and Watson U_n^2 tests statistics are computed and given in Table 4.2. The results show that the test statistics take the smallest values for the data set under Dagum distribution with regard to the other distribution. Thus, the proposed model offers an attractive alternative to the Marshall-Olkin extended Lomax and Fisk models. Figure 3 displays the empirical and fitted densities for breaking stress of carbon fibers data. Estimated survivals for the data are shown in Figure 4. The Dagum distribution approximately provide an

adequate fit for the data.

$$D_n = \max_i \left[\frac{1}{n} - F(x_i, \hat{\beta}, \hat{\delta}, \hat{\lambda}), F(x_i, \hat{\beta}, \hat{\delta}, \hat{\lambda}) - \frac{i-1}{n} \right].$$

$$W_n^2 = \frac{1}{12n} + \sum_{i=1}^n \left[F(x_i, \hat{\beta}, \hat{\delta}, \hat{\lambda}) - \frac{2i-1}{2n} \right]^2.$$

$$A_n^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left[\log(F(x_i, \hat{\beta}, \hat{\delta}, \hat{\lambda})) + \log(1 - F(x_i, \hat{\beta}, \hat{\delta}, \hat{\lambda})) \right]^2.$$

$$U_n^2 = W_n^2 + \sum_{i=1}^n \left[\frac{F(x_i, \hat{\beta}, \hat{\delta}, \hat{\lambda})}{n} - \frac{1}{2} \right]^2$$

Table 3. MLEs (standard errors in parentheses) for breaking stress of carbon fibers data.

Models	Estimates				
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\lambda}$
MOEL	96.3824 (53.5906)	0.0434 (0.0560)	43.2227 (49.8839)		
Dagum		0.3360 (0.0544)		7.4247 (0.6811)	8379.85 (7937.685)
Fisk	2.4948 (0.1053)	4.1156 (0.3438)			

Table 4. Goodness-of-fit tests and the measures AIC, BIC, HQIC and CAIC for breaking stress of carbon fibers data.

Models	Statistics				Measures			
	D_n	W_n^2	A_n^2	U_n^2	AIC	BIC	HQIC	CAIC
Dagum	0.0527	0.0437	0.2896	24.5491	288.28	296.09	291.44	288.53
MOEL	0.0655	0.0625	0.3815	24.5648	290.30	298.11	293.46	290.55
Fisk	0.0916	0.1579	1.0553	24.6613	296.42	301.63	298.53	296.55

Table 5. Estimates of the parameters of Dagum distribution for breaking stress of carbon fibers data.

Method	Estimates			Statistics	
	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\delta}$	D_n	W_n^2
MLE	0.33605	8379.186	7.42465	0.05265	0.04370
MME	0.34986	8426.464	7.52115	0.05578	0.05567
LME	0.58326	8377.958	7.39817	0.22728	2.26952
MPS	0.30096	8379.185	7.16180	0.05974	0.05605
OLEs	0.31897	8379.217	7.30835	0.05207	0.04077
WLES	0.33130	8379.186	7.41441	0.05572	0.04241
PCE	0.31481	8379.180	7.24698	0.05406	0.04479
CME	0.32363	8379.186	7.34435	0.05272	0.04016
ADE	0.32686	8379.186	7.37612	0.05428	0.04063
RTADE	0.09138	8379.191	5.57370	0.42108	6.33107

5. Conclusion

In this article, we provide explicit expressions for the quantiles, moments, moment generating function, conditional moments, hazard rate, mean residual lifetime, mean past lifetime, mean deviation about mean and median, various entropies, order statistics and Bonferroni and Lorenz curve. The model parameters are estimated by ten methods of estimation, namely, maximum likelihood, moments, L-moments, percentile, least squares, weighted least squares, maximum product of spacing, Cramér-von-Mises, Anderson-Darling and right tailed Anderson-Darling. We have performed an extensive

