

The Odd Log-Logistic Dagum Distribution: Properties and Applications

La distribución Odd Log-Logística de Dagum: propiedades y aplicaciones

FILIPPO DOMMA^{1,a}, ABBAS EFTEKHARIAN^{2,b}, AHMED Z. AFIFY^{3,c},
MORAD ALIZADEH^{4,d}, INDRANIL GHOSH^{5,e}

¹DEPARTMENT OF ECONOMICS, STATISTICS AND FINANCE, UNIVERSITY OF CALABRIA, RENDE
CS, ITALY

²DEPARTMENT OF STATISTICS, UNIVERSITY OF HORMOZGAN, BANDAR ABBAS, IRAN

³DEPARTMENT OF STATISTICS, MATHEMATICS AND INSURANCE, BENHA UNIVERSITY, BENHA,
EGYPT

⁴DEPARTMENT OF STATISTICS, FACULTY OF SCIENCES, PERSIAN GULF UNIVERSITY, BUSHEHR,
IRAN

⁵DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NORTH CAROLINA
WILMINGTON CITY, UNITED STATES

Abstract

This paper introduces a new four-parameter lifetime model called the odd log-logistic Dagum distribution. The new model has the advantage of being capable of modeling various shapes of aging and failure criteria. We derive some structural properties of the model odd log-logistic Dagum such as order statistics and incomplete moments. The maximum likelihood method is used to estimate the model parameters. Simulation results to assess the performance of the maximum likelihood estimation are discussed. We prove empirically the importance and flexibility of the new model in modeling real data.

Key words: Dagum Distribution; Maximum Likelihood; Odd Log-Logistic Family; Order Statistics.

^aPhD. E-mail: filippo.domma@unical.it

^bPhD. E-mail: ab2005eftekharian@gmail.com

^cPhD. E-mail: ahmed.afify@fcom.bu.edu.eg

^dPhD. E-mail: moradalizadeh78@gmail.com

^ePhD. E-mail: ghoshi@uncw.edu

Resumen

Este artículo introduce un nuevo modelo de sobrevivencia de cuatro parámetros llamado la distribución Odd Log-Logística de Dagum. El nuevo modelo tiene la ventaja de ser capaz de modelar varias formas de envejecimiento y criterios de falla. Derivamos propiedades estructurales del modelo Odd Log-Logística de Dagum tales como estadísticos de orden y momentos incompletos. El método de máxima verosimilitud es usado para estimar los parámetros del modelo. Se discuten resultados de algunas simulaciones, las cuales permitieron establecer la eficiencia del método de máxima verosimilitud. Probamos empíricamente la importancia y flexibilidad del nuevo modelo a través de un ejemplo en datos reales.

Palabras clave: distribución Dagum; familia Odd-Log-Logística; máxima verosimilitud; estadísticas de orden.

1. Introduction

In 1977, the Professor Camilo Dagum has derived the distribution function, which is called Dagum distribution, from a set of assumptions characterizing the observed regularities in the income distributions from both developed and developing countries (Dagum 1977). Then the Dagum distribution has been extensively and successfully used in researches on income and wealth. Based on this distribution, many authors have extensively studied the characteristics and properties of the Dagum model. In addition to, many works published by the same Dagum, we point out the papers of Kleiber (1996, 1999), Quintano & D'Agostino (2006), Domma (2007), Pérez & Alaiz (2011), just to name a few. Kleiber & Kotz (2003) and Kleiber (2008) provided an accurate and in-depth reviews of the genesis of Dagum distribution and its empirical applications.

Domma (2002) showed that the hazard rate of the Dagum distribution, according to the values of parameters, can be monotonically decreasing, upside-down bathtub, and, finally, bathtub then upside-down bathtub. This particular flexibility of the hazard rate has led to several authors apply the Dagum distribution in different fields, such as survival analysis and reliability theory (see, for example, Domma, Giordano & Zenga, 2011, 2012).

In order to increase the flexibility of the Dagum distribution, some authors have proposed different transformations of such distribution. Domma & Condino (2013) introduced a five parameter Beta-Dagum distribution, the special case of four-parameter of Beta Dagum has been used by Domma & Condino (2016) for modeling hydrologic data. Oluyede & Rajasooriya (2013) proposed the six parameter Mc-Dagum Distribution. Among the various transformations found in the literature, we remember the Exponentiated Kumaraswamy Dagum distribution (Huang & Oluyede 2014), the Weighted Dagum (Oluyede & Ye 2014), the transmuted Dagum distribution (Elbatal & Aryal 2015), the Weibull Dagum distribution (Tahir, Cordeiro, Mansoor, Zubair & Alizadeh 2016), the Dagum Poisson Distribution (Oluyede, Motsewabagale, Huang, Warahena-Liyanage & Pararai 2016).

Recently, Gleaton & Lynch (2006) defined a new transformation of distribution function called the odd log-logistic-G (OLL-G) family with one additional shape parameter $\alpha > 0$ by the cumulative distribution function (cdf)

$$F(x; \alpha, \varphi) = \frac{G(x; \varphi)^\alpha}{G(x; \varphi)^\alpha + \overline{G}(x; \varphi)^\alpha}, \quad (1)$$

where $\overline{G}(x; \varphi) = 1 - G(x; \varphi)$. The corresponding probability density function (pdf) of (1) is given by

$$f(x; \alpha, \varphi) = \frac{\alpha g(x; \varphi) G(x; \varphi)^{\alpha-1} \overline{G}(x; \varphi)^{\alpha-1}}{[G(x; \varphi)^\alpha + \overline{G}(x; \varphi)^\alpha]^2}. \quad (2)$$

In general a random variable X with pdf (2) is denoted by $X \sim \text{OLL-G}(\alpha, \varphi)$.

An interpretation of the OLL family (1) can be given as follows. Let T be a random variable describing a stochastic system by the cdf $G(x)$. If the random variable X represents the odds, the risk that the system following the lifetime T will be not working at time x is given by $G(x)/[1 - G(x)]$. If we are interested in modeling the randomness of the odds by the log-logistic pdf $r(t) = \alpha t^{\alpha-1}/(1+t^\alpha)^2$ (for $t > 0$), the cdf of X is given by

$$Pr(X \leq x) = \Pi\left(\frac{G(x)}{1 - G(x)}\right),$$

where $\Pi(t) = \frac{t^\alpha}{1+t^\alpha}$ is the cdf of log-logistic distribution.

For generating data from this distribution, if $u \sim u(0, 1)$ then

$$X_u = G^{-1}\left\{\frac{u^{\frac{1}{\alpha}}}{u^{\frac{1}{\alpha}} + (1-u)^{\frac{1}{\alpha}}}\right\}$$

has cdf (1).

The aim of this paper is to define and study a new lifetime model called the odd log-logistic Dagum (OLLDa) distribution. Its main feature is that one additional shape parameter is inserted in (3) to provide greater flexibility for the generated distribution. Based on the odd log-logistic-G (OLL-G) family of distributions, we construct the four-parameter OLLDa model and give a comprehensive description of some of its mathematical properties in order that it will attract wider applications in reliability, engineering and other areas of research. In fact, the OLLDa model can provide better fits than other competitive models.

This paper is organized as follows. In Section 2, we define the OLLDa distribution. We derive useful mixture representations for the pdf and cdf and give some plots for its pdf and hazard rate function (hrf) in Section 3. We provide in Section 4 some mathematical properties of the OLLDa distribution including quantile function, ordinary and incomplete moments, moment generating function (mgf), mean deviations, probability weighted moments (PWMs), moments of the residual life, reversed residual life and order statistics and their moments are determined. The maximum likelihood estimates (MLEs) of the unknown parameters

are obtained in Section 5. In Section 6, the OLLDa distribution is applied to two real data sets to illustrate its potentiality. Finally, in Section 7, we provide a simulation study.

2. The OLLDa Model

The cdf and pdf of the Dagum distribution are given by (for $x > 0$)

$$G(x; a, b, c) = \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \quad (3)$$

and

$$g(x; a, b, c) = bca^b x^{-b-1} \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c-1}, \quad (4)$$

respectively, where b and c are positive shape parameters and a is a scale parameter.

Using (3) and (1), the cdf and pdf of the OLLDa distribution are given (for $x > 0$) by

$$F(x) = \frac{\left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-\alpha c}}{\left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-\alpha c} + \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha}, \quad (5)$$

and

$$f(x) = \frac{\alpha bca^b x^{-b-1} \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-\alpha c-1} \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^{\alpha-1}}{\left(\left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-\alpha c} + \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha \right)^2}, \quad (6)$$

respectively, where b, c and α are positive shape parameters and $a > 0$ is a scale parameter. If X is a random variable with pdf (6), we denote $X \sim \text{OLLDa}(\alpha, a, b, c)$. For $\alpha = 1$ we obtain Dagum distribution.

