STATISTICS IN TRANSITION new series, September 2021
Vol. 22, No. 3, pp. 99–121, DOI 10.21307/stattrans-2021-029
Received – 24.01.2020; accepted – 01.06.2021

A new reciprocal Rayleigh extension: properties, copulas, different methods of estimation and a modified right-censored test for validation

Haitham M. Yousof[1], M. Masoom Ali[2], Hafida Goual[3], Mohamed Ibrahim[4]

ABSTRACT

In this article, a new reciprocal Rayleigh extension called the Xgamma reciprocal Rayleigh model is defined and studied. The relevant statistical properties are derived, and the useful results related to the convexity and concavity are addressed. We discussed the estimation of the parameters using different estimation methods such as the maximum likelihood estimation method, the ordinary least squares estimation method, the weighted least squares estimation method, the Cramer-Von-Mises estimation method, and the bootstrapping method. A simulation study was conducted to assess the performances of the proposed estimation methods are investigated through a simulation study. Many bivariate and multivariate type model have also been derived based on Farlie-Gumbel-Morgenstern copula, the Clayton copula, Renyi's entropy copula and the Ali-Mikhail-Haq copula. A modified Nikulin-Rao-Robson test for right-censored validation is applied to a censored real data set.

Key words: Xgamma model, reciprocal Rayleigh model, simulations, bootstrapping, Farlie Gumbel Morgenstern copula, least squares, Cramer-Von-Mises, bootstrapping, Ali-Mikhail-Haq copula, convexity, concavity.

1. Introduction

The probability density function (PDF) and cumulative distribution function (CDF) of the reciprocal Rayleigh (RR) distribution are given, respectively, by

$$g_{\theta_2}(y) = 2\theta_2^2 y^{-3} e^{-\left(\frac{\theta_2}{y}\right)^2}|_{y\in\mathbb{R}^+},$$

[1] Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Benha University, Benha 13518, Egypt. E-mail: haitham.yousof@fcom.bu.edu.eg. ORCID: https://orcid.org/0000-0003-4589-4944.

[2] Department of Mathematical Sciences Ball State University, Muncie, Indiana 47306, USA. E-mail: mali@bsu.edu. ORCID: https://orcid.org//0000-0002-0120-9442.

[3] Laboratory of Probability and Statistics, University of Badji Mokhtar, Annaba, Algeria. E-mail: goual.hafida@gmail.com. ORCID: https://orcid.org/0000-0003-4932-4332.

[4] Department of Applied, Mathematical and Actuarial Statistics, Faculty of Commerce, Damietta University, Damietta, Egypt. E-mail: mohamed_ibrahim@du.edu.eg. ORCID: https://orcid.org/0000-0003-4893-9669.

and

$$G_{\theta_2}(y) = e^{-\left(\frac{\theta_2}{y}\right)^2}|_{y\in\mathbb{R}^+},$$

where $\theta_2 > 0$ refers to the scale parameter. The RR model is a special case from the well-known inverse Weibull distribution. The RR model was originally proposed by Fréchet (1927). It has many applications in accelerated life testing, earthquakes, floods, wind speed, horse racing, rainfall, queues in supermarkets, and sea waves. Gusmao et al. (2011) defined and studied the generalized reciprocal Rayleigh (GRR) distribution. Krishna et al. (2013) proposed some applications of the Marshall-Olkin reciprocal Rayleigh (MORR) distribution. Mahmoud and Mandouh (2013) proposed and studied the transmuted reciprocal Rayleigh (TRR) distribution. Haq et al. (2017) presented a new four-prarameter reciprocal Rayleigh version for modeling extreme values. Korkmaz et al. (2017) studied some theoretical and computational aspects of the odd Lindley reciprocal Rayleigh (OLRR) distribution. Yousof et al. (2018d) defined a new family called the odd reciprocal Rayleigh G (ORR-G) family of distributions. Yousof et al. (2019) defined a new compound version of the reciprocal Rayleigh (OBRR) distribution. Salah et al. (2020) defined and studied a new version of RR model called the odd Burr RR model with different copula, different estimation methods, applications and validation testing. Recently, Cordeiro (2020) proposed and studied the Xgamma-G (Xg-G) family of distribution with CDF and PDF (for $\theta_1 > 0$) given by

