Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

Research Article

Odd Pareto-G Family: Properties, Regression, Simulations and Applications

Mohamed Hussein[1], **Gauss M. Cordeiro**[2], **Luís H. de Santana**[3], **Gabriela M. Rodrigues**[4], **Edwin M. M. Ortega**[4*]

[1]Department of Mathematics and Computer Science, Alexandria University, Alexandria, Egypt
[2]Department of Statistics, Federal University of Pernambuco, Recife, Brazil
[3]Department of Technology, State University of Maringá, Umuarama, Brazil
[4]Department of Exact Sciences, University of São Paulo, Piracicaba, Brazil
E-mail: edwin@usp.br

Received: 17 October 2022; **Revised:** 29 January 2023; **Accepted:** 15 February 2023

Abstract: Several new families have been introduced in the last two decades to extend well-known distributions, and at the same time provide great flexibility for modelling real data. We propose the *Odd Pareto-G family* with two extra shape parameters, and obtain some of its mathematical properties, which give important information about its structure and may be useful in future research. We construct a new regression model based on a special distribution of the new family. Maximum likelihood estimation and simulations are addressed, and these results show that the estimates are consistent for different percentages of censorship. Four applications to real data show the usefulness of the new models, reveal that they are suitable for the presented data, and can be considered interesting alternatives to model censored and uncensored lifetime data.

***Keywords*:** maximum likelihood, Odd Pareto-G family, simulation, T-X family

MSC: 47N30, 62J05, 62J20

1. Introduction

Although there were already hundreds of distributions before the 90's decade, various families to generate new distributions by adding extra parameter(s) to a parent distribution have been published recently such as the Marshall-Olkin [1], exponentiated-G [2], beta-G [3], odd log-logistic-G [4], Kumaraswamy-G [5], McDonald-G [6], exponentiated generalized-G [7], transformer T-X [8], Weibull-G [9], Libby-Novick beta [10], exponentiated half-logistic [11], logistic-X [12], Lomax-G [13], Generalized Odd Log-Logistic (GOLL) [14], modified power [15], and modified alpha power [16], generalized family of exponentiated composite distributions [17], exponentiated Teissier distribution [18], among others. The process of generating new distributions aims to create a more adequate model for explaining the data under investigation than other models.

The most common technique for generating a new family is by composing the cumulative distribution function (cdf),

DOI: https://doi.org/10.37256/cm.4120232062

say $H(x)$, of a distribution with a function, $\Psi(\cdot)$, of a parent cdf $G(x)$ to produce the cdf (or survival function) of the new distribution [8]. For example, if $\Psi_1[x] = G(x)$ and $\Psi_2[x] = -\ln G(x)$, then

$$F(x) = H(\Psi_1(x)) = H(G(x))$$

is a new cdf, and

$$\bar{F}(x) = H(\Psi_2(x)) = H(-\ln G(x))$$

is a new survival function. For more details, see [19]. The odd Lindley-G family has been constructed in [20] by composing the Lindely cdf with $\Psi(x) = G(x)/\bar{G}(x)$.

The cdf of the Pareto (PA) random variable is

$$H(x) = 1 - \left(\frac{c}{x+c}\right)^{a} \quad \text{for } x \geq 0, \tag{1}$$

where $a > 0$ is the shape parameter, and $c > 0$ is the scale parameter.

Consider a parent model with cdf $G(x)$ and probability density function (pdf) $g(x) = dG(x)/dx$ having a parameter vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_s)^{\top}$. The cdf of the *odd Pareto-G* (for short "OPA-G") family (with two extra shape parameters) follows by composing (1) with $\Psi(x) = G(x)/\bar{G}(x)$

$$F(x) = F(x;a,c) = H\left[\Psi(x)\right] = 1 - \left[\frac{c\bar{G}(x)}{G(x)+c\bar{G}(x)}\right]^{a}, \tag{2}$$

where $\bar{G}(x) = 1 - G(x)$. For $a = c = 1$, it becomes the parent G.

Note that the OPAG family is equal to the Lehmann type II alternative of the Marshall-Olkin class pioneered by [21], who did not work on its properties but addressed only the cases for the semiBurr and Burr baselines. Further, the OPAG family density is a special case of the Exponentiated Generalized Marshall-Olkin (EGMO) discussed by [22], but their linear combination depends on the exponentiated Marshall-Olkin density. The linear representation here, based on the exponentiated baseline density, is more simple and easier to develop some of its properties which are not addressed before.

The pdf corresponding to (2) has the form

$$f(x) = f(x;a,c) = \frac{ac^{a}\,g(x)}{\bar{G}(x)^{2}}\left[\frac{\bar{G}(x)}{G(x)+c\bar{G}(x)}\right]^{a+1}. \tag{3}$$

The support of X in Equations (2) and (3) is the same of the baseline G distribution. From now on, let X ~OPAG(a, b, $\boldsymbol{\eta}$) be a random variable with pdf (3).

By inverting (2), the quantile function (qf) of X becomes

$$Q(u) = G^{-1}\left[\frac{c - c(1-u)^{1/a}}{c + (1-c)(1-u)^{1/a}}\right], \tag{4}$$

where $G^{-1}(\cdot) = Q_G(\cdot)$. So, we can easily generate OPAG variates from (4).

The sections follow as: Section 2 addressees two special models. Section 3 provides a linear representation in terms of exponentiated densities. The moments and qf are obtained in Section 4. Section 5 discusses inference based on likelihood theory. Section 6 examines the adequacy of the parameter estimators. Section 7 introduces a new regression model for censored data. Four applications to real data sets prove the importance of our findings in Section 8. Section 9 ends with some conclusions.

2. Special models

For illustrations, simulations, estimation and applications, we consider two special models of (2) by taking the Weibull and gamma as baseline distributions.

2.1 *Odd Pareto Weibull (OPAW) model*

The Weibull cdf with shape $k > 0$ and scale $b > 0$ is $G(x) = 1 - e^{-\left(\frac{x}{b}\right)^k}$ (for $x \geq 0$). The pdf, cdf and hazard rate function (hrf) of the OPAW(a, c, b, k) random variable follow from (2) and (3) as

$$f(x) = \frac{a b^{-k} k c^a x^{k-1} e^{-a\left(\frac{x}{b}\right)^k}}{\left[1 - \bar{c} e^{-\left(\frac{x}{b}\right)^k}\right]^{a+1}},$$

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

$$h(x) = \frac{a b^{-k} k x^{k-1}}{1 - \bar{c} e^{-\left(\frac{x}{b}\right)^k}}, \ x \geq 0.$$

where $\bar{c} = 1 - c$.

The pdf and hrf of the OPAW model are displayed in Figures 1 and 2, respectively.

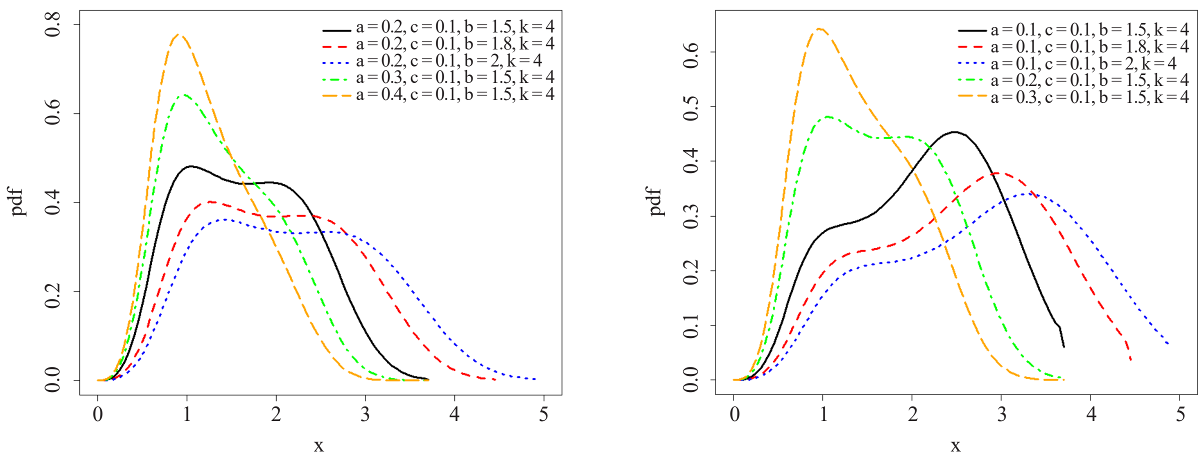


Figure 1. Plots of the OPAW density

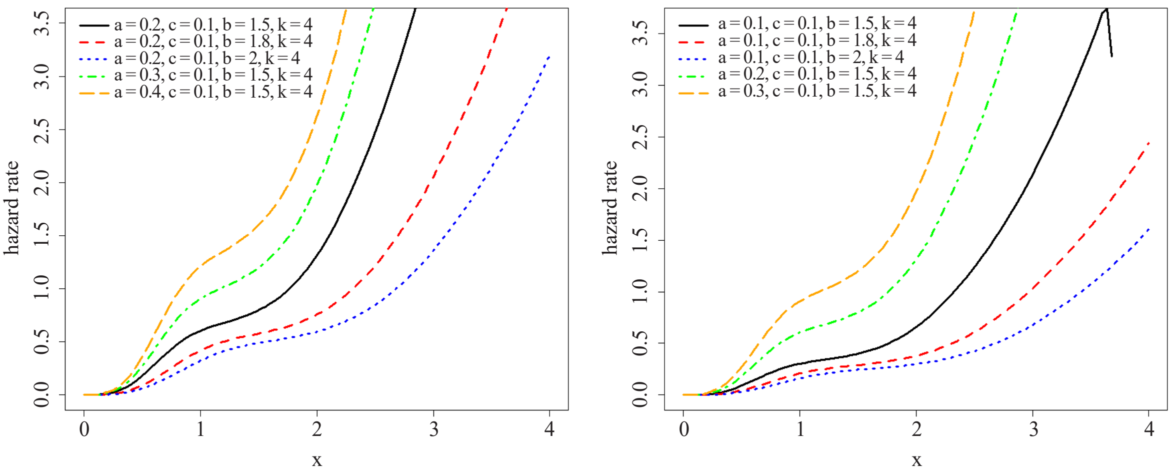
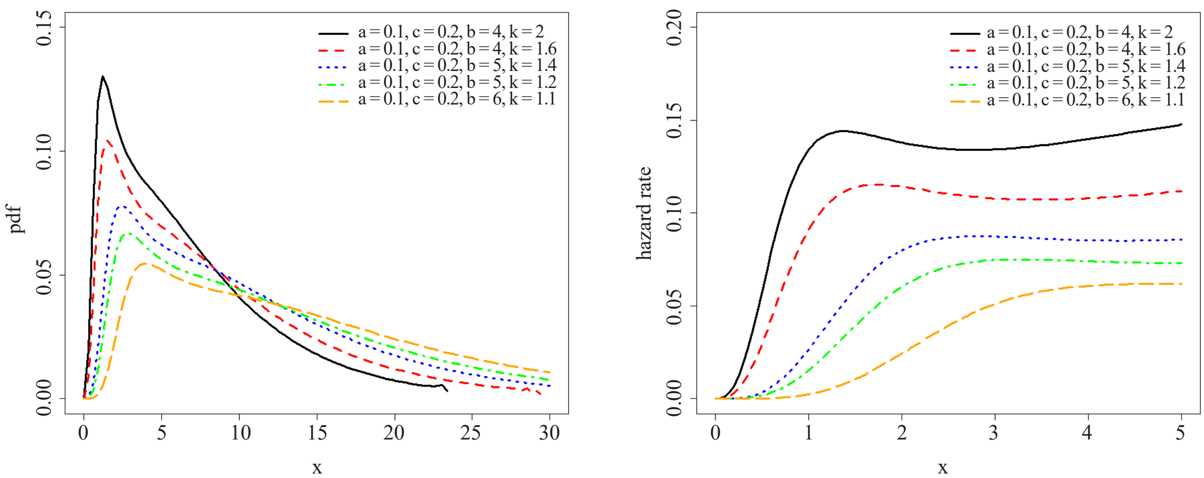


Figure 2. Plots of the OPAW hrf

Figures 1 and 2 illustrate some possible shapes of the pdf and hrf of the OPAW distribution, respectively. Its density can be unimodal, bimodal and left-right asymmetric. The hrf can be increasing and unimodal, which offers great capacity for real datasets.