The survival function, hrf and cumulative hazard rate function of X are, respectively, given by

$$S(x) = \frac{\left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha}{\left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-\alpha c} + \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha},$$

$$h(x) = \frac{\alpha bca^b x^{-b-1} \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-\alpha c-1} \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^{-1}}{\left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-\alpha c} + \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha},$$

and

$$H(x) = \log \left(\frac{\left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-\alpha c} + \left\{1 - \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c}\right\}^\alpha}{\left\{1 - \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c}\right\}^\alpha} \right).$$

Quantile function is considered as an important quantity in each distribution and is used widely for simulation of that distribution and finding key percentiles. One can obtain the quantile function of OLLDa distribution by inverting (5) as follows

$$Q_X(u) = a \left\{ -1 + \left[\frac{u^{\frac{1}{\alpha}}}{u^{\frac{1}{\alpha}} + (1-u)^{\frac{1}{\alpha}}} \right]^{\frac{-1}{c}} \right\}^{\frac{-1}{b}}. \tag{7}$$

An application of the quantile function is to generate numbers of the distribution. We can easily simulate the OLLDa random variable based on the quantile function. If U is a random variable from Uniform distribution on the unit interval, i.e. $U \sim u(0, 1)$, then the random variable $X = Q_X(U)$ follows (6), that is $X \sim \text{OLLDa}(\alpha, a, b, c)$. In addition, one is interested to analyze the variability of the kurtosis and skewness on the basis of the shape parameters of X . This purpose can be obtained using quantile function and based on the Moors kurtosis (Moors 1988) and Bowley skewness (Kenney & Keeping 1962).

2.1. Asymptotics of the OLL-Da Distribution Function, Density and Hazard Rate

The OLL-Da distribution has interesting asymptotic properties based on the cdf, pdf and hrf. The properties are obtained when either $x \rightarrow 0$ or $x \rightarrow \infty$. In what follows, we present two propositions for asymptotic properties of the OLL-Da distribution.

Proposition 1. *The asymptotics of OLL-Da distribution for cdf, pdf and hrf as $x \rightarrow 0$ are given by*

$$\begin{aligned} F(x) &\sim \left(\frac{x}{a}\right)^{bc\alpha}, \\ f(x) &\sim \frac{bc\alpha x^{bc\alpha-1}}{a^{bc\alpha}}, \\ h(x) &\sim \frac{bc\alpha x^{bc\alpha-1}}{a^{bc\alpha}}. \end{aligned}$$

Proposition 2. *The asymptotics of OLL-Da distribution for cdf, pdf and hrf as $x \rightarrow \infty$ are given by*

$$\begin{aligned} 1 - F(x) &\sim \frac{c^\alpha a^{b\alpha}}{x^{b\alpha}}, \\ f(x) &\sim \frac{b\alpha c^\alpha a^{b\alpha}}{x^{b\alpha+1}}, \\ h(x) &\sim \frac{b\alpha}{x}. \end{aligned}$$

These equations show the effect of parameters on tails of OLL-Da distribution.

3. Plots and Linear Representation

In this section, we provide some plots of the pdf and hrf of the OLLDa model to show its flexibility. Figure 1 displays some plots of the OLLDa density for some parameter values a, b, c and α . Plots of the hrf of the OLLDa model for selected parameter values are given in Figure 2, where the hrf can be bathtub, upside down bathtub (unimodal), increasing, decreasing or constant.

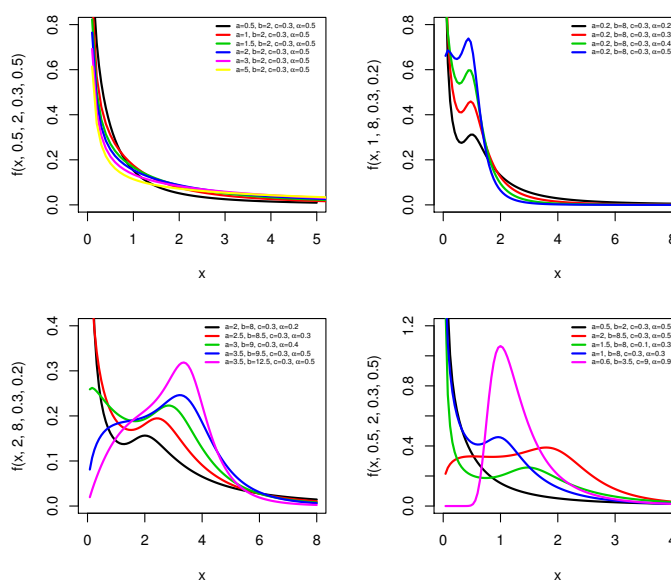


FIGURE 1: Plots of the OLLDa pdf for some parameter values of the first data set.

3.1. Linear Representation

Many authors paid attention to the mixture representations for the pdf and cdf of new family distributions, because using the mixture representations one can easily derive some mathematical properties such as incomplete moments and moments of residual life. To obtain these mixture representations of the OLLDa cdf, we need to the following remark.

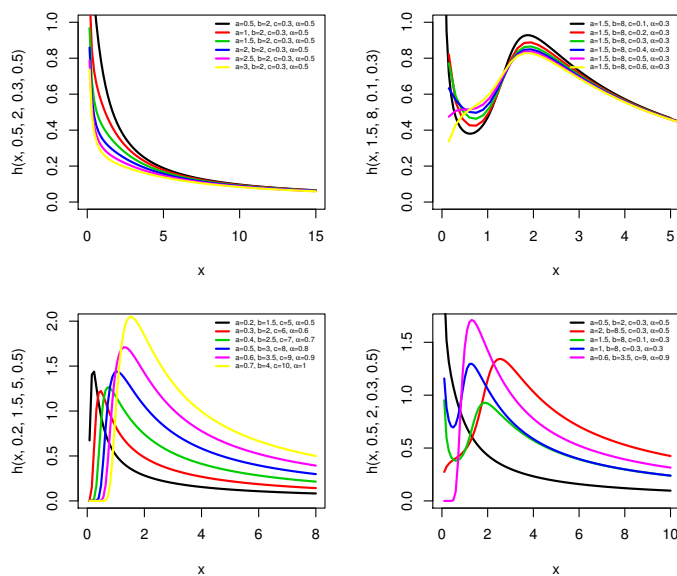


FIGURE 2: Plots of the OLLDa hrf for some parameter values of the first data set.

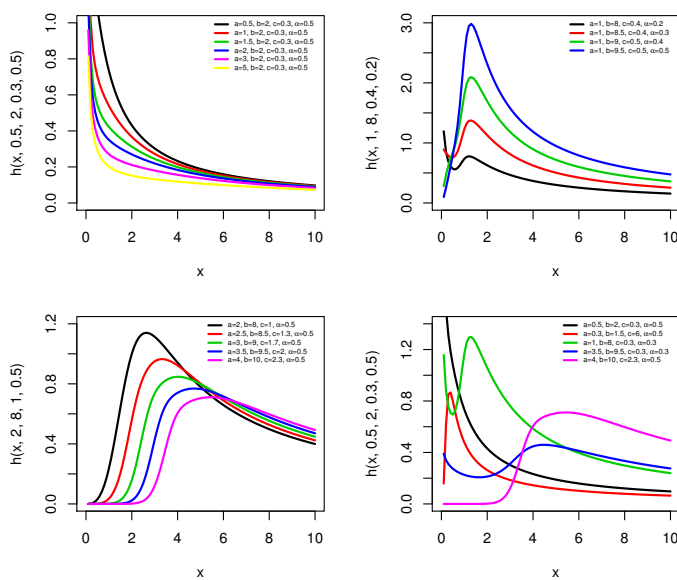


FIGURE 3: Plots of the OLLDa pdf for some parameter values of the second data set.

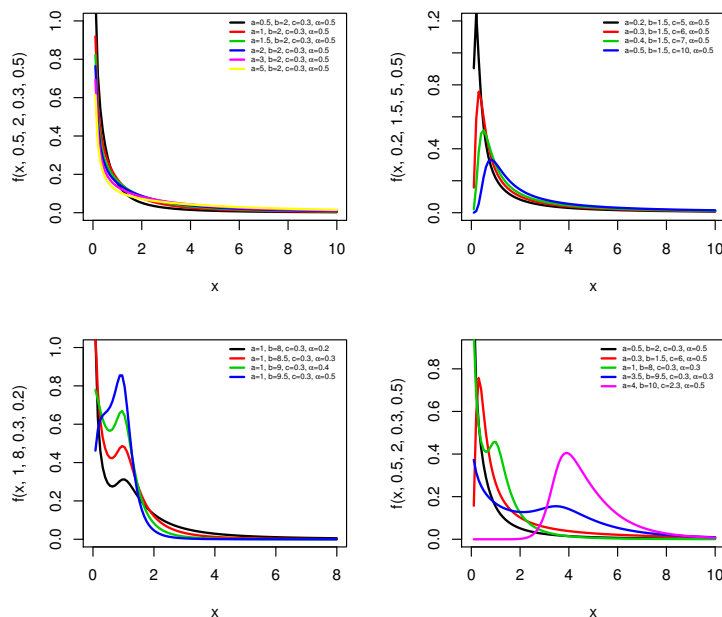


FIGURE 4: Plots of the OLLDa hrf for some parameter values of the second data set.