$$F_{\theta_1,\underline{\xi}}(y) = 1 - \frac{1+\theta_1-\theta_1\log\left[1-G_{\underline{\xi}}(y)\right]+\frac{1}{2}\theta_1^2\left\{\log\left[1-G_{\underline{\xi}}(y)\right]\right\}^2}{1+\theta}\left[1-G_{\underline{\xi}}(y)\right]^{\theta_1}|_{y\in\mathbb{R}}, \tag{1}$$

and

$$f_{\theta_1,\underline{\xi}}(y) = \frac{\theta_1}{1+\theta_1}g_{\underline{\xi}}(y)\left[1-G_{\underline{\xi}}(y)\right]^{\theta_1-1}\left(\theta+\frac{1}{2}\theta_1^2\left\{\log\left[1-G_{\underline{\xi}}(y)\right]\right\}^2\right)|_{y\in\mathbb{R}}, \tag{2}$$

respectively, where $g_{\underline{\xi}}(y)$ and $G_{\underline{\xi}}(y)$ are the baseline PDF and CDF respectively with a parameter vector $\underline{\boldsymbol{\xi}}$. To this end, we define the CDF of the Xgamma reciprocal Rayleigh (XgRR) model. Using (1), the CDF of the XgRR can be written as

$$F_{\theta_1,\theta_2}(y) = 1 - \frac{1}{1+\theta_1}\left[1-e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1}\left(\begin{array}{c}1+\theta_1-\theta_1\log\left[1-e^{-\left(\frac{\theta_2}{y}\right)^2}\right]\\ +\frac{1}{2}\theta_1^2\left\{\log\left[1-e^{-\left(\frac{\theta_2}{y}\right)^2}\right]\right\}^2\end{array}\right)|_{y\in\mathbb{R}^+}. \tag{3}$$

The PDF corresponding to (3) reduces to

$$f_{\theta_1,\theta_2}(y) = 2\theta_2^2\frac{\theta_1 y^{-3}e^{-\left(\frac{\theta_2}{y}\right)^2}}{1+\theta_1}\left[1-e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1-1}\left(\theta_1+\frac{1}{2}\theta_1^2\left\{\log\left[1-e^{-\left(\frac{\theta_2}{y}\right)^2}\right]\right\}^2\right)|_{y\in\mathbb{R}^+}. \tag{4}$$

The XgRR family density in (4) can be expressed as

$$f_{\theta_1,\theta_2}(y) = \frac{2\theta_1^2\theta_2^2}{1+\theta_1} y^{-3} e^{-\left(\frac{\theta_2}{y}\right)^2} \left[1 - e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1 - 1}$$

$$+ \left(\frac{\theta_1^3}{2(1+\theta_1)} \overbrace{2\theta_2^2 y^{-3} e^{-\left(\frac{\theta_2}{y}\right)^2}}^{g_{\theta_2}(y)} \left[1 - e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1 - 1} \underbrace{\left\{\log\left[1 - e^{-\left(\frac{\theta_2}{y}\right)^2}\right]\right\}^2}_{A(y;\theta_2)} \right) |_{y \in \mathbb{R}^+}. \tag{5}$$

Consider

$$\log\left(1 - \frac{a_1}{a_2}\right) = -\sum_{i=0}^{\infty} \frac{1}{i+1} \left(\frac{a_1}{a_2}\right)^{i+1} |_{\left|\frac{a_1}{a_2}\right| < 1}, \tag{6}$$

and the power series raised to a positive integer n (see Gradshteyn and Ryzhik (2002))

$$\left(\sum_{j=0}^{\infty} a_j u^j\right)^n = \sum_{j=0}^{\infty} c_{(n,j)} u^j, \tag{7}$$

where the coefficients $c_{(n,j)}$ (for $j = 1,2, \dots$) can be easily determined from the recurrence equation

$$c_{(n,j)} = (j a_0)^{-1} \sum_{m=1}^{j} [m(n+1) - j]\, a_m c_{(n,j-m)} \text{ and } c_{(n,0)} = a_0^n.$$