(a) pdf (b) hrf

Figure 3. Plots for the OPAGA distribution

2.2 *Odd Pareto Gamma (OPAGA) model*

The gamma cdf with shape $b > 0$ and rate $k > 0$ is $G(x) = \gamma(b, kx)/\Gamma(b)$ $(x \geq 0)$, where $\Gamma(b) = \int_0^\infty t^{b-1} \mathrm{e}^{-t} \mathrm{d}t$ and $\gamma(b, z) = \int_0^z t^{b-1} \mathrm{e}^{-t} \mathrm{d}t$ are the gamma and lower incomplete gamma functions, respectively.

The pdf, cdf and hrf of the OPAGA(a, c, b, k) random variable follow from (2) and (3) as

$$f(x) = \frac{ac^a k^b \Gamma(b) x^{b-1} \mathrm{e}^{-kx} \left[\Gamma(b) - \gamma(b, kx)\right]^{a-1}}{\left[c\Gamma(b) + \bar{c}\, \gamma(b, kx)\right]^{a+1}},$$

$$F\left(x\right)=1-c^{a}\left[\frac{\Gamma(b)-\gamma(b,kx)}{c\Gamma(b)+\bar{c}\gamma(b,kx)}\right]^{a},$$

$$h\left(x\right)=\frac{a k^{b}\,\Gamma(b)x^{b-1}\,\mathrm{e}^{-kx}}{\left[c\Gamma(b)+\bar{c}\,\gamma(b,kx)\right]\left[\Gamma(b)-\gamma(b,kx)\right]},\quad x\geq 0.$$

Figure 3 provides some shapes of the pdf and hrf of the OPAGA model, which can be unimodal in both cases. Further, the hrf has a large variety of shapes including increasing-decreasing-increasing shape.

We address the new OPAW and OPAGA models since they can be used in different areas, mainly in survival and reliability analysis.

3. Linear representation

A linear representation for the cdf (2) follows from two different manners. In the Marshall-Olkin way, we have

$$F(x)=1-\left[\frac{c\bar{G}(x)}{1-\bar{c}\,\bar{G}(x)}\right]^{a} \tag{5}$$

or

$$F(x)=1-\left[\frac{\bar{G}(x)}{1+d\,G(x)}\right]^{a}, \tag{6}$$

where $d=\bar{c}c$.

The power series holds (for any real a)

$$(1+z)^{a}=\sum_{k=0}^{\infty}\binom{a}{k}z^{k},\quad |z|<1, \tag{7}$$

where $\binom{a}{0}=1$, and (for $k\geq 1$) $\binom{a}{k}=a(a-1)\ldots(a-k+1)/k!$.

First, consider Equation (5) for $c<2$ or, equivalently, $|-\bar{c}\,|<1$. Then, $|-\bar{c}\bar{G}(x)|<1$, and it follows from Equation (7)

$$\left[1-\bar{c}\,\bar{G}(x)\right]^{a}=\sum_{j=0}^{\infty}A_{j}(a,c)G(x)^{j}, \tag{8}$$

where $A_{j}(a,c)=\sum_{k=j}^{\infty}(-1)^{j+k}\bar{c}^{k}\binom{a}{k}\binom{k}{j}$.

On the other hand,

$$c^{a}\,\bar{G}(x)^{a}=\sum_{j=0}^{\infty}B_{j}(a,c)G(x)^{j}, \tag{9}$$

where $B_j(a,c) = (-1)^j c^a \binom{a}{j}$.

Inserting (8) and (9) in Equation (5), and using the ratio of two power series

$$F(x) = 1 - \sum_{j=0}^{\infty} U_j(a,c) G(x)^j, \tag{10}$$

where $U_j = U_j(a, c)$ (for $j \geq 1$) is determined recursively by

$$U_j = \frac{1}{A_0}\left(B_j - \sum_{r=1}^{j} A_r U_{j-r}\right),\ U_0 = \frac{B_0}{A_0}.$$

Second, consider Equation (6) for $c > 12$. Then, $0 < c^{-1} < 2$, $-1 < c^{-1} - 1 < 1$, and $d \in (0, 1)$. Hence, $|dG(x)| < 1$. In particular, for $c \geq 2$, the last inequality holds, and from (7),

$$[1 + d\,G(x)]^a = \sum_{j=0}^{\infty} C_j(a,c) G(x)^j, \tag{11}$$

where $C_j(a,c) = (\bar{c}c)^j \binom{a}{j}$. Further,

$$\bar{G}(x)^a = \sum_{j=0}^{\infty} D_j(a) G(x)^j, \tag{12}$$

where $D_j(a) = (-1)^j \binom{a}{j}$. Inserting (12) and (11) in Equation (6),

$$F(x) = 1 - \sum_{j=0}^{\infty} V_j(a,c) G(x)^j, \tag{13}$$

where $V_j = V_j(a, c)$ ($j \geq 1$) follows recursively

$$V_k = \frac{1}{C_0}\left(D_j - \sum_{r=1}^{j} C_r V_{j-r}\right), V_0 = \frac{D_0}{C_0}.$$

The pdf of the *exponentiated-G* ("EXP-G") random variable W_α ~EXP-G(α) with power parameter $\alpha > 0$ is defined by $\pi_\alpha(x) = \alpha g(x) G(x)^{\alpha-1}$. By differentiating (10) and (13), the pdf of X reduces to

$$f(x) = \sum_{j=0}^{\infty} s_j \pi_{j+1}(x), \tag{14}$$

where

$$s_j = s_j(a,c) = \begin{cases} -U_j(a,c), & c < 2, \\ -V_j(a,c), & c \geq 2. \end{cases}$$

Hence, some properties of X can be found from Equation (14) and those EXP-G properties in several references listed in Tahir and Nadarajah's Table 1 [23].

4. Properties

4.1 *Moments*

Let W_{j+1} denote the EXP-G$(j+1)$ random variable. The nth moment of X comes from (14) as

$$\mu_n' = E(X^n) = \sum_{j=0}^{\infty} s_j\, E(W_{j+1}^n) = \sum_{j=0}^{\infty} (j+1) s_j\, \tau_{n,j},$$

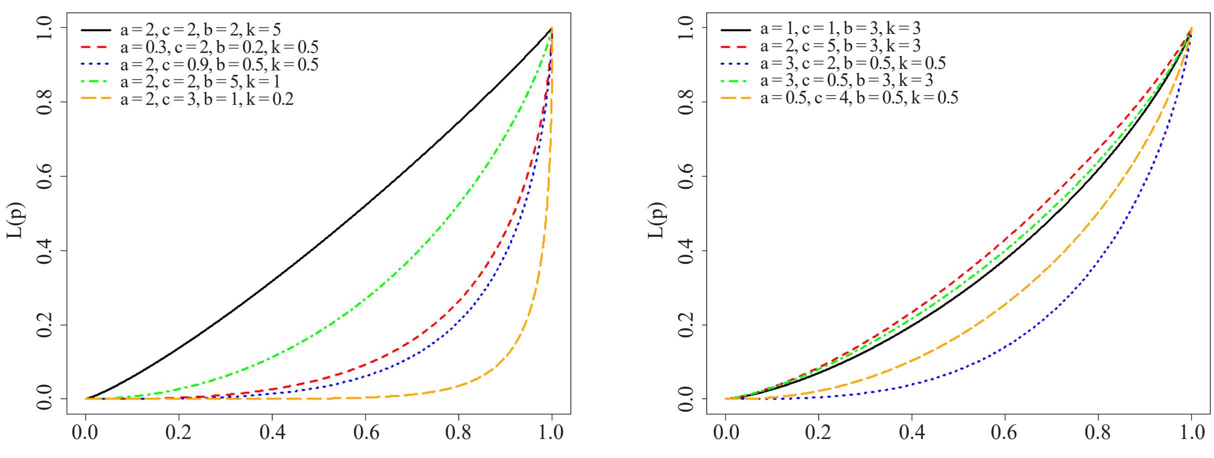
where $\tau_{n,j} == \int_0^1 Q_G(u)^n\, u^j\, du$.

In a similar manner, the nth incomplete moment of X is

$$m_n(y) = E\left(X^n \mid X < y\right) = \sum_{j=0}^{\infty} (j+1) s_j \int_0^{G(y)} Q_G(u)^n\, u^j\, du,$$

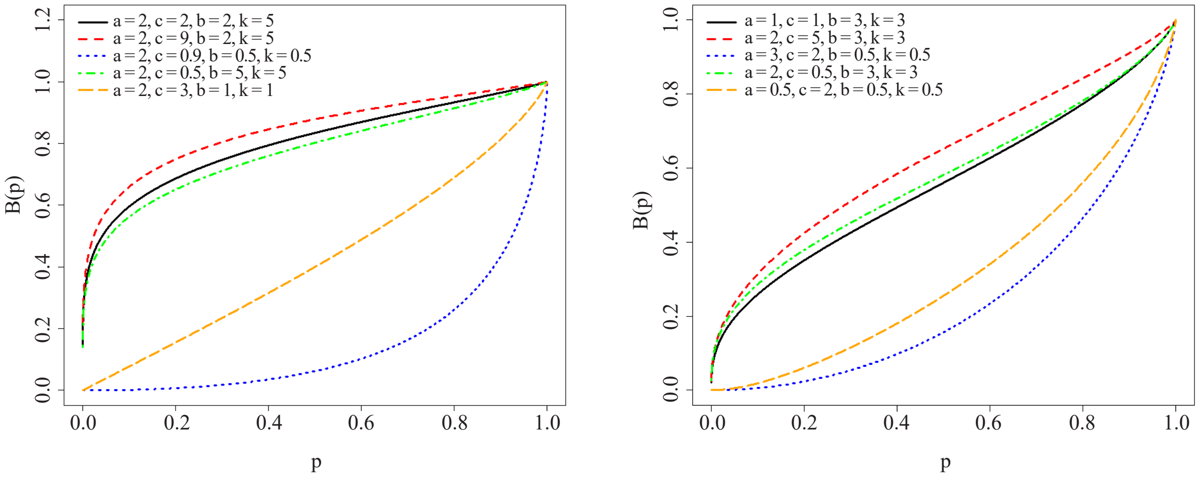
where the integral can be calculated numerically.