Remark 1. (Gradshteyn & Ryzhik 2007, p. 17). Division of two power series $\sum_{k=0}^{\infty} a_k x^k$ and $\sum_{k=0}^{\infty} b_k x^k$ is given by

$$\frac{\sum_{k=0}^{\infty} b_k x^k}{\sum_{k=0}^{\infty} a_k x^k} = \sum_{k=0}^{\infty} d_k x^k,$$

where $d_0 = \frac{a_0}{b_0}$ and for $k \geq 1$, we have

$$d_k = \frac{1}{b_0} \left[a_k - \frac{1}{b_0} \sum_{r=1}^k b_r d_{k-r} \right].$$

In what follows, we use (5) and Remark 1 to obtain the mixture representation of the OLLDa cdf. So, we get

$$\begin{aligned}
 \left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha &= \sum_{i=0}^{\infty} (-1)^i \binom{\alpha}{i} \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^i \\
 &= \sum_{i=0}^{\infty} \sum_{k=0}^i (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k} \left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^k \\
 &= \sum_{k=0}^{\infty} \sum_{i=k}^{\infty} (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k} \left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^k \\
 &= \sum_{k=0}^{\infty} a_k \left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^k,
 \end{aligned} \tag{8}$$

where $a_k = \sum_{i=k}^{\infty} (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k}$. Now, using (8) we can write

$$\left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha + \left\{ 1 - \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^\alpha = \sum_{k=0}^{\infty} b_k \left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^k, \tag{9}$$

where $b_k = a_k + (-1)^k \binom{\alpha}{k}$.

Therefore, according to equations (8) and (9) and using Remark 1, it is easy to show that

$$F(x) = \frac{\sum_{k=0}^{\infty} a_k \left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^k}{\sum_{k=0}^{\infty} b_k \left\{ \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-c} \right\}^k} = \sum_{k=0}^{\infty} d_k \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-ck},$$

so,

$$F(x) = \sum_{k=0}^{\infty} d_k G(x; a, b, ck), \tag{10}$$

where $G(x; a, b, ck) = \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-ck}$ denotes the Dagum cdf with parameters a, b and ck .

By differentiating of (10) we have

$$f(x) = \sum_{k=0}^{\infty} d_k g(x; a, b, ck), \tag{11}$$

where $g(x; a, b, ck) = bck a^b x^{-b-1} \left[1 + \left(\frac{x}{a} \right)^{-b} \right]^{-ck-1}$ denotes the pdf of the Dagum distribution with parameters a, b and ck . The mixture representation (11) is useful to calculate the moments and incomplete moments in an immediate way from those properties of the Dagum distribution.

Let Z be a random variable having the Dagum distribution (3) with parameters a, b and c . For $r \leq b$, the r th ordinary and incomplete moments of Z are given by Tahir et al. (2016)

$$\mu'_r = ca^r B(1 - r/b, c + r/b) \quad \text{and} \quad \varphi_r(t) = ca^r B_z(1 - r/b, c + r/b),$$

where $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$ and $B_z(a, b) = \int_0^z t^{a-1} (1-t)^{b-1} dt$ are the complete and incomplete beta functions, respectively. So, several structural properties of the OLLDa model can be obtained from (11) and those are properties of the Dagum distribution.

4. The OLLDa Properties

We investigate mathematical properties of the OLLDa distribution including ordinary and incomplete moments, PWMs and order statistics. It is better to obtain some structural properties of the OLLDa distribution by establishing algebraic expansions than computing those directly by numerical integration of its density function. To this end, we apply the obtained results in the last section.

4.1. Ordinary and Incomplete Moments

The n th ordinary moment of X is given by

$$\mu'_n = E(X^n) = \sum_{k=0}^{\infty} d_k \int_0^{\infty} x^n g(x; a, b, ck) dx.$$

For $n \leq b$, we get

$$\mu'_n = E(X^n) = \sum_{k=0}^{\infty} d_k ck a^n B(1 - n/b, ck + n/b). \tag{12}$$

Setting $n = 1$ in (12), we have the mean of X .

The s th central moment (M_s) and cumulants (κ_s) of X , are given by

$$M_s = E(X - \mu'_1)^s = \sum_{i=0}^s (-1)^i \binom{s}{i} (\mu'_1)^s \mu'_{s-i}$$

and

$$\kappa_s = \mu'_s - \sum_{r=0}^{s-1} \binom{s-1}{r-1} \kappa_r \mu'_{s-r},$$

respectively, where $\kappa_1 = \mu'_1$. The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships.

Using definition of the moment generating function we have

$$M(t) = \sum_{k=0}^{\infty} (k+1) d_{k+1} \int_0^1 u^k \exp [tQ_X(u)] du.$$

The n th incomplete moment, say $\varphi_n(t)$, of the OLLDa distribution is given by $\varphi_n(t) = \int_0^t x^n f(x) dx$. Therefore, we can write from equation (11)

$$\varphi_n(t) = \sum_{k=0}^{\infty} d_k \int_0^t x^n g(x; a, b, ck) dx,$$

and then using the lower incomplete gamma function, we obtain (for $n \leq b$)

$$\varphi_n(t) = \sum_{k=0}^{\infty} d_k c k a^n B_z(1 - n/b, ck + n/b). \tag{13}$$

The first incomplete moment of X , denoted by $\varphi_1(t)$, is simply determined from the equation (13) by setting $s = 1$. It should be noted that the n th incomplete moment of X can be determined on the basis of the quantile function

$$\varphi_n(t) = \int_0^t x^n f(x) dx = \sum_{k=0}^{\infty} (k+1) d_{k+1} \int_0^{[1+(\frac{t}{a})^{-b}]^{-c}} u^k Q_G(u)^n du, \tag{14}$$

but using (7) and properties of power series (Gradshteyn & Ryzhik 2007, p.17), we can write

$$\varphi_n(t) = \sum_{k,i=0}^{\infty} \frac{(k+1) d_{k+1} w_i^*}{i+k+1} \left\{ \left[1 + \left(\frac{t}{a} \right)^{-b} \right]^{-c} \right\}^{i+k+1},$$

where $\omega_0^* = \omega_0^n$, $\omega_m^* = \frac{1}{m\omega_0} \sum_{j=1}^m (j(n+1) - m)\omega_j \omega_{m-j}^*$, for $m \geq 1$. The first incomplete moment has important applications related to the Bonferroni and Lorenz curves and the mean residual life and the mean waiting time. Furthermore, the amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. The mean deviations, about the mean and about the median of X , depend on $\varphi_1(t)$.

4.2. Residual and Reversed Residual Lives

The n th moment of the residual life, denoted by $m_n(t) = E[(X - t)^n | X > t]$ for $t > 0$ and $n = 1, 2, \dots$, uniquely determines $F(x)$ (see Navarro, Franco & Ruiz 1998). The n th moment of the residual life of X is given by

$$m_n(t) = \frac{1}{1 - F(t)} \int_t^{\infty} (x - t)^n dF(x). \tag{15}$$

Based on (15) we can write

$$m_n(t) = \frac{1}{R(t)} \sum_{i=0}^n \sum_{k=0}^{\infty} \frac{(-1)^{n-i} t^{n-i}}{i!} (n)_i d_k c k a^i B_z(1 - i/b, ck + i/b),$$

where $R(t) = 1 - F(t)$ and $(n)_i = \Gamma(n+1)/\Gamma(n-i+1)$ is the falling factorial.

Another interesting function is the mean residual life function or the life expectation at age x defined by $m_1(t) = E[(X - t) | X > t]$, which represents the expected additional life length for a unit which is alive at age t . The mean residual life of the OLLDa distribution can be obtained by setting $n = 1$ in the last equation.

Navarro et al. (1998) proved that the n th moment (for $n = 1, 2, \dots$) of the reversed residual life, say $M_n(t) = E[(t - X)^n | X \leq t]$ for $t > 0$, uniquely determines $F(x)$. By definition, the n th moment of the reversed residual life is given by

$$M_n(t) = \frac{1}{F(t)} \int_0^t (t - x)^n dF(x).$$

So, it is clear that the n th moment of the reversed residual life of X is as follows

$$M_n(t) = \frac{1}{F(t)} \sum_{i=0}^n \sum_{k=0}^{\infty} \frac{(-1)^i}{i!} (n)_i d_k c k a^i t^{n-i} B_z(1 - i/b, ck + i/b).$$

The mean inactivity time or mean waiting time, also called the mean reversed residual life function, say $M_1(t) = E[(t - X) | X \leq t]$, represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in $(0, t)$. The mean inactivity time of X can be obtained by setting $n = 1$ in the above equation.