The coefficient $c_{(n,j)}$ can be calculated from $c_{(n,0)}, \dots, c_{(n,j-1)}$ and hence from the quantities $a_0, \dots, a_j$. For $\left|\frac{a_1}{a_2}\right| < 1$ and $a_3 > 0$, the power series holds

$$\left(1 - \frac{a_1}{a_2}\right)^{a_3} = \sum_{j=0}^{\infty} \frac{\Gamma(1+a_3)}{j!\,\Gamma(a_3 - j + 1)} \left(-\frac{a_1}{a_2}\right)^j. \tag{8}$$

Applying (6) to the quantity $A(y;\theta_2)$ in the PDF in (5), the PDF can be expressed as

$$f_{\theta_1,\theta_2}(y) = \frac{\theta_1^2}{1+\theta_1} \overbrace{2\theta_2^2 y^{-3} e^{-\left(\frac{\theta_2}{y}\right)^2}}^{g_{\theta_2}(y)} \left[1 - e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1 - 1}$$

$$+ \left(\frac{\theta_1^3}{2(1+\theta_1)} \overbrace{2\theta_2^2 y^{-3} e^{-\left(\frac{\theta_2}{y}\right)^2}}^{g_{\theta_2}(y)} \left[1 - e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1 - 1} \left[e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^2 \underbrace{\left\{\sum_{i=0}^{\infty} \frac{1}{i+1} e^{-i\left(\frac{\theta_2}{y}\right)^2}\right\}^2}_{B(y;\theta_2)} \right).$$

Expanding the quantity $B(y;\theta_2)$ using (7), the $f_{\theta_1,\theta_2}(y)$ can be written as

$$f_{\theta_1,\theta_2}(y)=\frac{\theta_1^2}{1+\theta_1}\overbrace{2\theta_2^2 y^{-3}e^{-\left(\frac{\theta_2}{y}\right)^2}}^{g_{\theta_2}(y)}\underbrace{\left[1-e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1-1}}_{C(y;\theta_1,\theta_2)}$$

$$+\left\{\frac{\theta_1^3}{2(1+\theta_1)}\overbrace{2\theta_2^2 y^{-3}e^{-\left(\frac{\theta_2}{y}\right)^2}}^{g_{\theta_2}(y)}\sum_{i=0}^{\infty}c_{2,i}\,e^{-(2+i)\left(\frac{\theta_2}{y}\right)^2}\underbrace{\left[1-e^{-\left(\frac{\theta_2}{y}\right)^2}\right]^{\theta_1-1}}_{C(y;\theta_1,\theta_2)}\right\},$$

where $c_{2,i}=1/(i+1)$. Applying the power series (8) to the quantity $C(y;\theta_1,\theta_2)$, we obtain

$$f(y)=\sum_{j=0}^{\infty}\left[\nabla_j\,\pi_{1+j}(y)+\sum_{i=0}^{\infty}\nabla_{i,j}\;\pi_{3+i+j}(y)\right], \tag{9}$$

where

$$\nabla_j=\frac{(-1)^j\theta_1^2\Gamma(\theta_1)}{(1+j)(1+\theta_1)\Gamma(\theta_1-j)},\ \nabla_{i,j}=\frac{(-1)^j\theta_1^3\Gamma(\theta_1)\ c_{2,i}}{2(1+\theta_1)(3+i+j)j!\,\Gamma(\theta_1-j)},$$

and $\pi_\varsigma(y)$ is the RR density with scale parameter $\theta_2\varsigma^{\frac{1}{2}}$ and shape parameter 2. So, the density of Y is a linear combination of RR densities. The CDF of Y follows by integrating (8) as

$$F_{\theta_1,\theta_2}(y)=\sum_{j=0}^{\infty}\left[\nabla_j\,H_{1+j}(y)+\sum_{i=0}^{\infty}\nabla_{i,j}\;H_{3+i+j}(y)\right], \tag{10}$$

where $H_\varsigma(y)$ is the RR density with scale parameter $\theta_2\varsigma^{\frac{1}{2}}$ and shape parameter 2. Equations (9) and (10) are the main results of this section. We provide some plots of the PDF and hazard rate function (HRF) of the XgRR model to show its flexibility. Figure 1 displays some plots of the XgRR density for selected parameter values. These plots reveal that the new density can be right skewed with different flexible shapes. The HRF plots of the XgRR distribution can be upside down or increasing. Many other useful real data sets can be found in Aryal and Yousof (2017), Merovci et al. (2017 and 2020), Korkmaz et al. (2017), Hamedani et al. (2017), Brito et al. (2017), Alizadeh et al. (2018), Korkmaz et al. (2018), Yousof et al. (2018a-d), Hamedani et al. (2018), Cordeiro et al. (2018), Hamedani et al. (2019), Ibrahim (2019), Nascimento et al. (2017, 2018 and 2019), Ibrahim et al. (2019), Goual and Yousof (2019), Korkmaz et al. (2019), Alizadeh et al. (2019) and Goual et al. (2020), Ibrahim (2020 a and b) and Yadav et al. (2020).