Lorenz and Bonferroni curves, widely used as inequality measures ([24] and [25]), are given by $L(p) = m_1(q)/\mu_1'$ and $B(p) = m_1(q)/(p\mu_1')$, respectively, where $q = Q(p)$ follows from Equation (4). The Lorenz curve is a graphical representation of the distribution of wealth in an economy. It is used to measure the degree of inequality in a society by plotting the cumulative percentage of total income earned against the cumulative percentage of population. This curve can be interpreted as a measure of how evenly wealth is distributed among members of a society. A perfectly equal distribution would result in a straight line, while an unequal distribution would result in a curved line. The further away from the straight line, the greater the degree of inequality. The area between the Lorenz curve and the straight line represents the amount of wealth that is unequally distributed among members of society (unbalance degree). Figures 4a and 4b display the Lorenz curves for the OPAW and OPAGA distributions and indicate that the OPAW parameters have more influence on the inequality than those of the OPAGA model. The Bonferroni curve is a slight alteration of the Lorenz curve. Figures 5a and 5b display plots of the Bonferroni curve for the OPAW and OPAGA distributions, respectively.



(a) OPAW (b) OPAGA

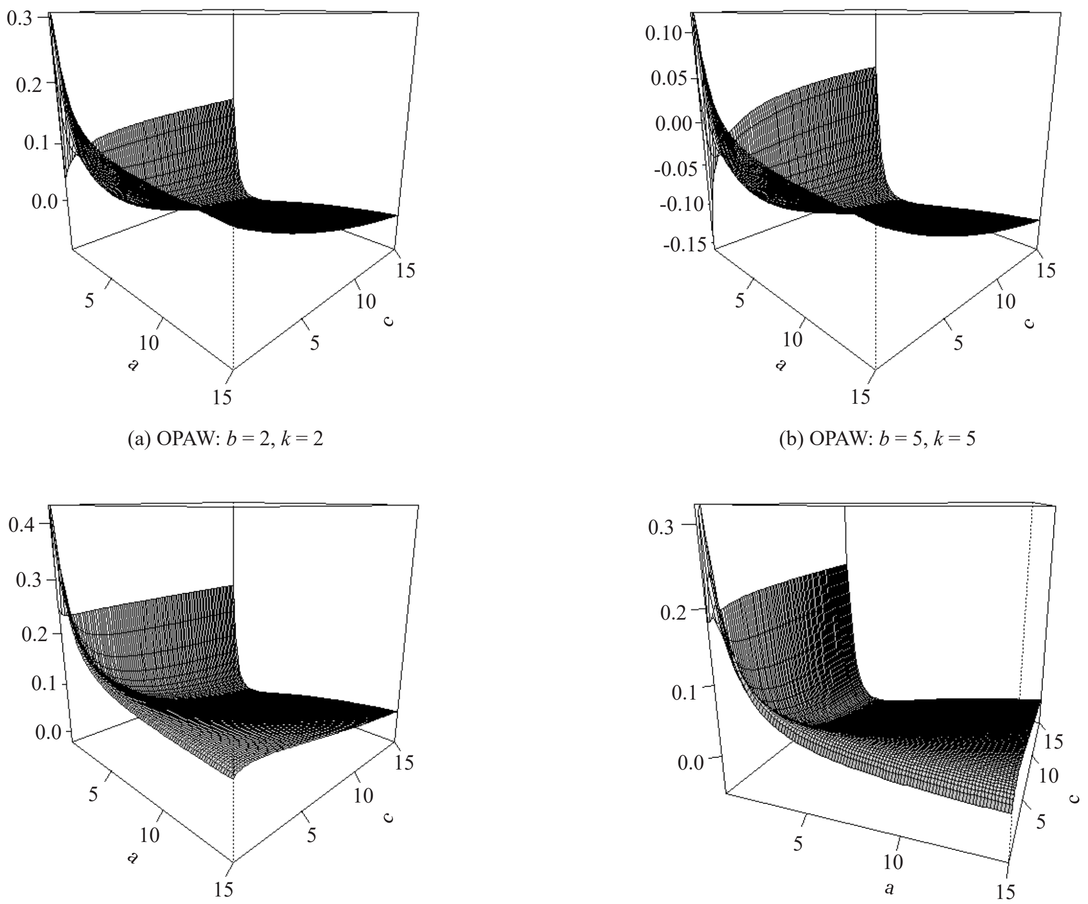
Figure 4. Lorenz curves for the OPAW and OPAGA distributions



(a) OPAW

(b) OPAGA

Figure 5. Bonferroni curves for the OPAW and OPAGA distributions



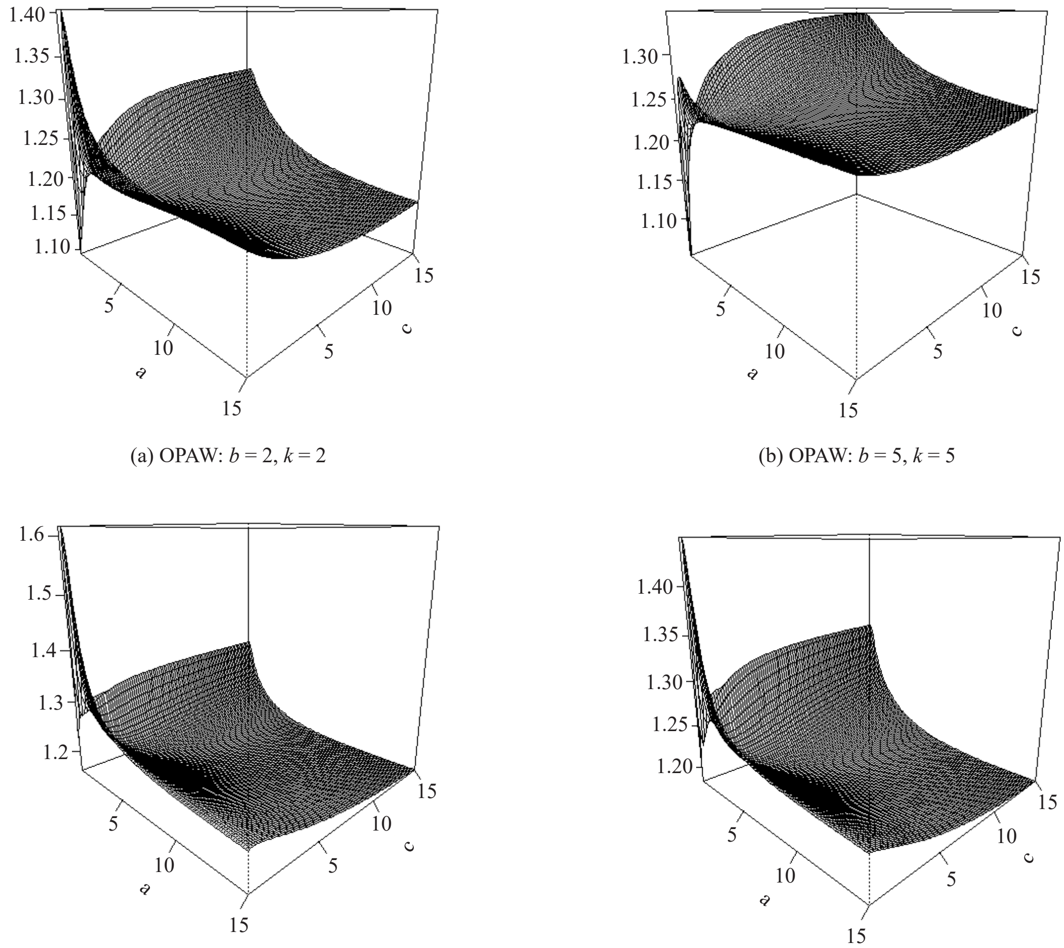
(a) OPAW: $b = 2$, $k = 2$

(b) OPAW: $b = 5$, $k = 5$

(c) OPAGA: $b = 2$, $k = 2$

(d) OPAGA: $b = 5$, $k = 5$

Figure 6. Bowley skewness for the OPAW and OPAGA models



(a) OPAW: $b = 2$, $k = 2$ (b) OPAW: $b = 5$, $k = 5$

(c) OPAGA: $b = 2$, $k = 2$ (d) OPAGA: $b = 5$, $k = 5$

Figure 7. Moors' kurtosis for the OPAW and OPAGA models

4.2 *Quantile measures*

The Bowley skewness [26] and Moors' kurtosis [27] of X are functions of the octiles and they easily follow from (4). These quantile measures for the OPAW and OPAGA models are displayed in Figures 6 and 7.

The plots in Figure 6 reveal that the OPAW and OPAGA models could be asymmetric (positive or negative) or symmetric, whereas those in Figure 7 show moderate variations of these distributions.