4.3. Mean deviations and Bonferroni and Lorenz Curves

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. If X has the OLLDa distribution, we can derive the mean deviations about the mean $\nu = E(X)$ and about the median M from the relations $\delta_1 = \int_0^\infty |x - \nu| f(x) dx = 2\nu F(\nu) - 2\varphi_1(\nu)$ and $\delta_2 = \int_0^\infty |x - M| f(x) dx = \nu - 2\varphi_1(M)$, respectively, where $F(\nu)$ is simply calculated from (5), $\varphi_1(\nu)$ is the first incomplete moment given by

$$\varphi_1(\nu) = \sum_{k=0}^{\infty} d_k c k a B_z(1 - 1/b, ck + 1/b). \quad (16)$$

Bonferroni, $B(p)$, and Lorenz, $L(p)$, curves have applications in economics, to study income and poverty, reliability, demography, insurance and medicine. They are defined by $B(p) = \varphi_1(q)/p\nu$ and $L(p) = \varphi_1(q)/\nu$, respectively, where $q = Q(p) = a(-1 + \{p^{1/\alpha}/[p^{1/\alpha} + (1-p)^{1/\alpha}]\}^{-1/c})^{-1/b}$. These measures can be calculated immediately from equation (16) evaluated at $\nu = q$.

4.4. Probability Weighted Moments

The PWMs are used to derive estimators of the parameters and quantiles of generalized distributions. These moments have low variance and no severe bias, and they compare favorably with estimators obtained by the maximum likelihood method. The (s, r) th PWM of X (for $r \geq 1, s \geq 0$) is formally defined by

$$\rho_{r,s} = E[X^r F(X)^s] = \int_0^\infty x^r F(x)^s f(x) dx.$$

Using equation (10), we can write

$$f(x)F(x)^s = \sum_{k,l=0}^{\infty} q_{k,l} [\alpha(s+k+1)+l] b c a^b x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c-1} \left\{ \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha(s+k+1)+l-1},$$

where

$$q_{k,l} = (-1)^k \binom{s+k+1}{k} \frac{\Gamma(\alpha(s+k+1)+l)}{l! \Gamma(\alpha(s+k+1)+1)}.$$

The last equation can be rewritten as follows

$$f(x)F(x)^s = \sum_{k,l=0}^{\infty} q_{k,l} g(x; a, b, c [\alpha(s+k+1)+l]).$$

Then, we have

$$\rho_{r,s} = \sum_{k,l=0}^{\infty} q_{k,l} \int_0^{\infty} x^r g(x; a, b, c [\alpha(s+k+1)+l]) dx.$$

The above equation reveals that the (s, r) th PWM of X is a linear combination of the Da densities. So, we can write

$$\rho_{r,s} = a^r \sum_{k,l=0}^{\infty} q_{k,l} c [\alpha(s+k+1)+l] B(1-r/b, c [\alpha(s+k+1)+l] + r/b).$$

4.5. Order Statistics

In this part, we obtain the pdf of the i th order statistic from a random sample with size n . We denote the i th order statistic and corresponding pdf from OLLD distribution by $X_{i:n}$ and $f_{i:n}(x)$, respectively. By suppressing the parameters, we have (for $i = 1, 2, \dots, n$)

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1},$$

but based on the equations (5) and (6) we have

$$\begin{aligned}
 f_{i:n}(x) &= \frac{\alpha b c a^b x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c-1}}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \\
 &\quad \times \frac{\left\{ \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha(i+j)-1} \left\{ 1 - \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha-1}}{\left[\left\{ \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha} + \left\{ 1 - \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha} \right]^{i+j+1}} \\
 &= \frac{\alpha b c a^b x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c-1}}{B(i, n-i+1)} \sum_{j=0}^{n-i} \sum_{k=0}^{\infty} (-1)^{(j+k)} \binom{n-i}{j} \binom{i+j+k}{k} \\
 &\quad \times \left\{ \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha(i+j+k)-1} \left\{ 1 - \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{-\alpha(i+j+k)-1} \\
 &= \frac{\alpha b c a^b x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c-1}}{B(i, n-i+1)} \sum_{j=0}^{n-i} \sum_{k,l=0}^{\infty} (-1)^{(j+k)} \binom{n-i}{j} \binom{i+j+k}{k} \\
 &\quad \times \binom{\alpha(i+j+k)+l}{l} \left\{ \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha(i+j+k)+l-1},
 \end{aligned}$$

and then

$$\begin{aligned}
 f_{i:n}(x) &= \sum_{k,l=0}^{\infty} w_{k,l} [\alpha(i+j+k)+l] b c a^b x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c-1} \\
 &\quad \left\{ \left[1 + \left(\frac{x}{a}\right)^{-b}\right]^{-c} \right\}^{\alpha(i+j+k)+l-1}, \tag{17}
 \end{aligned}$$

where

$$w_{k,l} = \frac{1}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^{j+k} \binom{n-i}{j} \binom{i+j+k}{k} \frac{\Gamma(\alpha(i+j+k)+l)}{l! \Gamma(\alpha(i+j+k)+1)}.$$

Using (17), it can be easily shown that

$$f_{i:n}(x) = \sum_{k,l=0}^{\infty} w_{k,l} g(x; a, b, c [\alpha(i+j+k)+l]).$$

It is clear that the density function of the i th order statistic from the OLLD distribution is a linear combination of the Dagum densities. So, we can easily obtain the mathematical properties for $X_{i:n}$. For example, the r th moment of $X_{i:n}$ follows from

$$E(X_{i:n}^r) = a^r \sum_{k,l=0}^{\infty} w_{k,l} c [\alpha(i+j+k)+l] B(1-r/b, c[\alpha(i+j+k)+l] + r/b).$$

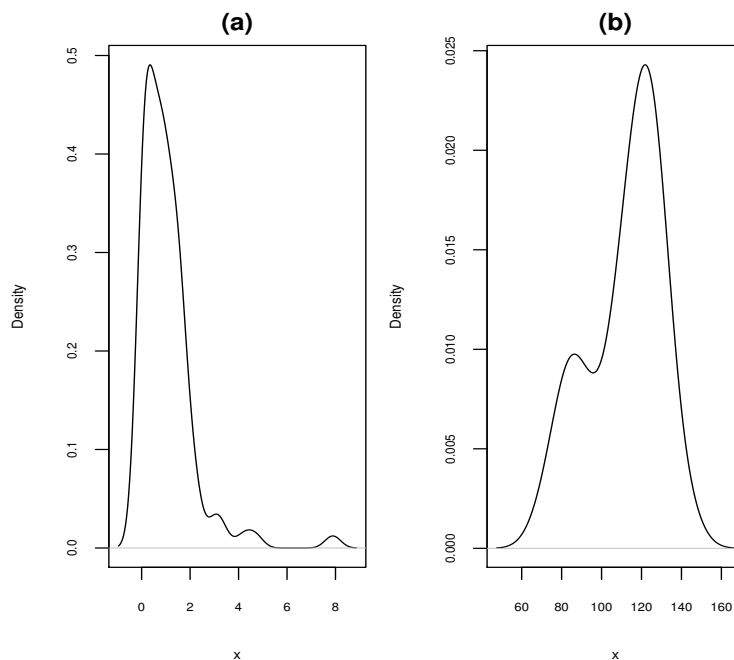


FIGURE 5: Gaussian kernel density estimation for: (a) the failure times of Kevlar and (b) birth weights of newborn babies.

5. Estimation

Several methods to estimate parameters were proposed in the literature, but the most popular method is maximum likelihood method. Based on this method, one obtains the maximum likelihood estimations (MLEs) for the parameters. After that, using asymptotic property of MLE, it is easy to derive a $100(1 - \gamma)\%$ approximate confidence interval for each parameter. Here, for simplify, we use $g(x)$ and $G(x)$ as the pdf and cdf of Dagum distribution and obtain the MLEs of parameters for OLLDa(α, a, b, c) distribution. To this purpose, let X_1, X_2, \dots, X_n be a random sample of size n from the OLLDa distribution given by (6). The log-likelihood function of the vector parameters $\theta = (\alpha, a, b, c)$ is as follows

$$\ell(\theta) = n \ln \alpha + \sum_{i=1}^n \{ \ln g(x_i) + (\alpha - 1) [\ln G(x_i) + \ln \bar{G}(x_i)] - 2 \ln [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] \}.$$

It can be shown that

$$\dot{\ell}_\alpha(\theta) = \frac{n}{\alpha} + \sum_{i=1}^n [\ln G(x_i) + \ln \bar{G}(x_i)] - 2 \sum_{i=1}^n \frac{G(x_i)^\alpha \ln G(x_i) + \bar{G}(x_i)^\alpha \ln \bar{G}(x_i)}{G(x_i)^\alpha + \bar{G}(x_i)^\alpha},$$

$$\dot{\ell}_a = \sum_{i=1}^n \frac{\dot{g}_a(x_i)}{g(x_i)} + \sum_{i=1}^n \dot{G}_a(x_i) \left\{ (\alpha - 1) \frac{1 - 2G(x_i)}{G(x_i)\bar{G}(x_i)} - 2\alpha \frac{G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}}{G(x_i)^\alpha + \bar{G}(x_i)^\alpha} \right\},$$

$$\dot{\ell}_b = \sum_{i=1}^n \frac{\dot{g}_b(x_i)}{g(x_i)} + \sum_{i=1}^n \dot{G}_b(x_i) \left\{ (\alpha - 1) \frac{1 - 2G(x_i)}{G(x_i)\bar{G}(x_i)} - 2\alpha \frac{G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}}{G(x_i)^\alpha + \bar{G}(x_i)^\alpha} \right\},$$

$$\dot{\ell}_c = \sum_{i=1}^n \frac{\dot{g}_c(x_i)}{g(x_i)} + \sum_{i=1}^n \dot{G}_c(x_i) \left\{ (\alpha - 1) \frac{1 - 2G(x_i)}{G(x_i)\bar{G}(x_i)} - 2\alpha \frac{G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}}{G(x_i)^\alpha + \bar{G}(x_i)^\alpha} \right\},$$

where $\dot{\ell}_j(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial j}$, $\dot{g}_j(x) = \frac{\partial g(x)}{\partial j}$ and $\dot{G}_j(x) = \frac{\partial G(x)}{\partial j}$ for any $j = \alpha, a, b, c$. For more details, see Appendix.