After studying the mathematical properties of the XgRR model, we discussed the estimation of the parameters using different estimation methods such as maximum likelihood estimation method, ordinary least squares estimation method, weighted least-squares estimation method, Cramer-Von-Mises estimation method and bootstrapping method. Then, the Nikulin-Rao-Robson (N.R.R) statistic and modified N.R.R are discussed. In particular, the modified chi-squared test for composite hypothesis for complete samples was first considered by Nikulin (1973a, 1973b and 1973c), Rao and Robson (1974), among others. On the other hand, several goodness-of-fit tests have been suggested by the statisticians for censored data.

2. Properties

2.1. Moments

Let Y_ς be a rv having density $\pi_\varsigma(y)$. The r^{th} ordinary moment of Y, say $\mu'_{r,Y}$, follows from (9) as

$$\mu'_{r,Y} = E(Y^r) = \sum_{j=0}^{\infty}\left[\nabla_j\, E\left(Y_{1+j}^r\right) + \sum_{i=0}^{\infty}\nabla_{i,j}\; E\left(Y_{3+i+j}^r\right)\right].$$

Therefore,

$$\mu'_{r,Y} = \theta_2^r\Gamma\left(1-\frac{r}{2}\right)\sum_{j=0}^{\infty}\left[\nabla_j^{(r,1+j)} + \sum_{i=0}^{\infty}\nabla_{i,j}^{(r,3+i+j)}\right]\,|_{(2>r)}, \tag{11}$$

where

$$\nabla_j^{(r,1+j)} = b_j\,(1+j)^{\frac{r}{2}} \text{ and } \nabla_{i,j}^{(r,3+i+j)} = b_{i,j}\,(3+i+j)^{\frac{r}{2}},$$

and

$$\Gamma(1+\tau)|_{(\tau\in R^+)} = \tau! = \prod_{w=0}^{\tau-1}(\tau-w) = \int_0^\infty y^\tau\, eyp(-y)\, dy.$$

Setting $r = 1$ in (11) gives the mean of Y

$$E(Y) = \theta_2\Gamma\left(1-\frac{1}{2}\right)\sum_{j=0}^{\infty}\left[\nabla_j\,(1+j)^{\frac{1}{2}} + \sum_{i=0}^{\infty}\nabla_{i,j}\;(3+i+j)^{\frac{1}{2}}\right].$$

2.2. Incomplete moments

The r^{th} incomplete moment of Y is defined by

$$m_{r,Y}(t) = \int_{-\infty}^{t} y^r\, f(y)dy.$$

We can write from (9)

$$m_{r,Y}(t) = \sum_{j=0}^{\infty}\left[\nabla_j\, m_{r,Y,1+j}(y) + \sum_{i=0}^{\infty}\nabla_{i,j}\; m_{r,Y,3+i+j}(y)\right].$$

Therefore,

$$m_{r,Y}(t) = \theta_2^r\sum_{j=0}^{\infty}\left[\begin{array}{c}\nabla_j^{(r,1+j)}\gamma\left(1-\frac{r}{2},(1+j)\left(\frac{\theta_2}{t}\right)^2\right)\\ +\sum_{i=0}^{\infty}\nabla_{i,j}^{(r,3+i+j)}\gamma\left(1-\frac{r}{2},(3+i+j)\left(\frac{\theta_2}{t}\right)^2\right)\end{array}\right]|_{(2>r)}, \tag{12}$$

where

$$\gamma(\tau,q)=\int_0^q t^{\tau-1}e^{-t}dt=\frac{q^\tau}{\tau}\{F_{1:1}[\tau;\tau\theta_2+1;-q]\}=\sum_{j=0}^{\infty}\frac{(-1)^j q^{\tau+j}}{j!\,(\tau+j)}=\Gamma(\tau)-\Gamma(\tau,q),$$