5. Estimation

Let $x_1, \ldots, x_n$ be independent observations from the OPAG family, and $\boldsymbol{\theta} = (a, c_1, \boldsymbol{\eta}^\top)^\top$, where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_s)^\top$ and $c_1 = 1/c$. The total log-likelihood function for θ reduces to

$$\ell(\boldsymbol{\theta}|\ x) = n(\log a + \log c_1) - (a+1)\left\{\sum_{i=1}^{n} \log\left[\bar{G}(x_i) + c_1\, G(x_i)\right] - \sum_{i=1}^{n} \log \bar{G}(x_i)\right\}$$

$$+\sum_{i=1}^{n}\log g(x_i)-2\sum_{i=1}^{n}\log\bar{G}(x_i). \tag{15}$$

The score components can be expressed as

$$\frac{\partial \ell}{\partial a}=\frac{n}{a}-\sum_{i=1}^{n}\ln\left[\bar{G}(x_i)+c_1 G(x_i)\right]+\sum_{i=1}^{n}\ln\bar{G}(x_i),$$

$$\frac{\partial \ell}{\partial c_1}=\frac{n}{c_1}-(a+1)\sum_{i=1}^{n}\frac{G(x_i)}{\bar{G}(x_i)+c_1 G(x_i)},$$

$$\frac{\partial \ell}{\partial \boldsymbol{\eta}_k}=(a+1)\sum_{i=1}^{n}\left[\frac{1-c_1}{\bar{G}(x_i)+c_1 G(x_i)}-\frac{1}{\bar{G}(x_i)}\right]\frac{\partial G(x_i)}{\partial \boldsymbol{\eta}_k}$$

$$+\sum_{i=1}^{n}\frac{1}{g(x_i)}\frac{\partial g(x_i)}{\partial \boldsymbol{\eta}_k}+2\sum_{i=1}^{n}\frac{1}{\bar{G}(x_i)}\frac{\partial G(x_i)}{\partial \boldsymbol{\eta}_k},\ k=1,\dots,s.$$

Setting the score components to zero, the Maximum Likelihood Estimates (MLEs) of a, c_1, η_k (k = 1, ... s) are the solution of a nonlinear system of equations, which has no explicit solution for these parameters. Otherwise, the MLEs of these parameters can be determined by maximizing (15) using the *AdequacyModel* library in R software [28]. We can also use the functions optim in R, MaxBFGS in Ox or PROC NLMIXED in SAS to find these estimates.

The inference on the parameters is based on standard likelihood theory. The observed information matrix is given in Appendix A.

6. Simulations

A simulation study is done for the OPAW distribution to assess the estimators performance. The Average Estimates (AEs), Absolute Biases (ABs), and Mean Square Errors (MSEs) of the MLEs are calculated from 1,000 samples for some scenarios reported in Tables 1, 2 and 3, respectively. These findings indicate that the AEs tend to the true values, and the ABs and MSEs decay when n increases, which ensures that the MLEs of the OPAW parameters are consistent.

7. OPAW regression model

Regression analysis of lifetime data involves specification of the lifetime distribution of X given a vector of covariates $\mathbf{v} = (\mathbf{v}_1, ..., \mathbf{v}_p)^\top$. We introduce the OPAW regression model for censored data under two systematic components. The most important form of this model defines the parameters b and k depending on $\mathbf{v}$ by the logarithm link functions (i = 1, ..., n), $b_i = \exp(\mathbf{v}_i^T\boldsymbol{\beta}_1)$ and $k_i = \exp(\mathbf{v}_i^T\boldsymbol{\beta}_2)$, respectively, where $\boldsymbol{\beta}_1 = (\beta_{11}, ..., \beta_{1p})^\top$ and $\boldsymbol{\beta}_2 = (\beta_{21}, ..., \beta_{2p})^\top$ are the vectors of regression coefficients and $\mathbf{v}_i^\top = (v_{i1}, ..., v_{ip})$. The OPAW regression model, where both parameters b and k depend on $\mathbf{v}$, is very useful in many practical situations.

Table 1. Averages from simulations of the OPAW distribution

Parameters					AEs			
a	c_1	b	k	n	$\hat{a}$	$\hat{c}_1$	$\hat{b}$	$\hat{k}$
0.5	0.5	0.5	0.5	20	0.5946	0.6012	0.6915	0.5916
				50	0.4435	0.5828	0.6234	0.5567
				100	0.4502	0.5686	0.5639	0.5410
				200	0.5010	0.5023	0.5049	0.5387
0.5	0.5	2	2	20	0.3311	0.8719	1.6121	2.3676
				50	0.6503	0.7241	1.7654	2.2237
				100	0.5286	0.4626	1.8229	2.1369
				200	0.4906	0.4912	1.9558	2.0979
2	2	2	2	20	2.1450	1.7785	1.6504	2.4716
				50	2.0441	1.8830	1.8391	2.1087
				100	2.0302	2.0211	1.9655	2.0875
				200	2.0131	2.0065	1.9954	2.0118
2	2	0.5	0.5	20	1.7470	1.5797	0.2517	0.6164
				50	1.8333	2.3947	0.2751	0.5924
				100	2.0136	2.0858	0.3315	0.5901
				200	2.0001	1.9524	0.4992	0.5334
5	4	5	2	20	4.3268	3.2155	4.7421	3.0411
				50	4.7567	3.5559	5.1241	2.8656
				100	4.8925	3.8862	4.8896	2.4445
				200	5.0071	4.0107	4.9798	2.1269
0.5	3	3	3	20	0.2531	2.5056	1.8872	2.0006
				50	0.4507	3.3043	3.9200	3.1656
				100	0.4877	2.9041	3.3585	3.0072
				200	0.5083	2.9504	2.9511	3.0035
3	0.5	3	3	20	2.1823	0.6893	2.4840	3.4025
				50	2.7182	0.6575	3.3712	3.2159
				100	2.7470	0.5255	3.0645	2.9877
				200	3.0058	0.5001	3.2140	2.9921

Table 2. ABs from simulations of the OPAW distribution

Parameters					AEs			
a	c_1	b	k	n	$\hat{a}$	$\hat{c}_1$	$\hat{b}$	$\hat{k}$
0.5	0.5	0.5	0.5	20	0.0946	0.1012	0.1915	0.0916
				50	0.0565	0.0828	0.1234	0.0567
				100	0.0498	0.0686	0.0639	0.0410
				200	0.0010	0.0023	0.0049	0.0387
0.5	0.5	2	2	20	0.1689	0.3719	0.3879	0.3676
				50	0.1503	0.2241	0.2346	0.2237
				100	0.0286	0.0374	0.1771	0.1369
				200	0.0094	0.0088	0.0442	0.0979
2	2	2	2	20	0.1450	0.2215	0.3496	0.4716
				50	0.0441	0.1170	0.1609	0.1087
				100	0.0302	0.0211	0.0345	0.0875
				200	0.0131	0.0065	0.0046	0.0118
2	2	0.5	0.5	20	0.2530	0.4203	0.2483	0.1164
				50	0.1667	0.3947	0.2249	0.0924
				100	0.0136	0.0858	0.1685	0.0901
				200	0.0001	0.0476	0.0008	0.0334
5	4	5	2	20	0.6732	0.7845	0.2579	1.0411
				50	0.2433	0.4442	0.1241	0.8656
				100	0.1075	0.1138	0.1104	0.4445
				200	0.0071	0.0107	0.0202	0.1269
0.5	3	3	3	20	0.2469	0.4944	1.1128	0.9994
				50	0.0493	0.3043	0.9195	0.1656
				100	0.0123	0.0959	0.3585	0.0072
				200	0.0083	0.0496	0.0489	0.0035
3	0.5	3	3	20	0.8177	0.1893	0.5160	0.4025
				50	0.2818	0.1575	0.3712	0.2159
				100	0.2530	0.02547	0.0645	0.0123
				200	0.0058	0.0001	0.2140	0.0079

Table 3. MSEs from simulations of the OPAW distribution

Parameters					MSEs			
a	c_1	b	k	n	$\hat{a}$	$\hat{c}_1$	$\hat{b}$	$\hat{k}$
0.5	0.5	0.5	0.5	20	0.2027	0.3851	0.3520	0.7408
				50	0.1766	0.1707	0.2497	0.3183
				100	0.0283	0.0566	0.1304	0.0960
				200	0.0130	0.0331	0.0794	0.0540
0.5	0.5	2	2	20	0.4480	0.7213	0.6200	0.0612
				50	0.1500	0.0955	0.2584	0.0349
				100	0.0759	0.0641	0.0453	0.0187
				200	0.0139	0.0490	0.0359	0.0115
2	2	2	2	20	2.0449	1.0364	0.1509	0.5079
				50	0.2411	0.2614	0.1413	0.1844
				100	0.0214	0.0619	0.1055	0.0924
				200	0.0115	0.0103	0.0315	0.0685
2	2	0.5	0.5	20	0.7333	1.2086	0.3205	1.8455
				50	0.2029	0.1947	0.2551	0.6124
				100	0.1796	0.1844	0.0347	0.0354
				200	0.0668	0.1654	0.0043	0.0297
5	4	5	2	20	1.2136	0.9086	1.6549	0.8187
				50	1.1918	0.2946	1.2551	0.6124
				100	0.9647	0.0863	0.456	0.0354
				200	0.1654	0.0262	0.239	0.0264
0.5	3	3	3	20	2.0653	1.2446	1.2457	1.0989
				50	1.1446	1.1246	1.1758	0.9997
				100	0.7605	0.5460	1.1344	0.9200
				200	0.064	0.0246	1.1015	0.8000
3	0.5	3	3	20	1.8322	1.5421	1.0257	1.1587
				50	0.9370	0.3251	0.8015	1.0185
				100	0.2810	0.1254	0.1549	0.9254
				200	0.1630	0.0542	0.12354	0.1175

The survival function of X_i given $\mathbf{v}_i$ follows from (2) as

$$S(x_i) = \left\{ \frac{c\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]}{1-\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]+c\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]} \right\}^a . \tag{16}$$

Let F and C be the sets of individuals for which x_i is the lifetime or censoring c_i, respectively. The total log-likelihood function for $\boldsymbol{\theta} = (a, c, \boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ from model (16) has the form

$$l(\boldsymbol{\theta}) = r\log(ac^a) - \sum_{i\in F} k_i \log(b_i) + \sum_{i\in F}\log(k_i) + \sum_{i\in F}(k_i-1)\log(x_i) - a\sum_{i\in F}\left(\frac{x_i}{b_i}\right)^{k_i}$$

$$-(a+1)\sum_{i\in F}\log\left\{1-\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]+c\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]\right\}$$

$$+a\sum_{i\in C}\log\left\{ \frac{c\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]}{1-\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]+c\exp\left[-\left(\frac{x_i}{b_i}\right)^{k_i}\right]} \right\}, \tag{17}$$

where r is the number of failures.

The MLE $\hat{\boldsymbol{\theta}}$ of the vector of unknown parameters can be found by maximizing Equation (17) via R software or NLMixed procedure in SAS. Equation (17) is twice differentiable with respect to the variable x_i since it is a composite function built by two times differentiable functions. Then, the standard regularity conditions for maximum likelihood estimation are satisfied, and the existence of the MLEs with desirable properties follow.

For the OPAW regression model, we consider the quantile residuals (qrs), namely

$$\widehat{rq}_i = \Phi^{-1}\left(1-\left\{ \frac{\hat{c}\exp\left[-\left(\frac{x_i}{\hat{b}_i}\right)^{\hat{k}_i}\right]}{1-\exp\left[-\left(\frac{x_i}{\hat{b}_i}\right)^{\hat{k}_i}\right]+\hat{c}\exp\left[-\left(\frac{x_i}{\hat{b}_i}\right)^{\hat{k}_i}\right]} \right\}^{\hat{a}}\right).$$

[36] suggested the construction of envelopes to enable better interpretation of the probability normal plot of the residuals. These envelopes are simulated confidence bands that contain the residuals such that if the model is well-fitted, the majority of points will be within these bands and randomly distributed.