The MLEs can be obtained using the solutions of above equations, but these equations are nonlinear and one can be solved them iteratively.

On the other hand, to obtain the $100(1 - \gamma)\%$ approximate confidence interval for each parameter, we need to the Fisher information matrix and for this purpose, we should obtain the following expectation

$$I(\boldsymbol{\Theta}) = -E \left[\ddot{\ell}(\boldsymbol{\theta}) \right],$$

where

$$\ddot{\ell}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \begin{bmatrix} \ddot{\ell}_{\alpha\alpha} & \ddot{\ell}_{\alpha a} & \ddot{\ell}_{\alpha b} & \ddot{\ell}_{\alpha c} \\ \cdot & \ddot{\ell}_{aa} & \ddot{\ell}_{ab} & \ddot{\ell}_{ac} \\ \cdot & \cdot & \ddot{\ell}_{bb} & \ddot{\ell}_{bc} \\ \cdot & \cdot & \cdot & \ddot{\ell}_{cc} \end{bmatrix}$$

with considering $\ddot{\ell}_{ij} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial i \partial j}$. The $\ddot{\ell}(\boldsymbol{\theta})$ is in Appendix.

6. Application

In this section, we show that the OLLDa distribution can be a better model than the Dagum, Beta Burr XII (BBXII) (Paranaíba, Ortega, Cordeiro & Pescim 2011), Beta Dagum (BDa) Domma & Condino (2013), Mc-Dagum (Mc-Da) Oluyede & Rajasooriya (2013) and Dagum-Poisson (DaP) Oluyede et al. (2016) distributions using two real data sets. To this end, we compare various models using criteria like log likelihood, Akaike information criterion (AIC), Cramer-von Mises (W^*) and Anderson-Darling (A^*) statistics based on the data sets. Cramer-von Mises (W^*) and Anderson-Darling (A^*) statistics are defined as follows Evans, Drew & Leemis (2008)

$$W^* = \sum_{i=1}^n \left(\hat{F}(x_{(i)}) - \frac{i - 0.5}{n} \right)^2 + \frac{1}{12n}$$

and

$$A^* = - \sum_{i=1}^n \frac{2i-1}{n} \left(\ln \left(\hat{F}(x_{(i)}) \right) + \ln \left(1 - \hat{F}(x_{(n+1-i)}) \right) \right) - n.$$

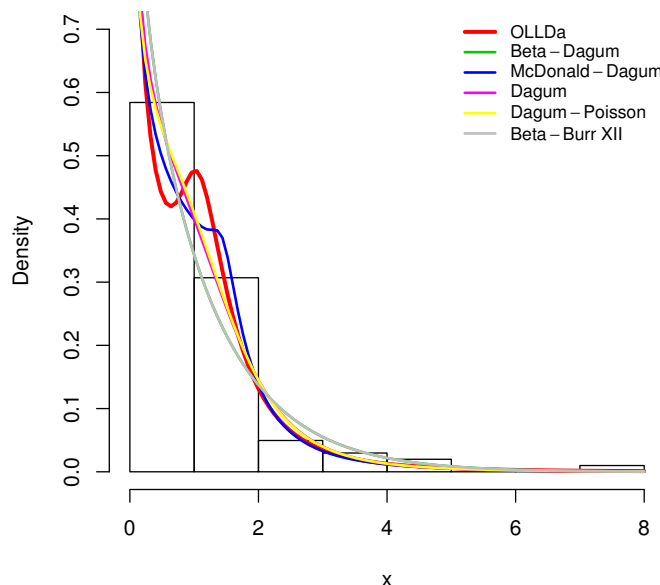


FIGURE 6: Fitted pdf for the first data set.

If a model has smaller values of -log likelihood, AIC, W^* and A^* , then it is better than other models.

The numerical results of MLEs and their corresponding standard errors (in parentheses) for the parameters of model are reported in Tables 1 and 2 for the first and second data sets, respectively. Figure 5(a) and (b) displays the fitted densities for the first and second data sets obtained using kernel density estimation based on Gaussian kernel function, respectively. Suppose that X_1, \dots, X_n is a random sample of independent and identically distributed random variables with an unknown pdf f . Then the kernel density estimator of f is derived by

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x - x_i}{h}\right),$$

where $k(\cdot)$ is the kernel function and h is a smoothing parameter or bandwidth. Silverman (1986) is an appropriate reference for more details of kernel estimation and its properties. In Figure 9, we have considered the standard normal as kernel function, $k(\cdot)$, with bandwidth $h = (4\hat{\sigma}/3n)^{1/5}$, where $\hat{\sigma}$ is the standard deviation of the interested sample.

In addition, the plots of the fitted distributions for the data sets are exhibited in Figures 6 and 8. These results illustrate that the OLLDa model is superior.

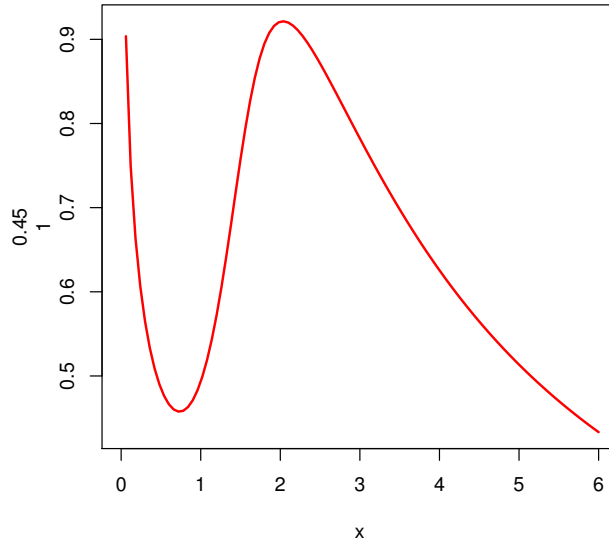


FIGURE 7: Fitted hef for the first data set.

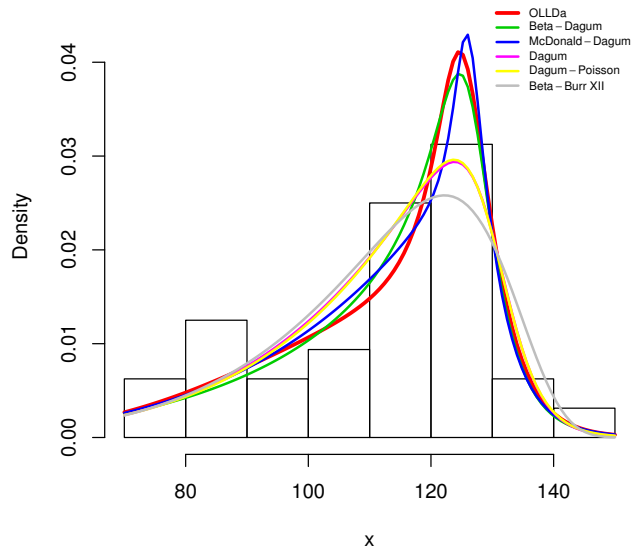


FIGURE 8: Fitted pdf for the second data set.

TABLE 1: MLEs of the parameters (standard errors in parentheses) and goodness-of-fit statistics for the first data set.

Model	Estimates			$-\ell(\hat{\theta})$	W^*	A^*	AIC	
$OLLDa(\alpha, a, b, c)$	0.32075 (0.19761)	1.05828 (0.21205)	8.26769 (4.43169)	0.30079 (0.07867)	97.80727	0.03108	0.24482	203.6145
$BDa(\alpha, \beta, a, b, c)$	0.04739 (0.00591)	0.13966 (0.02196)	1.45572 (0.02180)	18.86106 (0.03052)	0.75390 (0.05289)	0.03848	0.30413	206.545
$DaP(a, b, c, \lambda)$	1.88924 (0.24366)	3.38728 (0.67200)	0.00246 (0.30829)	87.75564 (2.47235)	100.1593	0.07384	0.49855	208.3186
$Da(a, b, c)$	1.90091 (0.24583)	3.42692 (0.69055)	0.21099 (0.05484)		100.1069	0.07416	0.49795	206.2137
$Mc - Da(a, b, c, \alpha, \beta, \delta)$	1.45268 (0.00100)	20.19398 (0.00101)	0.97457 (0.00140)	0.04254 (0.00457)	98.13383	0.03977	0.30759	207.8895
$BBXII(a, b, c, \alpha, \beta)$	1.74414 (0.01351)	6.52824 (0.01567)	0.17468 (0.01381)	0.10716 (0.01127)	98.52696	0.05317	0.37154	207.0539

TABLE 2: MLEs of the parameters (standard errors in parentheses) and goodness-of-fit statistics for the second data set.