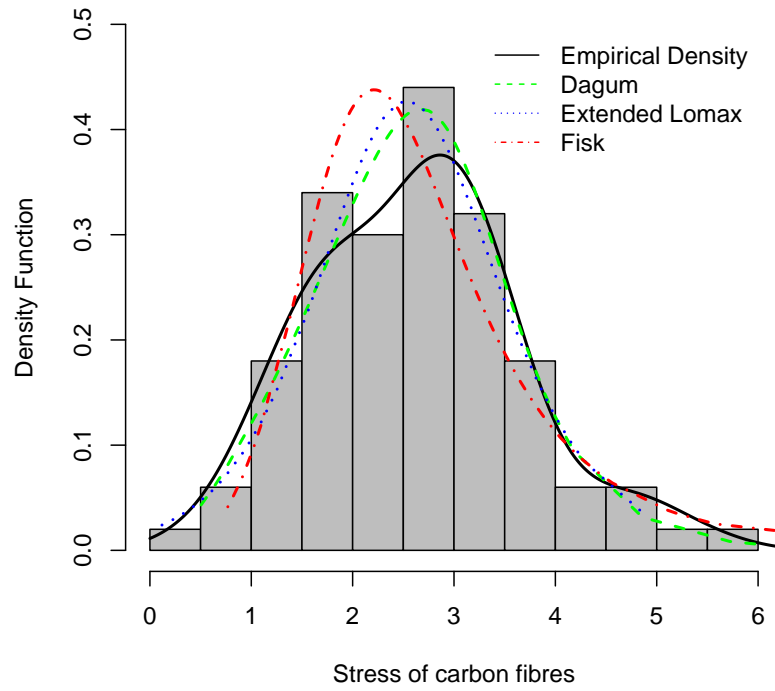


Figure 3. Estimated densities for breaking stress of carbon fibers data.

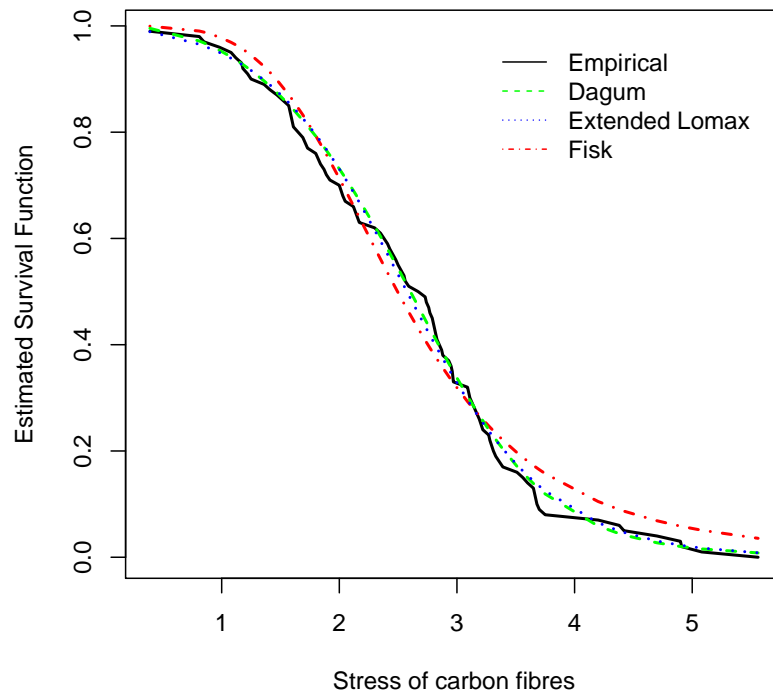


Figure 4. Estimated survivals for breaking stress of carbon fibers data.

simulation study to compare these methods. We have compared estimators with respect to bias and root mean-squared error. We have also compared estimators using a real data applications. The simulation results show that maximum product spacing estimators, ordinary least square estimators, L-moment estimators and Cramér-von-Mises estimators are equally produces the least biases of the estimates with least RMSE for most of the configurations considered in our studies.