8. Applications

We consider two sub-models of the new family to compare with the baseline distributions and six other extended models (with two extra parameters) by means of four real datasets. The models are: the Beta-G (BG) [3], Exponentiated Generalized-G (EGG) [7], Kumaraswamy-G (KG) [5], Log Gamma-G I and II (LGGI and LGGII) [30], and Weibull-G (WG) [8]. The maximum log-likelihood (l_{max}), four classical statistics denoted by their initials (AIC, CAIC, BIC, HQIC), and Anderson-Darling (A^*) and Cramér-von Mises (W^*) are adopted as measures of the adequacy of the fitted models.

Dataset I: 1,150 heights measurements of the surface roughness of rollers (available for download [31]).

Dataset II: 74 gauge lengths of 20 mm [32].

The descriptive statistics for both datasets are reported in Table 4

Table 4. Descriptive analysis

Data	Mean	Median	Mode	Variance	Skewness	Kurtosis	Min	Max
I	3.535	3.614	3.750	0.422	-0.987	1.863	0.237	5.150
II	2.477	2.513	2.750	0.238	-0.154	-0.049	1.312	3.585

Table 5. Findings from the fitted distributions

Dataset	Model	Parameter	MLE	SE
I	OPAW	a	1.139	0.229
		c_1	0.031	0.019
		b	2.419	0.242
		k	2.976	0.427
II	OPAGA	a	1.314	1.089
		c_1	0.149	0.173
		b	17.14	6.095
		k	8.361	2.835

We consider the Weibull and gamma baselines for fitting eight distributions to datasets I and II. The MLEs, their Standard Errors (SEs) for the OPAW and OPAGA models are reported in Table 5. The adequacy statistics of the eight fitted distributions are given in Tables 6 and 7. Figures 8 and 9 illustrate the closeness of the estimated pdfs and cdfs for both models to their empirical counterparts. Further, the Probabilityprobability (PP) and Quantile-Quantile (QQ) plots are close to the first bisector line.

Table 6. Adequacy measures for dataset I

Model	AIC	CAIC	BIC	HQIC	W^{*}	A^{*}	$-\ell_{max}$
OPAW	2120.614	2120.649	2140.804	2128.236	0.074	0.463	1056.307
Weibull	2179.853	2179.863	2189.948	2183.663	0.651	4.004	1087.926
BW	2164.985	2165.020	2185.175	2172.606	0.645	3.618	1078.492
EGW	2166.374	2166.409	2186.564	2173.995	0.663	3.719	1079.187
KW	2152.446	2152.481	2172.636	2160.067	0.517	2.860	1072.223
LGWI	2164.924	2164.959	2185.114	2172.545	364.9	2284	1078.462
LGWII	2143.234	2143.269	2163.424	2150.855	0.425	2.333	1067.617
WW	2183.853	2183.888	2204.043	2191.474	0.833	4.817	1087.926

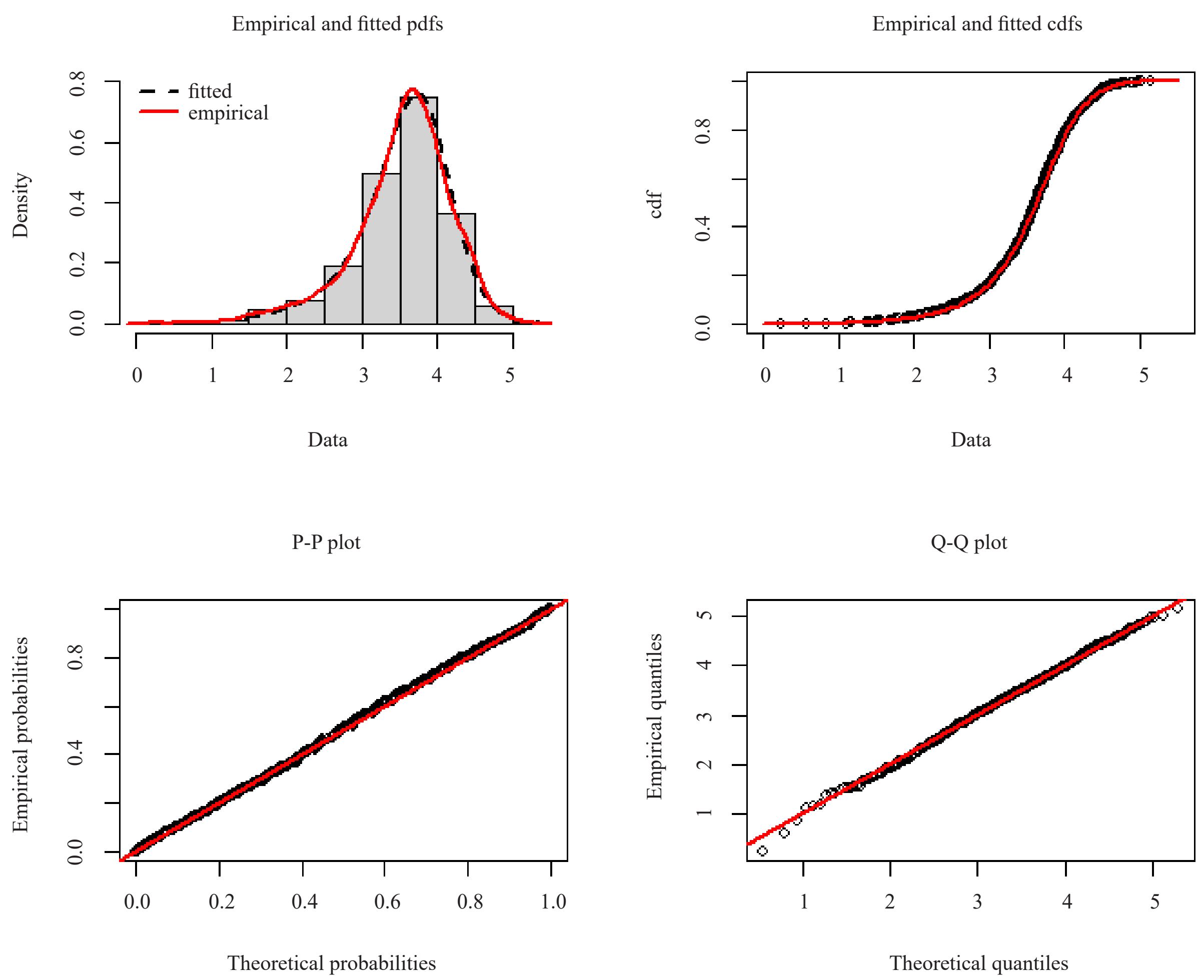


Figure 8. Results from the OPAW distribution fitted to dataset I

Table 7. Adequacy measures for dataset II

Model	AIC	CAIC	BIC	HQIC	W^*	A^*	$-\ell_{max}$
OPAW	109.896	110.476	119.112	113.572	0.023	0.182	50.948
Weibull	110.330	110.499	114.938	112.168	0.086	0.564	53.165
BW	110.227	110.807	119.443	113.903	0.026	0.209	51.113
EGW	110.260	110.839	119.476	113.936	0.027	0.213	51.130
KW	110.314	110.894	119.530	113.990	0.029	0.219	51.157
LGWI	110.227	110.807	119.443	113.904	24.50	148.0	51.114
LGWII	112.406	112.986	121.623	116.083	0.061	0.414	52.203
WW	110.273	110.853	119.489	113.950	0.028	0.223	51.137

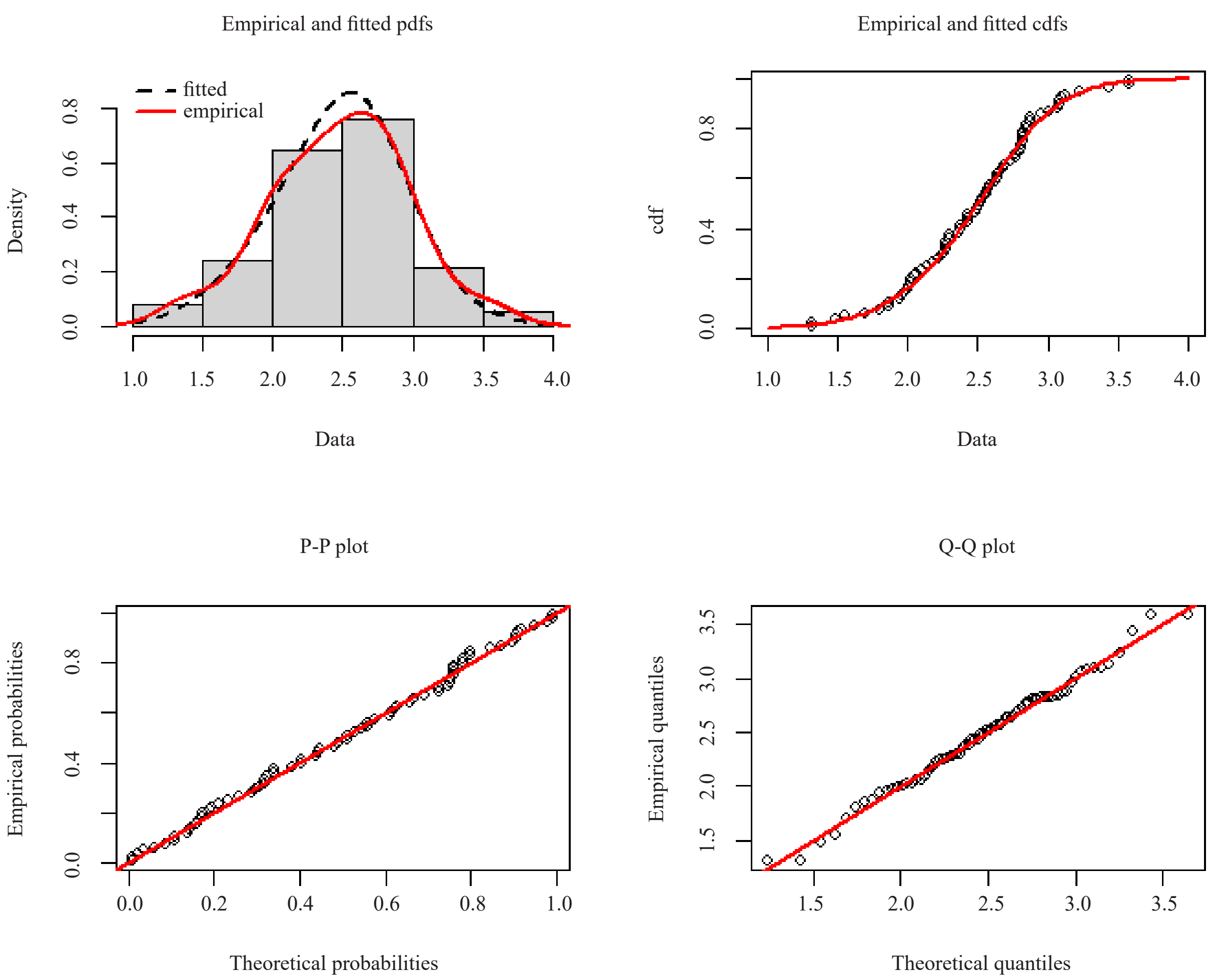


Figure 9. Results from the OPAGA distribution fitted to dataset II

8.1 *Datasets III and IV: Bimodal data and regression model*

The third application refers to a bimodal dataset to show the flexibility of the OPAW distribution, and the fourth application considers a OPAW regression model for censored data. We compare the results of their fits with those of non-nested KW and BW models. The computations are performed with the gamlss script in R software [28].