Model	Estimates			$-\ell(\hat{\theta})$	W^*	A^*	AIC	
$OLLDa(\alpha, a, b, c)$	0.41815 (0.28662)	124.20145 (3.50074)	74.47492 (49.74483)	0.18921 (0.09041)	133.6301	0.02556	0.21589	275.2602
$BDa(\alpha, \beta, a, b, c)$	28.51289 (86.0744)	2.66105 (1.17964)	146.01419 (5.31973)	141.42377 (403.8166)	135.2869	0.08438	0.54244	280.5738
$DaP(a, b, c, \theta)$	130.7950 (2.79497)	39.06672 (15.33588)	0.00205 (0.000253)	70.93378 (34.11743)	134.465	0.03926	0.30550	276.9332
$Da(a, b, c)$	130.89813 (2.82271)	39.26002 (15.59890)	0.14302 (0.06822)		134.4357	0.03992	0.30764	274.8713
$Mc - Da(a, b, c, \alpha, \beta, \delta)$	125.70633 (1.85821)	145.38390 (18.87298)	0.18310 (1.54146)	0.08375 (0.06321)	133.5487 (19.24204)	0.02544	0.22320	279.0974
$BBXII(a, b, c, \alpha, \beta)$	263.627601 (11.18899)	16.36626 (8.97002)	6.779.20227 (909.67053)	0.35600 (0.25964)	134.971	0.05523	0.40027	279.9421

The sample size of the first data set equals to 101 observations and the data set has been taken based on the failure times of Kevlar 49/epoxy strands with pressure at 90%. It has been originally presented by Barlow, Toland & Freeman (1984) in which failure times are in hours. The data set has been previously analyzed by Andrews & Herzberg (2012) and Cooray & Ananda (2008). Recently, Oluyede et al. (2016) used it to fit based on the Dagum-Poisson distribution.

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.

The second data set contains 32 observations corresponding to the birth weights of newborn babies in ounces (from Armitage & Berry, 1987). Some authors studied this data set and fitted some models for the data set such as Thode (2002, p. 342) and Torabi, Montazeri & Grané (2016).

72, 80, 81, 84, 86, 87, 92, 94, 103, 106, 107, 111, 112, 115, 116, 118, 119, 122, 123, 123, 114, 125, 126, 126, 126, 127, 118, 128, 128, 132, 133, 142.

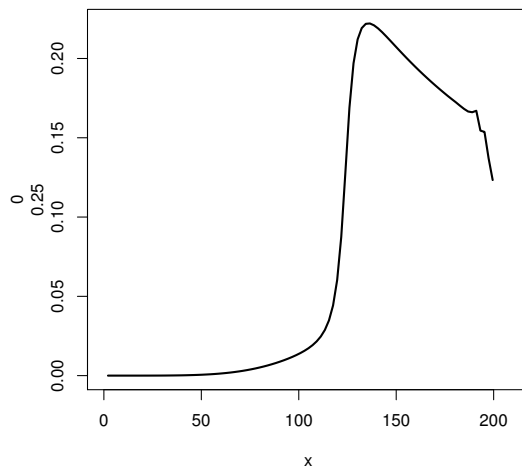


FIGURE 9: Fitted hrf for the second data set.

7. Simulation Study

In order to assess the performance of the MLEs, a small simulation study is performed using the statistical software R through the package (stats4), command

MLE. The number of Monte Carlo replications was 20,000. For maximizing the log-likelihood function, we use the MaxBFGS subroutine with analytical derivatives. The evaluation of the estimates was performed based on the following quantities for each sample size: the empirical mean squared errors (MSEs) are calculated using the R package from the Monte Carlo replications. The MLEs are determined for each simulated data, say, $(\hat{\alpha}_i, \hat{\beta}_i, \hat{a}_i, \hat{b}_i)$ for $i = 1, 2, \dots, 10,000$ and the biases and MSEs are computed by

$$bias_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{h}_i - h),$$

and

$$MSE_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{h}_i - h)^2,$$

for $h = a, b, c, \alpha$. We consider the sample sizes at $n = 100, 150$ and 300 and consider different values for the parameters. The empirical results are given in Table 3. The figures in Table 3 indicate that the estimates are quite stable and, more importantly, are close to the true values for the these sample sizes. Furthermore, as the sample size increases, the MSEs decreases as expected.

TABLE 3: Bias and MSE of the estimates under the maximum likelihood method.

Sample Size	Actual Value				Bias				MSE			
	<i>a</i>	<i>b</i>	<i>c</i>	α	\hat{a}	\hat{b}	\hat{c}	$\hat{\alpha}$	\hat{a}	\hat{b}	\hat{c}	$\hat{\alpha}$
100	0.5	0.5	2	4	-0.4173	-0.4198	0.3554	-1.4923	0.0518	0.0460	0.0538	0.972
	0.5	0.5	3	5	-0.7738	0.3239	-0.2145	-0.3426	0.0180	0.0426	0.0978	0.626
	0.7	0.8	4	3	0.4891	-0.2460	-0.6227	0.4821	0.0154	0.1207	0.1065	0.167
	0.9	0.7	6	4	0.1883	0.9799	-0.5091	0.0563	0.0481	0.0222	0.0442	0.1138
	1	1.5	0.9	0.6	0.1780	-0.4983	-0.4292	-0.5454	0.4271	0.0282	0.0917	0.4956
	1.5	2	0.6	0.8	-0.0841	-0.3634	-0.4059	-0.2203	0.9537	0.0183	0.0737	0.572
150	0.5	0.5	2	4	0.0717	0.3617	0.0499	0.0729	0.0216	0.0230	0.0248	0.3132
	0.5	0.5	3	5	0.5182	0.1845	0.0843	0.1153	0.0084	0.0229	0.0457	0.5781
	0.7	0.8	4	3	0.3165	0.1594	0.0505	-0.3298	0.0065	0.0591	0.0441	0.1587
	0.9	0.7	6	4	0.1375	-0.0493	0.1309	-0.0318	0.0181	0.0105	0.0203	0.0955
	1	1.5	0.9	0.6	0.1258	0.4753	0.0866	0.2424	0.0646	0.0131	0.0281	0.1472
	1.5	2	0.6	0.8	0.0343	0.1737	0.2243	-0.1509	0.4015	0.0089	0.0367	0.432
300	0.5	0.5	2	4	-0.04609	-0.02898	-0.0368	-0.0475	0.0094	0.0114	0.0111	0.0842
	0.5	0.5	3	5	-0.0512	-0.1110	-0.0355	0.0023	0.0041	0.0102	0.0229	0.0182
	0.7	0.8	4	3	-0.0730	-0.0527	0.0467	-0.02271	0.0039	0.0315	0.0241	0.0364
	0.9	0.7	6	4	-0.1023	-0.0208	-0.0986	-0.0235	0.0086	0.0053	0.0102	0.0217
	1	1.5	0.9	0.6	-0.0783	-0.0519	-0.0169	0.0026	0.0270	0.0071	0.0123	0.0154
	1.5	2	0.6	0.8	0.0078	-0.0691	0.0660	-0.0832	0.1345	0.0049	0.0128	0.01142

Acknowledgments

The authors express their sincere thanks to the anonymous referees and the associate editor for their useful comments and constructive criticisms on the original version of this manuscript, which led to this considerably improved version.