$$\Gamma(\tau,q)|_{(y>0)}=\int_q^\infty t^{\tau-1}\,eyp(-t)\,dt,$$

and $F_{1:1}[\cdot,\cdot,\cdot]$ is a confluent hypergeometric function (see Johnson et al. (2005)). Setting $r=1$ in (12) gives the first incomplete moment

$$m_{1,Y}(t)=\theta_2\sum_{j=0}^{\infty}\left[\begin{array}{c}\nabla_j^{(1,1+j)}\gamma\left(\frac{1}{2},(1+j)\left(\frac{\theta_2}{t}\right)^2\right)\\+\sum_{i=0}^{\infty}\nabla_{i,j}^{(1,3+i+j)}\gamma\left(\frac{1}{2},(3+i+j)\left(\frac{\theta_2}{t}\right)^2\right)\end{array}\right]$$

Two important applications of the $m_{1,Y}(t)$ are related to the mean deviation about the mean (S_1) and median and to the Bonferroni and Lorenz curves. The mean deviation about the mean

$$S_{1,Y}=E(|Y-E(Y)|)=2\mu'_{1,Y}F(E(Y))-2m_{1,Y}(E(Y))$$

and about the median

$$S_{2,Y}=E(|Y-M|)=E(Y)-2m_{1,Y}(M)$$

where $E(Y), M=Q(u)=F^{-1}(u)$ is the median of Y, $F(\mu'_{1,Y})$ is easily calculated.

2.3. Moment generating function

The moment generating function (MGF) of Y, say $M_Y(t)=E(e^{tY})$, is obtained from (9) as

$$M_Y(t)=\sum_{j=0}^{\infty}\left[\nabla_j\,M_{Y,1+j}(t)+\sum_{i=0}^{\infty}\nabla_{i,j}\;M_{Y,3+i+j}(t)\right].$$

Therefore,

$$M_Y(t)=\sum_{j,r=0}^{\infty}\left[\begin{array}{c}\nabla_j[(t\theta_2)^r/r!](1+j)^{\frac{r}{2}}\Gamma\left(1-\frac{r}{2}\right)\\+\sum_{i=0}^{\infty}\nabla_{i,j}\;[(t\theta_2)^r/r!](3+i+j)^{\frac{r}{2}}\Gamma\left(1-\frac{r}{2}\right)\end{array}\right]|_{(2>r)}.$$

2.4. Convexity and concavity

Convex densities play an important role in several areas of mathematics. They are important in studying the “problems of optimization” where they are distinguished by several convenient characteristics. In mathematical analysis, a certain density defined on a certain n-dimensional interval is called “convex density” if the line between any two points on the graph of the density lies above the graph between the two points. The PDF in (5) is said to be “concave density” if for any $Y_1\sim$ XgRR (θ_1,θ_2) and $Y_2\sim$ XgRR (c_1,c_2) the PDF satisfies

$$f(\mathbf{b}y_1+\bar{\mathbf{b}}y_2)\geq\mathbf{b}f_{\theta_1,\theta_2}(y_1)+\bar{\mathbf{b}}f_{c_1,c_2}(y_2)|_{0\leq\mathbf{b}\leq1\text{ and }\bar{\mathbf{b}}=1-\mathbf{b}}.$$

If the function $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is twice differentiable, then if $f^{//}(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2) < 0, \forall\, y \in \mathbb{R}^+$, $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "strictly convex density". If $f^{//}(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2) \leq 0, \forall\, y \in \mathbb{R}^+$, then $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "convex". The density in (5) is said to be "convex density" if for any $Y_1 \sim$ XgRR (θ_1, θ_2) and $Y_2 \sim$ XgRR (c_1, c_2) the density satisfies

$$f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2) \leq \mathbf{b}f_{\theta_1,\theta_2}(y_1) + \bar{\mathbf{b}}f_{c_1,c_2}(y_2)|_{0\leq\mathbf{b}\leq 1 \text{ and } \bar{\mathbf{b}}=1-\mathbf{b}}.$$