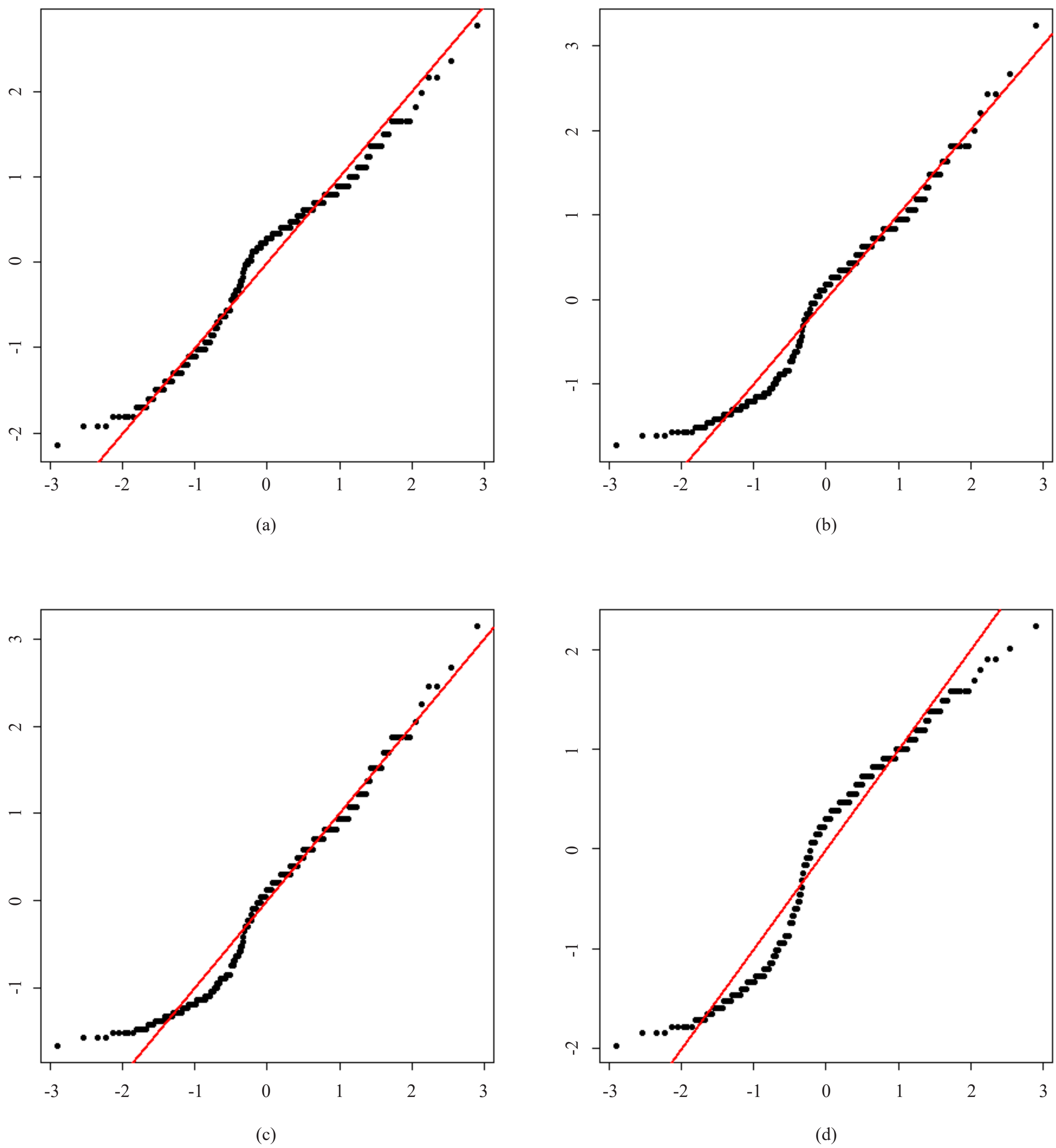


Figure 10. QQ plots from fitted models to dataset III: (a) OPAW (b) KW (c) BW (d) Weibull

8.1.1 *Dataset III*

The data consist of $n = 272$ waiting times (in minutes) between eruptions in Old Faithful geyser in Wyoming's Yellowstone National Park, USA (library datasets in the R software).

Table 8 reports the values of some statistics for the fitted models, which indicate that the OPAW distribution is the best model. Figure 10 displays the QQ plots of the qrs for the fitted models, thus supporting previous conclusion. Table 9 gives the MLEs and their SEs. The histogram and the estimated densities in Figure 11a, and the empirical and estimated cdfs in Figure 11b reveal that the OPAW distribution explains the current data.

Table 8. Measures for some models fitted to dataset III

Model	AIC	CAIC	BIC	HQIC	W^*	A^*	$-\ell_{max}$
OPAW	2127.87	2128.02	2142.30	2133.66	0.4828	3.4287	1059.94
KW	2157.62	2157.77	2172.04	2163.408	0.5796	3.5759	1074.81
BW	2155.96	2180.33	2170.38	2185.971	0.5153	3.4972	1073.98
Weibull	2174.02	2174.06	2181.23	2176.911	1.1559	6.4249	1085.01

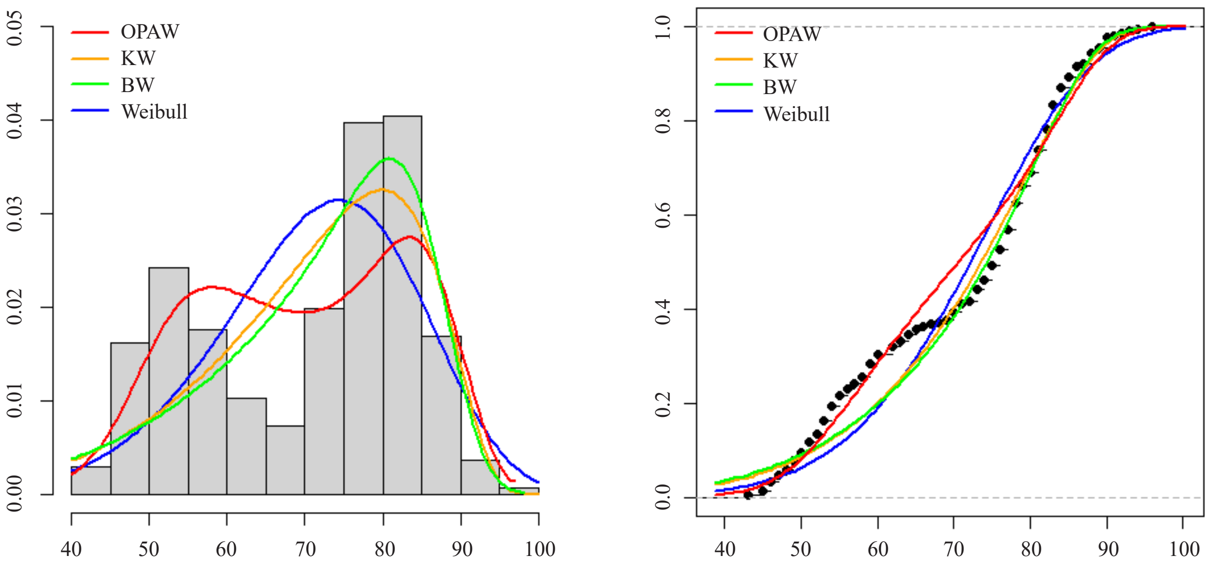


Figure 11. Dataset III. (a) Estimated densities. (b) Empirical and estimated cdfs

Table 9. Findings for the OPAW model fitted to dataset III

	b	k	a	c
MLEs	72.6028	12.4117	0.1614	0.0140
SEs	(0.0244)	(0.0931)	(0.3298)	(0.7348)

8.1.2 *Dataset IV*

The data refer to head cancer ($n = 96$), where 51 patients received radiotherapy (Arm A) and 45 patients received radiotherapy and chemotherapy (Arm B). In this study, we have approximately 24% of censored data. For more details, see [33-35].

Just one covariate v_1 is used: two-Arm (Arm $A = 0$ and Arm $B = 1$), and the systematic components for b and k are

$$b = \exp\left(\beta_{10} + \beta_{11} v_{i1}\right) \text{ and } k = \exp\left(\beta_{20} + \beta_{21} v_{i1}\right),\ i = 1\ldots 96,$$

where $\boldsymbol{\beta}_1 = (\beta_{10}, \beta_{11})^\top$ and $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21})^\top$.