[Received: July 2017 — Accepted: January 2018]

References

- Andrews, D. F. & Herzberg, A. M. (2012), *Data: a collection of problems from many fields for the student and research worker*, Springer Science & Business Media.
- Armitage, P. & Berry, G. (1987), Distribution-free methods, in ‘Statistical methods in medical research’, Blackwell Scientific Oxford, London, Edinburgh, Boston, Melbourne, pp. 408–420.
- Barlow, R., Toland, R. & 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’.
- Cooray, K. & Ananda, M. M. (2008), ‘A generalization of the half-normal distribution with applications to lifetime data’, *Communications in Statistics-Theory and Methods* **37**(9), 1323–1337.
- Da Cruz, J., Cordeiro, G., Ortega, E. & Alizadeh, M. (2014), ‘A new family of distributions: Theory and applications in survival analysis’, *Statistics* .
- Dagum, C. (1977), ‘New model of personal income-distribution-specification and estimation’, *Economie appliquée* **30**(3), 413–437.
- Domma, F. (2002), ‘L’andamento della hazard function nel modello di dagum a tre parametri’, *Quaderni di Statistica* **4**, 1–12.
- Domma, F. (2007), ‘Asymptotic distribution of the maximum likelihood estimators of the parameters of the right-truncated dagum distribution’, *Communications in Statistics-Simulation and Computation* **36**(6), 1187–1199.
- Domma, F. & Condino, F. (2013), ‘The beta-dagum distribution: definition and properties’, *Communications in Statistics-Theory and Methods* **42**(22), 4070–4090.
- Domma, F. & Condino, F. (2016), ‘Use of the beta-dagum and beta-singh-maddala distributions for modeling hydrologic data’, *Stochastic Environmental Research and Risk Assessment* pp. 1–15.
- Domma, F., Giordano, S. & Zenga, M. (2011), ‘Maximum likelihood estimation in dagum distribution with censored samples’, *Journal of Applied Statistics* **38**(12), 2971–2985.
- Domma, F., Latorre, G. & Zenga, M. (2012), ‘The dagum distribution in reliability analysis.’, *Statistica & Applicazioni* **10**(2).
- Elbatal, I. & Aryal, G. (2015), ‘Transmuted dagum distribution with applications’, *Chilean Journal of Statistics (ChJS)* **6**(2).

- Evans, D. L., Drew, J. H. & Leemis, L. M. (2008), 'The distribution of the kolmogorov-smirnov, cramer-von mises, and anderson-darling test statistics for exponential populations with estimated parameters', *Communications in Statistics-Simulation and Computation* **37**(7), 1396–1421.
- Gleaton, J. & Lynch, J. (2006), 'Properties of generalized log-logistic families of lifetime distributions', *Journal of Probability and Statistical Science* **4**(1), 51–64.
- Gradshteyn, I. & Ryzhik, I. (2007), *Table of Integrals, Series, and Products*, Elsevier/Academic Press.
- Huang, S. & 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.
- Kenney, J. & Keeping, E. (1962), *Mathematics of Statistics*, Vol. 1, Princeton, NJ: Van Nostrand.
- Kleiber, C. (1996), 'Dagum vs. singh-maddala income distributions', *Economics Letters* **53**(3), 265–268.
- Kleiber, C. (1999), 'On the lorenz order within parametric families of income distributions', *Sankhyā: The Indian Journal of Statistics, Series B* pp. 514–517.
- Kleiber, C. (2008), A guide to the dagum distributions, in 'Modeling Income Distributions and Lorenz Curves', Springer, pp. 97–117.
- Kleiber, C. & Kotz, S. (2003), *Statistical size distributions in economics and actuarial sciences*, Vol. 470, John Wiley & Sons.
- Moors, J. (1988), 'A quantile alternative for kurtosis', *The Statistician* pp. 25–32.
- Navarro, J., Franco, M. & Ruiz, J. (1998), 'Characterization through moments of the residual life and conditional spacings', *Sankhyā: The Indian Journal of Statistics, Series A* pp. 36–48.
- Oluyede, B. O., Motsewabagale, G., Huang, S., Warahena-Liyanage, G. & Pararai, M. (2016), 'The dagum-poisson distribution: model, properties and application', *Electronic Journal of Applied Statistical Analysis* **9**(1), 169–197.
- Oluyede, B. O. & Rajasooriya, S. (2013), 'The mc-dagum distribution and its statistical properties with applications', *Asian Journal of Mathematics and Applications* **2013**.
- Oluyede, B. O. & Ye, Y. (2014), 'Weighted dagum and related distributions', *Afrika Matematika* **25**(4), 1125–1141.
- Paranaíba, P. F., Ortega, E. M., Cordeiro, G. M. & Pescim, R. R. (2011), 'The beta burr xii distribution with application to lifetime data', *Computational Statistics & Data Analysis* **55**(2), 1118–1136.

Pérez, C. G. & Alaiz, M. P. (2011), ‘Using the dagum model to explain changes in personal income distribution’, *Applied Economics* **43**(28), 4377–4386.

Quintano, C. & D’Agostino, A. (2006), ‘Studying inequality in income distribution of single-person households in four developed countries’, *Review of Income and Wealth* **52**(4), 525–546.

Silverman, B. W. (1986), *Density estimation for statistics and data analysis*, Vol. 26, CRC press.

Tahir, M., Cordeiro, G. M., Mansoor, M., Zubair, M. & Alizadeh, M. (2016), ‘The weibull-dagum distribution: properties and applications’, *Communications in Statistics-Theory and Methods* (just-accepted).

Thode, H. C. (2002), *Testing for normality*, Vol. 164, CRC press.

Torabi, H., Montazeri, N. H. & Grané, A. (2016), ‘A test for normality based on the empirical distribution function’, *SORT* **40**(1), 3–36.

Appendix

In this section, we report some needed derivatives in Section 5.

$$\dot{g}_a(x) = b^2 c a^{b-1} x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b} \right]^{-c-2} \left[1 - c \left(\frac{x}{a}\right)^{-b} \right]$$

$$\dot{g}_b(x) = c a^b x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b} \right]^{-c-2} \left\{ 1 + \left(\frac{x}{a}\right)^{-b} + b \ln \left(\frac{x}{a}\right) \left[(c+1) \left(\frac{x}{a}\right)^{-b} - 1 \right] \right\}$$

$$\dot{g}_c(x) = b a^b x^{-b-1} \left[1 + \left(\frac{x}{a}\right)^{-b} \right]^{-c-1} \left\{ 1 - c \ln \left[1 + \left(\frac{x}{a}\right)^{-b} \right] \right\}$$

$$\dot{G}_a(x) = -c b a^{b-1} x^{-b} \left[1 + \left(\frac{x}{a}\right)^{-b} \right]^{-c-1}$$

$$\dot{G}_b(x) = c \left(\frac{x}{a}\right)^{-b} \left[1 + \left(\frac{x}{a}\right)^{-b} \right]^{-c-1} \ln \left(\frac{x}{a}\right)$$

$$\dot{G}_c(x) = - \left[1 + \left(\frac{x}{a}\right)^{-b} \right]^{-c} \ln \left[1 + \left(\frac{x}{a}\right)^{-b} \right]$$