If the function $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is twice differentiable, then if $f^{//}(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2) > 0, \forall\, y \in \mathbb{R}^+$, $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "strictly convex density". If $f^{//}(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2) \geq 0, \forall\, y \in \mathbb{R}^+$, then $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "convex". If $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "convex" and c is a constant, then the function $cf(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "convex". If $f(\mathrm{b}y_1 + \bar{\mathbf{b}}y_2)$ is "convex density", then $[cf(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)]$ is convex for every $c > 0$. If $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ and $g(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ are "convex density", then $[f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2) + g(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)]$ is also "convex density". If $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ and $g(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ are "convex density", then $[f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2).\, g(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)]$ is also "convex density". If the function $-f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "convex density", then the function $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "convex density". If $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "concave density", then $\frac{1}{f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)}$ is "convex density" if $f(y) > 0$. If $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "concave density", $\frac{1}{f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)}$ is "convex density" if $f(y) < 0$. If $f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)$ is "concave density", $\frac{1}{f(\mathbf{b}y_1 + \bar{\mathbf{b}}y_2)}$ is "convex density".

3. Copulas

For modelling the bivariate real data sets, we can consider the bivariate XgRR type generated via the FGM copula, modified FGM copula, Clayton copula and Renyi's entropy copula. Many other types of copula could be considered in separate works. In this Section, we derive some new bivariate type XgRR (BXgRR) model using the theorems of FGM copula, modified FGM copula, Clayton copula and Renyi's entropy. The Multivariate XgRR (MvXgRR) type is also presented. However, future works may be allocated to study these new models. First, we consider the joint CDF (JCDF) of the FGM family, where $C_\rho(m, w) = mw(1 + \rho m^* w^*)|_{m^*=1-m, w^*=1-w}$, where the marginal function $m = F_1 = F_{\theta_1,\theta_2}(y_1)$, $w = F_2 = F_{a_1,a_2}(y_2)$, $\rho \in (-1,1)$ is a dependence parameter and for every $m, w \in (0,1)$, $C(m, 0) = C(0, w) = 0$, which is "grounded minimum", and $C(m, 1) = m$ and $C(1, w) = w$ which is "grounded maximum", $C(m_1, w_1) + C(m_2, w_2) - C(m_1, w_2) - C(m_2, w_1) \geq 0$.

3.1. Via FGM copula

A copula is continuous in m and w where

$$|C(m_2, w_2) - C(m_1, w_1)| \leq |m_2 - m_1| + |w_2 - w_1|,$$

is the stronger Lipschitz condition. For $0 \le m_1 \le m_2 \le 1$ and $0 \le w_1 \le w_2 \le 1$, we have

$$Pr(m_1 \le M \le m_2, w_1 \le W \le w_2) \\ = C(m_1, w_1) + C(m_2, w_2) - C(m_1, w_2) - C(m_2, w_1) \ge 0.$$

Then, setting

$$m^* = S_{\theta_1,\theta_2}(y_1) = 1 - F_{\theta_1,\theta_2}(y_1)|_{[m^*=(1-m)\in(0,1)]}$$

and

$$w^* = S_{a_1,a_2}(y_2) = 1 - F_{a_1,a_2}(y_2)|_{[w^*=(1-w)\in(0,1)]},$$

we can easily obtain the JCDF of the FGM family from

$$C_\rho(y_1, y_2) = F_{\theta_1,\theta_2}(y_1)F_{a_1,a_2}(y_2)\{1 + \rho[1 - F_{\theta_1,\theta_2}(y_1)][1 - F_{a_1,a_2}(y_2)]\},$$

where

$$F_{\theta_1,\theta_2}(y_1) = 1 - \frac{\hbar_{\theta_2}(y_1)^{\theta_1}}{1+\theta_1}\begin{Bmatrix} 1 + \theta_1 - \theta_1 \log \hbar_{\theta_2}(y_1) \\ +\frac{1}{2}\theta_1^2[\log \hbar_{\theta_2}(y_1)]^2 \end{Bmatrix},$$

$$\hbar_{\theta_2}(y_1) = 1 - e^{-\left(\frac{\theta_2}{y_1}\right)^2},$$

$$F_{a_1,a_2}(y_2) = 1 - \frac{\hbar_{a_2}(y_2)^{a_1}}{1+a_1}\begin{Bmatrix} 1 + a_1 - a_1 \log \hbar_{a_2}(y_2) \\ +\frac{1}{2}\theta_1^2[\log \hbar_{a_2}(y_2)]^2 \end{Bmatrix},$$