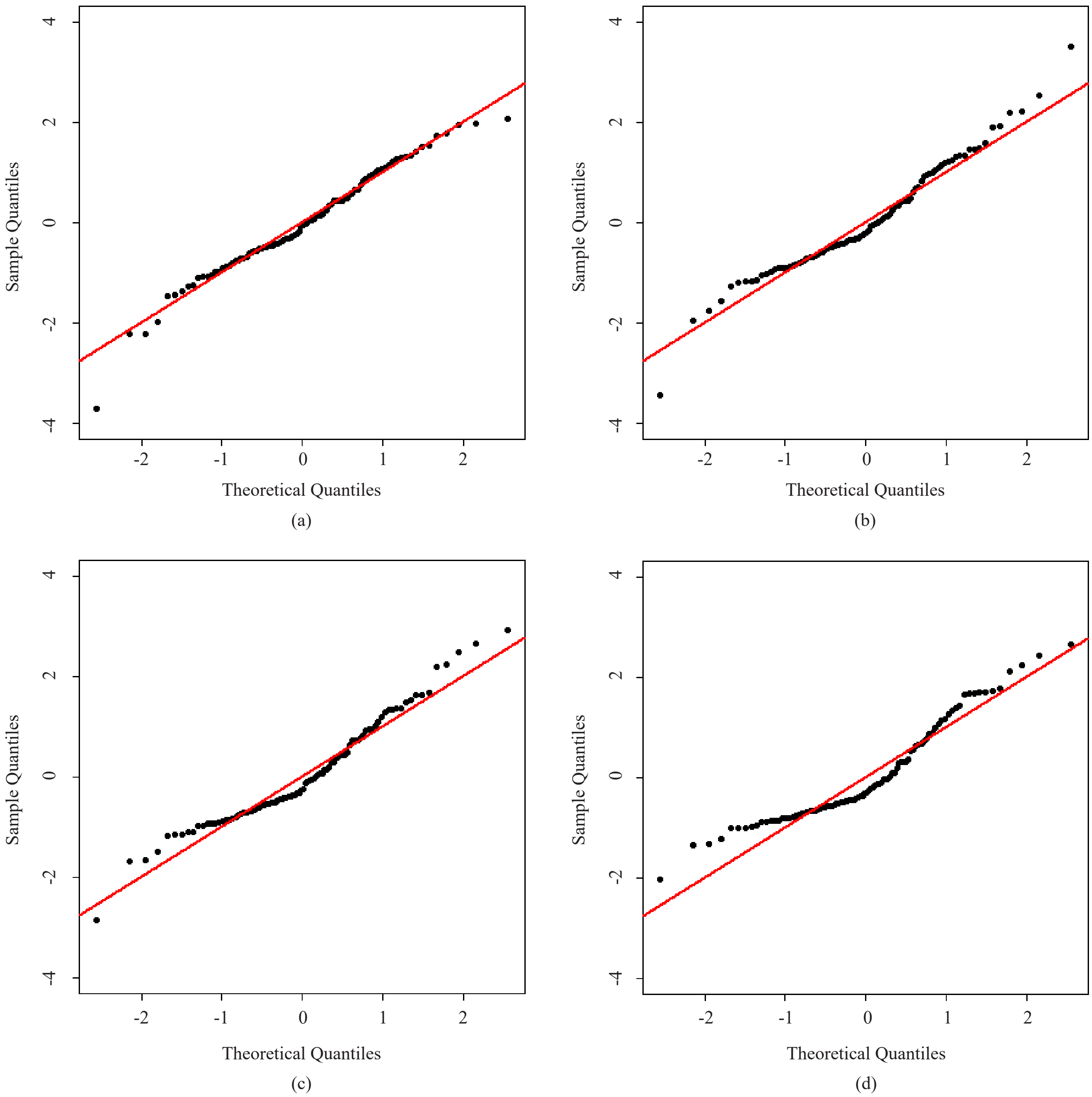


Figure 12. QQ plots from the fitted regression models to dataset IV: (a) OPAW (b) KW (c) BW (d) Weibull

Table 10 lists three statistics, thus showing that the OPAW regression model is the best for the current data. The QQ plots of the qrs for both models in Figure 12, and the plots of the empirical and estimated survival functions in Figure 13 support this conclusion. Table 11 provides the estimates, their SEs and p-values. We note that the covariate two-Arm is significant for both parameters. The residual index plot (Figure 14a), and the normal probability plot of the residuals with simulated envelope [36] (Figure 14b) show no evidence against the regression assumptions. So, the OPAW model is adequate to Efron's data.

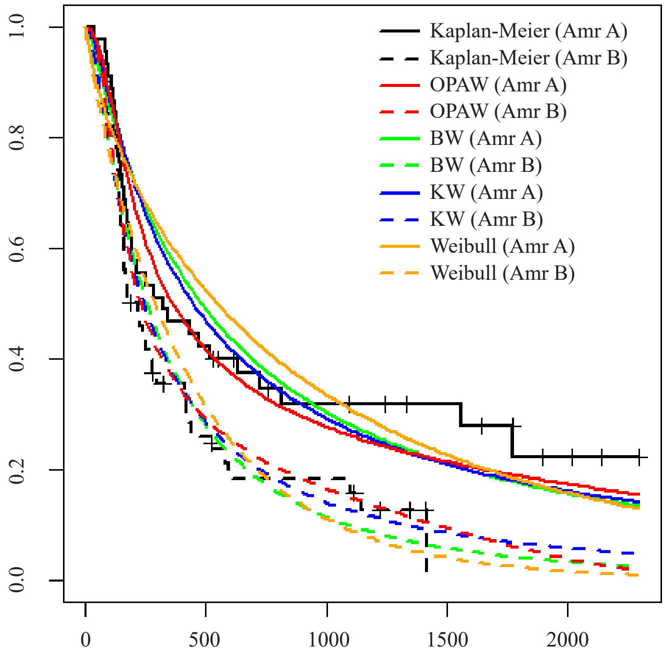


Figure 13. Dataset IV: Estimated survival and empirical functions

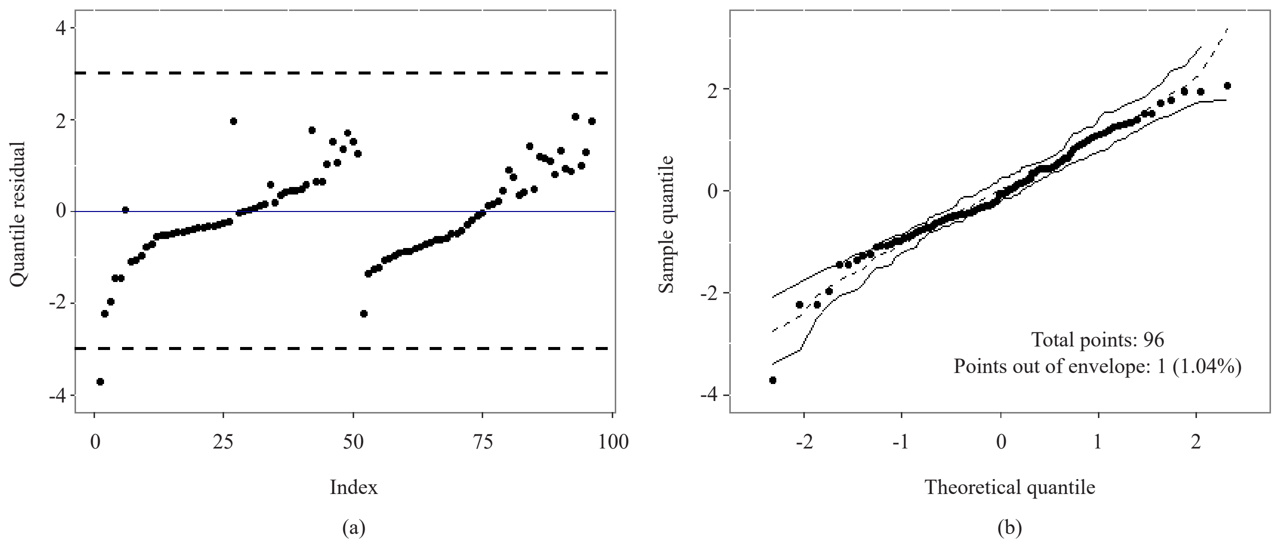


Figure 14. (a) Index plot of the qrs. (b) QQ plot for the qrs with envelope

Table 10. Adequacy measures for some regression models fitted to dataset IV

Model	AIC	CAIC	BIC
OPAW	1071.31	1081.57	1086.70
KW	1080.86	1096.68	1096.24
BW	1085.14	1100.96	1100.53
Weibull	1096.71	1116.53	1106.96

Table 11. Findings from the OPAW regression fitted to dataset IV

	MLEs	SEs	p-values
β_{10}	7.7440	0.2229	0.0000
β_{11}	-0.7642	0.2665	0.0051
β_{20}	0.9386	0.0737	0.0000
β_{21}	0.1812	0.0865	0.0388
$\log(a)$	-1.4866	0.1570	
$\log(c)$	-7.7486	0.3614	

9. Conclusions

Significant progress has been made in the last years towards constructing flexible lifetime regression models among applied statisticians. We provide a mathematical treatment of the distribution including the density of the quantile measures and give infinite expansions for the rth moment which hold in generality for any parameter values. We proposed the Odd Pareto-G (OPAG) family of distributions, and developed a regression model based on it. The inference was conducted based on likelihood theory. Four applications to real data proved empirically the utility of the new models. The main advantage of the OPAG family is that it can model bimodal data instead of using mixtures of distributions. Another advantage is the associated regression model presented in applications III and IV. For future researches, other regression models can be developed, such as random effects, semiparametric, bivariate regression models based on the new family.

Acknowledgments

We gratefully acknowledge financial support from CAPES and CNPq.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] Marshall AN, Olkin I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*. 1997; 84(3): 641-652. Available from: doi: 10.1093/biomet/84.3.641.

[2] Gupta RC, Gupta PL, Gupta RD. Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and Methods*. 1998; 27(4): 887-904. Available from: doi: 10.1080/03610929808832134.

[3] Eugene N, Lee C, Famoye F. Beta-normal distribution and its applications. *Commununications in Statistics, Theory and Methods*. 2002; 31(4): 497-512. Available from: doi: 10.1081/STA-120003130.

[4] Gleaton JU, Lynch JD. Properties of generalized log-logistic families of lifetime distributions. *Journal of Probability and Statistical Science*. 2006; 4(1): 51-64.

[5] Cordeiro GM, de Castro M. A new family of generalized distributions. *Journal of Statistical Computation and Simulation*. 2011; 81(7): 883-898. Available from: doi: 10.1080/00949650903530745.

[6] Alexander C, Cordeiro GM, Ortega EM, Sarabia JM. Generalized beta-generated distributions. *Computational Statistics & Data Analysis*. 2012; 56(6): 1880-1897. Available from: doi: 10.1016/j.csda.2011.11.015.

[7] Cordeiro GM, Ortega EMM, da Cunha DCC. The exponentiated generalized class of distributions. *Journal of Data Science*. 2013; 11(1): 1-27. Available from: doi: 10.6339/JDS.2013.11(1).1086.

[8] Alzaatreh A, Lee C, Famoye F. A new method for generating families of continuous distributions. *Metron*. 2013; 71(1): 63-79. Available from: doi: 10.1007/s40300-013-0007-y.

[9] Bourguignon M, Silva RB, Cordeiro GM. The Weibull-G family of probability distributions. *Journal of Data Science*. 2014; 12(1): 53-68. Available from: doi: 10.6339/JDS.2014.12(1).1210.

[10] Cordeiro GM, de Santana LH, Ortega EM, Pescim RR. A new family of distributions: Libby-Novick beta. *International Journal of Statistics and Probability*. 2014; 3(2): 63-80. Available from: doi: 10.5539/ijsp.v3n2p63.

[11] Cordeiro GM, Alizadeh M, Ortega EMM. The exponentiated half-logistic family of distributions: properties and applications. *Journal of Probability and Statistics*. 2014; 1-21. Available from: doi: 10.1155/2014/864396.

[12] Tahir MH, Cordeiro GM, Alzaatreh A, Zubair M, Mansoor M. The Logistic-X family of distributions and its applications. *Commununications in Statistics - Theory Methods*. 2014; 45(24): 7326-7349. Available from: doi: 10.1080/03610926.2014.980516.