$$\begin{aligned} \ddot{\ell}_{\alpha\alpha}(\boldsymbol{\theta}) = & -\frac{n}{\alpha^2} - 2 \sum_{i=1}^n \frac{1}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \left\{ [G(x_i)^\alpha (\ln G(x_i))^2 + \bar{G}(x_i)^\alpha (\ln \bar{G}(x_i))^2] \right. \\ & \left. \times [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] - [G(x_i)^\alpha \ln G(x_i) + \bar{G}(x_i)^\alpha \ln \bar{G}(x_i)]^2 \right\}, \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{\alpha a}(\boldsymbol{\theta}) = & \sum_{i=1}^n \left[\frac{\dot{G}_a(x_i)}{G(x_i)} - \frac{\dot{G}_a(x_i)}{\bar{G}(x_i)} \right] - 2 \sum_{i=1}^n \frac{\dot{G}_a(x_i)}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \left\{ \left[\alpha G(x_i)^{\alpha-1} \ln G(x_i) \right. \right. \\ & \left. \left. + G(x_i)^{\alpha-1} - \alpha \bar{G}(x_i)^{\alpha-1} \ln \bar{G}(x_i) - \bar{G}(x_i)^{\alpha-1} \right] [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] \right. \\ & \left. - \alpha [G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}] [G(x_i)^{\alpha-1} \ln G(x_i) - \bar{G}(x_i)^{\alpha-1} \ln \bar{G}(x_i)] \right\}, \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{\alpha b}(\boldsymbol{\theta}) = & \sum_{i=1}^n \left[\frac{\dot{G}_b(x_i)}{G(x_i)} - \frac{\dot{G}_b(x_i)}{\bar{G}(x_i)} \right] - 2 \sum_{i=1}^n \frac{\dot{G}_b(x_i)}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \left\{ \left[\alpha G(x_i)^{\alpha-1} \ln G(x_i) \right. \right. \\ & \left. \left. + G(x_i)^{\alpha-1} - \alpha \bar{G}(x_i)^{\alpha-1} \ln \bar{G}(x_i) - \bar{G}(x_i)^{\alpha-1} \right] [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] \right. \\ & \left. - \alpha [G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}] [G(x_i)^{\alpha-1} \ln G(x_i) - \bar{G}(x_i)^{\alpha-1} \ln \bar{G}(x_i)] \right\}, \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{\alpha c}(\boldsymbol{\theta}) = & \sum_{i=1}^n \left[\frac{\dot{G}_c(x_i)}{G(x_i)} - \frac{\dot{G}_c(x_i)}{\bar{G}(x_i)} \right] - 2 \sum_{i=1}^n \frac{\dot{G}_c(x_i)}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \left\{ \left[\alpha G(x_i)^{\alpha-1} \ln G(x_i) \right. \right. \\ & \left. \left. + G(x_i)^{\alpha-1} - \alpha \bar{G}(x_i)^{\alpha-1} \ln \bar{G}(x_i) - \bar{G}(x_i)^{\alpha-1} \right] [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] \right. \\ & \left. - \alpha [G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}] [G(x_i)^{\alpha-1} \ln G(x_i) - \bar{G}(x_i)^{\alpha-1} \ln \bar{G}(x_i)] \right\}, \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{aa}(\boldsymbol{\theta}) = & \sum_{i=1}^n \frac{\ddot{g}_{aa}(x_i)g(x_i) - \dot{g}_a(x_i)^2}{g(x_i)^2} \\ & + (\alpha - 1) \sum_{i=1}^n \left\{ \frac{\ddot{G}_{aa}(x_i) - 2\ddot{G}_{aa}(x_i)G(x_i) - 2\dot{G}_a(x_i)^2}{[G(x_i)\bar{G}(x_i)]} \right. \\ & \left. - \frac{\dot{G}_a(x_i)^2 [1 - 2G(x_i)]^2}{[G(x_i)\bar{G}(x_i)]^2} \right\} - 2\alpha \sum_{i=1}^n \frac{1}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \left\{ \left[\ddot{G}_{aa}(x_i)[G(x_i)^{\alpha-1} \right. \right. \\ & \left. \left. - \bar{G}(x_i)^{\alpha-1}] + (\alpha - 1)\dot{G}_a(x_i)^2 [G(x_i)^{\alpha-2} + \bar{G}(x_i)^{\alpha-2}] \right] [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] \right. \\ & \left. - \alpha \dot{G}_a(x_i)^2 [G(x_i)^{\alpha-1} + \bar{G}(x_i)^{\alpha-1}]^2 \right\}, \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{ab}(\boldsymbol{\theta}) &= \sum_{i=1}^n \frac{\ddot{g}_{ab}(x_i) g(x_i) - \dot{g}_a(x_i) \dot{g}_b(x_i)}{g(x_i)^2} + (\alpha - 1) \sum_{i=1}^n \left\{ \frac{\ddot{G}_{ab}(x_i) - 2\dot{G}_{ab}(x_i) G(x_i)}{[G(x_i) \bar{G}(x_i)]} \right. \\ &\quad \left. - 2 \frac{\dot{G}_a(x_i) \dot{G}_b(x_i)}{[G(x_i) \bar{G}(x_i)]} - \frac{\dot{G}_a(x_i) \dot{G}_b(x_i) [1 - 2G(x_i)]^2}{[G(x_i) \bar{G}(x_i)]^2} \right\} - 2\alpha \sum_{i=1}^n \frac{1}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \\ &\quad \times \left\{ \left[[G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}] \ddot{G}_{ab}(x_i) + (\alpha - 1) \dot{G}_a(x_i) \dot{G}_b(x_i) [G(x_i)^{\alpha-2} + \bar{G}(x_i)^{\alpha-2}] \right] \right. \\ &\quad \left. \times [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] - \alpha \dot{G}_a(x_i) \dot{G}_b(x_i) [G(x_i)^{\alpha-1} + \bar{G}(x_i)^{\alpha-1}]^2 \right\} \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{ac}(\boldsymbol{\theta}) &= \sum_{i=1}^n \frac{\ddot{g}_{ac}(x_i) g(x_i) - \dot{g}_a(x_i) \dot{g}_c(x_i)}{g(x_i)^2} + (\alpha - 1) \sum_{i=1}^n \left\{ \frac{\ddot{G}_{ac}(x_i) - 2\dot{G}_{ac}(x_i) G(x_i)}{[G(x_i) \bar{G}(x_i)]} \right. \\ &\quad \left. - 2 \frac{\dot{G}_a(x_i) \dot{G}_c(x_i)}{[G(x_i) \bar{G}(x_i)]} - \frac{\dot{G}_a(x_i) \dot{G}_c(x_i) [1 - 2G(x_i)]^2}{[G(x_i) \bar{G}(x_i)]^2} \right\} - 2\alpha \sum_{i=1}^n \frac{1}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \\ &\quad \times \left\{ \left[[G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}] \ddot{G}_{ac}(x_i) + (\alpha - 1) \dot{G}_a(x_i) \dot{G}_c(x_i) [G(x_i)^{\alpha-2} + \bar{G}(x_i)^{\alpha-2}] \right] \right. \\ &\quad \left. \times [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] - \alpha \dot{G}_a(x_i) \dot{G}_c(x_i) [G(x_i)^{\alpha-1} + \bar{G}(x_i)^{\alpha-1}]^2 \right\} \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{bb}(\boldsymbol{\theta}) &= \sum_{i=1}^n \frac{\ddot{g}_{bb}(x_i) g(x_i) - \dot{g}_b(x_i)^2}{g(x_i)^2} + (\alpha - 1) \sum_{i=1}^n \left\{ \frac{\ddot{G}_{bb}(x_i) - 2\dot{G}_{bb}(x_i) G(x_i) - 2\dot{G}_b(x_i)^2}{[G(x_i) \bar{G}(x_i)]} \right. \\ &\quad \left. - \frac{\dot{G}_b(x_i)^2 [1 - 2G(x_i)]^2}{[G(x_i) \bar{G}(x_i)]^2} \right\} - 2\alpha \sum_{i=1}^n \frac{1}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \left\{ \left[\dot{G}_{bb}(x_i) [G(x_i)^{\alpha-1} \right. \right. \\ &\quad \left. \left. - \bar{G}(x_i)^{\alpha-1}] + (\alpha - 1) \dot{G}_b(x_i)^2 [G(x_i)^{\alpha-2} + \bar{G}(x_i)^{\alpha-2}] \right] [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] \right. \\ &\quad \left. - \alpha \dot{G}_b(x_i)^2 [G(x_i)^{\alpha-1} + \bar{G}(x_i)^{\alpha-1}]^2 \right\}, \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{bc}(\boldsymbol{\theta}) &= \sum_{i=1}^n \frac{\ddot{g}_{bc}(x_i) g(x_i) - \dot{g}_b(x_i) \dot{g}_c(x_i)}{g(x_i)^2} + (\alpha - 1) \sum_{i=1}^n \left\{ \frac{\ddot{G}_{bc}(x_i) - 2\dot{G}_{bc}(x_i) G(x_i)}{[G(x_i) \bar{G}(x_i)]} \right. \\ &\quad \left. - 2 \frac{\dot{G}_b(x_i) \dot{G}_c(x_i)}{[G(x_i) \bar{G}(x_i)]} - \frac{\dot{G}_b(x_i) \dot{G}_c(x_i) [1 - 2G(x_i)]^2}{[G(x_i) \bar{G}(x_i)]^2} \right\} - 2\alpha \sum_{i=1}^n \frac{1}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \\ &\quad \times \left\{ \left[[G(x_i)^{\alpha-1} - \bar{G}(x_i)^{\alpha-1}] \ddot{G}_{bc}(x_i) + (\alpha - 1) \dot{G}_b(x_i) \dot{G}_c(x_i) [G(x_i)^{\alpha-2} + \bar{G}(x_i)^{\alpha-2}] \right] \right. \\ &\quad \left. \times [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] - \alpha \dot{G}_b(x_i) \dot{G}_c(x_i) [G(x_i)^{\alpha-1} + \bar{G}(x_i)^{\alpha-1}]^2 \right\} \end{aligned}$$

$$\begin{aligned} \ddot{\ell}_{cc}(\boldsymbol{\theta}) = & \sum_{i=1}^n \frac{\ddot{g}_{cc}(x_i)g(x_i) - \dot{g}_c(x_i)^2}{g(x_i)^2} + (\alpha - 1) \sum_{i=1}^n \left\{ \frac{\ddot{G}_{cc}(x_i) - 2\ddot{G}_{cc}(x_i)G(x_i) - 2\dot{G}_c(x_i)^2}{[G(x_i)\bar{G}(x_i)]} \right. \\ & \left. - \frac{\dot{G}_c(x_i)^2 [1 - 2G(x_i)]^2}{[G(x_i)\bar{G}(x_i)]^2} \right\} - 2\alpha \sum_{i=1}^n \frac{1}{[G(x_i)^\alpha + \bar{G}(x_i)^\alpha]^2} \left\{ \left[\ddot{G}_{cc}(x_i)[G(x_i)^{\alpha-1} \right. \right. \\ & \left. \left. - \bar{G}(x_i)^{\alpha-1}] + (\alpha - 1)\dot{G}_c(x_i)^2 [G(x_i)^{\alpha-2} + \bar{G}(x_i)^{\alpha-2}] \right] [G(x_i)^\alpha + \bar{G}(x_i)^\alpha] \right. \\ & \left. - \alpha \dot{G}_c(x_i)^2 [G(x_i)^{\alpha-1} + \bar{G}(x_i)^{\alpha-1}]^2 \right\}, \end{aligned}$$

where $\ddot{g}_{jk}(x) = \frac{\partial^2 g(x)}{\partial j \partial k}$ and $\ddot{G}_{jk}(x) = \frac{\partial^2 G(x)}{\partial j \partial k}$ for any $j, k = \alpha, a, b, c$.