The JPDF can then be derived from

$$c_\rho(m, w) = 1 + \rho(1 - 2m)(1 - 2w)$$

or from

$$f(y_1, y_2) = C(F_1, F_2)f_1 f_2.$$

3.2. Via modified FGM copula

The modified FGM copula is defined as

$$C_\rho(m, w) = mw[1 + \rho V(m)A(w)]|_{\rho\in(-1,1)}$$

or

$$C_\rho(m, w) = mw + \rho\dot{V}(m)\dot{A}(w)|_{\rho\in(-1,1)},$$

where $\dot{V}(m) = mV(m)$, and $\dot{A}(w) = wA(w)$, where $V(m)$ and $A(w)$ are being two continuous functions on $(0,1)$ where $V(0) = V(1) = A(0) = A(1) = 0$. Let

$$c_1 = inf\left\{\frac{\partial}{\partial m}\dot{V}(m)|\boldsymbol{q}_1(m)\right\} < 0, c_2 = sup\left\{\frac{\partial}{\partial m}\dot{V}(m)|\boldsymbol{q}_1(m)\right\} < 0,$$

$$d_1 = inf\left\{\frac{\partial}{\partial w}\dot{A}(w)|\boldsymbol{q}_2(w)\right\} > 0$$

and

$$d_2 = sup\left\{\frac{\partial}{\partial w}\dot{A}(w)|\boldsymbol{q}_2(w)\right\} > 0.$$

Then, $1 \leq min(\mathrm{c}_1\mathrm{c}_2, \mathrm{d}_1\mathrm{d}_2) \leq \infty$, where

$$\frac{\partial}{\partial m}V(m) = V(m) + m\frac{\partial}{\partial m}V(m),$$

$$\boldsymbol{q}_1(u) = \left\{m: m \in (0,1)| \frac{\partial \dot{V}(m)}{\partial m} \text{ exists}\right\}$$

and

$$\boldsymbol{q}_2(w) = \left\{w: w \in (0,1)| \frac{\partial \dot{A}(w)}{\partial w} \text{ exists}\right\}.$$

Type-I

Recall the following functional forms for both $V(m)$ and $A(w)$. Then, the BXgRR-FGM (Type-**I**) can be derived from

$$C_\rho(y_1, y_2) = F_{\theta_1,\theta_2}(y_1)F_{a_1,a_2}(y_2) + \rho\dot{V}(m)\dot{A}(y_2)|_{\rho\in(-1,1)}$$

where

$$\dot{V}(y_1) = F_{\theta_1,\theta_2}(y_1)S_{\theta_1,\theta_2}(y_1)$$

and

$$\dot{A}(y_2) = F_{a_1,a_2}(y_2)S_{a_1,a_2}(y_2).$$

Type-II

Let $V(y_1)^*$ and $A(y_2)^*$ be two functional forms satisfying all the conditions stated earlier where

$$V(y_1)^*|_{(\rho_1>0)} = S_{\theta_1,\theta_2}(y_1)^{1-\rho_1}F_{\theta_1,\theta_2}(y_1)^{\rho_1}$$

and

$$A(y_2)^*|_{(\rho_2>0)} = S_{a_1,a_2}(y_2)^{1-\rho_2}F_{a_1,a_2}(y_2)^{\rho_2}.$$

Then, the corresponding BXgRR-FGM (Type-**II**) can be derived from

$$C_{\rho,\rho_1,\rho_2}(y_1, y_2) = F_{\theta_1,\theta_2}(y_1)F_{a_1,a_2}(y_2)[1 + \rho V(y_1)^* A(y_2)^*].$$