[13] Cordeiro GM, Ortega EMM, Popović B, Pescim RR. The Lomax generator of distributions: properties, minification process and regression model. *Applied Mathematics and Computation*. 2014; 247: 465-486. Available from: doi: 10.1016/j.amc.2014.09.004.

[14] Cordeiro GM, Alizadeh M, Ozel G, Hosseini B, Ortega EMM, Altun E. The generalized odd log-logistic family of distributions: properties, regression models and applications. *Journal of Statistical Computation and Simulation*. 2017; 87(5): 908-932. Available from: doi: 10.1080/00949655.2016.1238088.

[15] Hussein M, Cordeiro GM. A modified power family of distributions: Properties, simulations and applications. *Mathematics*. 2022; 10(7): 1035. Available from: doi: 10.3390/math10071035.

[16] Hussein M, Elsayed H, Cordeiro GM. A new family of continuous distributions: Properties and estimation. *Symmetry*. 2022; 14(2): 276. Available from: doi: 10.3390/sym14020276.

[17] Liu B, Ananda M. A generalized family of exponentiated composite distributions. *Mathematics*. 2022; 10(11): 1895. Available from: doi: 10.3390/math10111895.

[18] Sharma VK, Singh SV, Shekhawat K. Exponentiated Teissier distribution with increasing, decreasing and bathtub hazard functions. *Journal of Applied Statistics*. 2022; 49(2): 371-393. Available from: doi: 10.1080/02664763.2020.1813694.

[19] Al-Hussaini EK. Composition of cumulative distribution functions. *Journal of Statistical Theory and Applications*. 2012; 11(4): 323-336.

[20] Gomes-Silva FS, Percontini A, Brito E, Ramos MW, Venâncio R, Cordeiro GM. The odd Lindley-G family of distributions. *Austrian Journal of Statistics*. 2017; 46(1): 65-87. Available from: doi: 10.17713/ajs.v46i1.222.

[21] Jayakumar K, Mathew T. On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution. *Statistical Papers*. 2008; 49(2): 421-439. Available from: doi: 10.1007/s00362-006-0024-5.

[22] Handique L, Chakraborty S, de Andrade TA. The exponentiated generalized Marshall-Olkin family of distribution:

its properties and applications. *Annals of Data Science*. 2019; 6(3): 391-411. Available from: doi: 10.1007/s40745-018-0166-z.
[23] Tahir MH, Nadarajah S. Parameter induction in continuous univariate distributions: Wellestablished G families. *Anais da Academia Brasileira de Ciências*. 2015; 87(2): 539-568. Available from: doi: 10.1590/0001-3765201520140299.
[24] Gao D, Ye J, Hu Y, Dong Y, Zhu Z, Yuan H, Wang S. Application of Lorenz-curve model to stratified water injection evaluation. *Petroleum Exploration and Development*. 2015; 42(6): 861-868. Available from: doi: 10.1016/S1876-3804(15)30083-5.
[25] Giorgi GM, Crescienzi M. A look at the Bonferroni inequality measure in a reliability framework. *Statistica*. 2001; 61(4): 571-583. Available from: doi: 10.6092/issn.1973-2201/6824.
[26] Bowley AL. *Elements of Statistics*. 4th ed. London: P. S. King & Son, Ltd.; New York: C. Scribner's Sons; 1920.
[27] Moors JJA. The meaning of kurtosis: Darlington reexamined. *The American Statistician*. 1986; 40(4): 283-284. Available from: doi: 10.1080/00031305.1986.10475415.
[28] Team RC. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2022. Available from: https://www.r-project.org/ [Accessed 24th June 2022].
[29] Dunn P, Smyth G. Randomized quantile residuals. *Journal of Computational and Graphical Statistics*. 1996; 5(3): 236-244. Available from: doi: 10.1080/10618600.1996.10474708.
[30] Amini M, MirMostafaee S, Ahmadi J. Log-gamma-generated families of distributions. *Statistics: A Journal of Theoretical and Applied Statistics*. 2014; 48(4): 913-932. Available from: doi: 10.1080/02331888.2012.748775.
[31] StatLib — JASA Data Archive. Available from: http://lib.stat.cmu.edu/jasadata/laslett [Accessed 14th May 2022].
[32] Kundu D, Raqab MZ. Estimation of $R = P(Y < X)$ for three parameter Weibull distribution. *Statistics & Probability Letters*. 2009; 79(17): 1839-1846. Available from: doi: 10.1016/j.spl.2009.05.026.
[33] Efron B. Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American Statistical Association*. 1988; 83(402): 414-425. Available from: doi: 10.1080/01621459.1988.10478612.
[34] Mudholkar GS, Srivastava DK, Kollia GD. A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*. 1996; 91(436): 1575-1583. Available from: doi: 10.1080/01621459.1996.10476725.
[35] Cordeiro GM, Nadarajah S, Ortega EM. General results for the beta Weibull distribution. *Journal of Statistical Computation and Simulation*. 2013; 83(6): 1082-1114. Available from: doi: 10.1080/00949655.2011.649756.
[36] Atkinson AC. *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostics Regression Analysis*. 2nd ed. Oxford: Clarendon Press; 1987.

Appendix A

For the OPAW(a, c_1, b, k) and OPAGA(a, c_1, b, k) models, the observed information matrix is given by

$$J(\theta) = J\left(a, c_1, b, k\right) = -\begin{bmatrix} L_{aa} & L_{ac_1} & L_{ab} & L_{ak} \\ L_{ac_1} & L_{c_1c_1} & L_{c_1b} & L_{c_1k} \\ L_{ab} & L_{c_1b} & L_{bb} & L_{bk} \\ L_{ak} & L_{c_1k} & L_{bk} & L_{kk} \end{bmatrix},$$

whose elements are

$$L_{aa} = -\frac{n}{a^2},$$

$$L_{ac_1} = -\sum_{i=1}^{n} \frac{G(x_i)}{\overline{G}(x_i) + c_1 G(x_i)},$$

$$L_{ab} = \sum_{i=1}^{n} \left[\frac{1-c_1}{\overline{G}(x_i) + c_1 G(x_i)} - \frac{1}{\overline{G}(x_i)} \right] \frac{\partial G(x_i)}{\partial b},$$

$$L_{ak} = \sum_{i=1}^{n} \left[\frac{1-c_1}{\overline{G}(x_i) + c_1 G(x_i)} - \frac{1}{\overline{G}(x_i)} \right] \frac{\partial G(x_i)}{\partial k},$$

$$L_{c_1c_1} = -\frac{n}{c_1^2} + (a+1) \sum_{i=1}^{n} \left[\frac{G(x_i)}{\overline{G}(x_i) + c_1 G(x_i)} \right]^2,$$

$$L_{c_1b} = -(a+1) \sum_{i=1}^{n} \frac{1}{\left[\overline{G}(x_i) + c_1 G(x_i) \right]^2} \frac{\partial G(x_i)}{\partial b},$$

$$L_{c_1k} = -(a+1) \sum_{i=1}^{n} \frac{1}{\left[\overline{G}(x_i) + c_1 G(x_i) \right]^2} \frac{\partial G(x_i)}{\partial k},$$

$$L_{bb} = (a+1) \left\{ \sum_{i=1}^{n} \left[\frac{1-c_1}{\overline{G}(x_i) + c_1 G(x_i)} - \frac{1}{\overline{G}(x_i)} \right] \frac{\partial^2 G(x_i)}{\partial b^2} \right.$$

$$\left. + \sum_{i=1}^{n} \left[\frac{(1-c_1)^2}{[\overline{G}(x_i) + c_1 G(x_i)]^2} - \frac{1}{\overline{G}(x_i)^2} \right] \frac{\partial^2 G(x_i)}{\partial b^2} \right\}$$

$$+ \sum_{i=1}^{n} \frac{1}{g(x_i)} \frac{\partial^2 g(x_i)}{\partial b^2} - \sum_{i=1}^{n} \left[\frac{1}{g(x_i)} \frac{\partial g(x_i)}{\partial b} \right]^2$$

$$+2\left\{\sum_{i=1}^{n}\frac{1}{\bar{G}(x_i)}\frac{\partial^2 G(x_i)}{\partial b^2}+\sum_{i=1}^{n}\left[\frac{1}{\bar{G}(x_i)}\frac{\partial G(x_i)}{\partial b}\right]^2\right\},$$

$$L_{bk}=(a+1)\left\{\sum_{i=1}^{n}\left[\frac{1-c_1}{\bar{G}(x_i)+c_1 G(x_i)}-\frac{1}{\bar{G}(x_i)}\right]\frac{\partial^2 G(x_i)}{\partial k\partial b}\right.$$

$$\left.+\sum_{i=1}^{n}\left[\frac{(1-c_1)^2}{[\bar{G}(x_i)+c_1 G(x_i)]^2}-\frac{1}{\bar{G}(x_i)^2}\right]\frac{\partial G(x_i)}{\partial k}\frac{\partial G(x_i)}{\partial b}\right\}$$

$$+\sum_{i=1}^{n}\frac{1}{g(x_i)}\frac{\partial^2 g(x_i)}{\partial k\partial b}-\sum_{i=1}^{n}\frac{1}{g(x_i)^2}\frac{\partial g(x_i)}{\partial k}\frac{\partial g(x_i)}{\partial b}$$

$$+2\left\{\sum_{i=1}^{n}\frac{1}{\bar{G}(x_i)}\frac{\partial^2 G(x_i)}{\partial k\partial b}+\sum_{i=1}^{n}\frac{1}{\bar{G}(x_i)^2}\frac{\partial G(x_i)}{\partial k}\frac{\partial G(x_i)}{\partial b}\right\},$$

$$L_{kk}=(a+1)\left\{\sum_{i=1}^{n}\left[\frac{1-c_1}{\bar{G}(x_i)+c_1 G(x_i)}-\frac{1}{\bar{G}(x_i)}\right]\frac{\partial^2 G(x_i)}{\partial k^2}\right.$$

$$\left.+\sum_{i=1}^{n}\left[\frac{(1-c_1)^2}{[\bar{G}(x_i)+c_1 G(x_i)]^2}-\frac{1}{\bar{G}(x_i)^2}\right]\frac{\partial^2 G(x_i)}{\partial k^2}\right\}$$

$$+\sum_{i=1}^{n}\frac{1}{g(x_i)}\frac{\partial^2 g(x_i)}{\partial k^2}-\sum_{i=1}^{n}\left[\frac{1}{g(x_i)}\frac{\partial g(x_i)}{\partial k}\right]^2$$

$$+\left\{\sum_{i=1}^{n}\frac{1}{\bar{G}(x_i)}\frac{\partial^2 G(x_i)}{\partial k^2}+\sum_{i=1}^{n}\left[\frac{1}{\bar{G}(x_i)}\frac{\partial G(x_i)}{\partial k}\right]^2\right\}.$$