Acknowledgment This study is a part of the Master Thesis of the third named author whose work was supervised by the second named author.

References

- Alkawasbeh, M. R., & Raqab, M. Z. (2009). Estimation of the generalized logistic distribution parameters: comparative study. *Statistical Methodology*, 6(3), 262–279.
- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, 193-212.
- Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American statistical association*, 49(268), 765-769.
- Bonferroni, C. E. (1930). *Elementi di statistica generale*, Libreria Seber: Firenze.
- Casella, G. & Berger, R. L. (1990), *Statistical Inference*, Wadsworth & Brooks/Cole.
- Cheng, R. C. H., & Amin, N. A. K. (1979). Maximum product of spacings estimation with application to the lognormal distribution. *Math. Report*, 79(1).
- Cheng, R. C. H., & Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society. Series B (Methodological)*, 394-403.
- Dagum, C. (1977). A New Model for Personal Income Distribution: Specification and Estimation. *Economic Applique*, 30(3), 413–437.
- Dey, S., Dey, T., & Kundu, D. (2014). Two-parameter Rayleigh distribution: different methods of estimation. *American Journal of Mathematical and Management Sciences*, 33(1), 55-74.
- Domma, F. (2002). L'andamento della Hazard function nel modello di Dagum a tre parametri. *Quaderni di Statistica*, 4, 1–12.
- Domma, F., Giordano, S. & Zenga, M. M. (2011). Maximum likelihood estimation in Dagum distribution with censored sample. *Journal of Applied Statistics*, 38(12), 2971-2985.
- Domma, F., & Condino, F. (2013). The Beta-Dagum distribution: definition and properties. *Communications in Statistics-Theory and Methods*, 42(22), 4070-4090.
- Gupta, R. D. & Kundu, D. (2001). Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, 69(4) 315-337.
- Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association*, 75(371), 667-672.
- Gupta, R. D., & Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical Planning and Inference*, 137(11), 3537-3547.
- Hosking, J. R. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 105-124.
- Hosking, J. R. (1994). The four-parameter kappa distribution. *IBM Journal of Research and Development*, 38(3), 251-258.
- Kundu, D., & Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimations. *Computational Statistics and Data Analysis*, 49(1), 187-200.
- Kleiber, C., & Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. 470. John Wiley & Sons.
- Kleiber, C. (2008). A guide to the Dagum distributions. In *Modeling Income Distributions and Lorenz Curves*, 97-117. Springer New York.
- Macdonald, P. D. M. (1971). Comment on "An estimation procedure for mixtures of distributions" by Choi and Bulgren. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33, 326-329.
- Mann, N. R., Singpurwalla, N. D., & Schafer, R. E. (1974). *Methods for statistical analysis of reliability and life data*. Nichols, M. D., & Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and reliability engineering international*, 22(2), 141-151.

- Oluyede, B. O., & Ye, Y. (2014). Weighted Dagum and related distributions. *Afrika Matematika*, 25(4), 1125-1141.
- Ranneby, B. (1984). The maximum spacing method. An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, 93-112.
- Rényi, A. (1961). On measures of entropy and information, in proceedings of the 4 th berkeley symposium on Mathematical Statistics and Probability, 1, 547C561.
- Shahzad, M. N., & Asghar, Z. (2013). Comparing TL-Moments, L-Moments and Conventional Moments of Dagum Distribution by Simulated data. *Revista Colombiana de Estadística*, 36(1), 79-93.
- Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell system technical journal*, 30(1), 50-64.
- Swain, J. J., Venkatraman, S., & Wilson, J. R. (1988). Least-squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation*, 29(4), 271-297.
- Teimouri, M., Hoseini, S. M., & Nadarajah, S. (2013). Comparison of estimation methods for the Weibull distribution. *Statistics*, 47(1), 93-109.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).