Type-III

Let $\ddot{V}(y_1)$ and $\ddot{A}(y_2)$ be two functional forms for satisfying all the conditions stated earlier where

$$\ddot{V}(y_1) = F_{\theta_1,\theta_2}(y_1)\log\left[1 + S_{\theta_1,\theta_2}(y_1)\right],$$

and

$$\ddot{A}(y_2) = F_{a_1,a_2}(y_2)\log\left[1 + S_{a_1,a_2}(y_2)\right].$$

In this case, one can also derive a closed form expression for the associated CDF of the BXgRR-FGM (Type-**III**) from

$$C_\rho(y_1, y_2) = F_{\theta_1,\theta_2}(y_1)F_{a_1,a_2}(y_2)\left[1 + \rho\ddot{V}(m)\,\ddot{A}(m)\right].$$

3.3. Via Clayton copula

The Clayton copula can be considered as

$$C(w_1, w_2) = [(1/w_1)^\rho + (1/w_2)^\rho - 1]^{-\rho^{-1}}|_{\rho\in(0,\infty)}.$$

Setting $w_1 = F_{\theta_1,\theta_2}(y_1) \in (0,1)$ and $w_2 = F_{a_1,a_2}(y_2) \in (0,1)$, the BXgRR type can be derived from $C(y_1, y_2) = C(F_{\theta_1,\theta_2}(y_1), F_{a_1,a_2}(y_2))$. Similarly, the MvXgRR (D-dimensional extension) from the above can be derived from

$$C(w_h) = \left(\sum_{h=1}^{D} w_h^{-\rho} + 1 - D\right)^{-\rho^{-1}}.$$

3.4. Via Renyi's entropy

Let $m \in (0,1) = F_{\theta_1,\theta_2}(y_1)$ and $w \in (0,1) = F_{a_1,a_2}(y_2)$. Then, the Renyi's entropy copula can be expressed as

$$C(y_1, y_2) = y_2 F_{\theta_1,\theta_2}(y_1) + y_1 F_{a_1,a_2}(y_2) - y_1 y_2.$$

Then, the associated BXgRR can be directly derived from

$$C(y_1, y_2) = C(F_{\theta_1,\theta_2}(y_1), F_{a_1,a_2}(y_2)).$$

3.5. Via Ali-Mikhail-Haq copula

Under the stronger Lipschitz condition, the Archimedean Ali-Mikhail-Haq copula can be expressed as

$$C_\rho(\mathrm{u}, \varpi) = \frac{\mathrm{u}\varpi}{1 - \rho\overline{\mathrm{u}}\overline{\varpi}}|_{\rho\in(-1,1)}.$$

Then, for $\overline{\mathrm{u}} = 1 - F_{\theta_1,\theta_2}(y_1)$, $\overline{\varpi} = 1 - F_{a_1,a_2}(y_2)$ we have the following Bv XgRR type

$$C_\rho(y_1, y_2) = \frac{F_{\theta_1,\theta_2}(y_1) F_{a_1,a_2}(y_2)}{1 - \rho S_{\theta_1,\theta_2}(y_1) S_{a_1,a_2}(y_2)}|_{\rho\in(-1,1)}.$$

4. Estimation

4.1. Maximum likelihood estimation (MLE)

Here, we consider the estimation of the unknown parameters of the new family from complete samples by maximum likelihood. Let $y_1, \cdots, y_n$ be a random sample from the XgRR model with a (2×1) parameter vector. The log-likelihood function for θ_1, θ_2 is given by

$$\ell_n(\theta_1, \theta_2) = n \log \theta_1 - n \log(1 + \theta_1) + n\, log\, 2 + 2n \log \theta_2 - 3\sum_{i=1}^{n} \log(y_i)$$
$$-\sum_{i=1}^{n}\left(\frac{\theta_2}{y_i}\right)^2 + (\theta_1 - 1)\sum_{i=1}^{n} \log[\hbar_{\theta_2}(y_i)] + \sum_{i=1}^{n} \log\left(\theta_1 + \frac{1}{2}\theta_1^2\{\log[\hbar_{\theta_2}(y_i)]\}^2\right).$$

where $\hbar_{\theta_2}(y_i) = 1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}$. The log-likelihood function in $\ell_n(\theta_1, \theta_2)$ can be maximized numerically by using R (optim), SAS (PROC NLMIXED) or Ox program (sub-routine MaxBFGS), among others. For interval estimation of the parameters, the elements of the 2×2 observed information matrix $\mathbf{J}(\theta_1, \theta_2)$ can be evaluated numerically.

4.2. Ordinary and weighted least-squares estimators

The theory of least square estimation and weighted least square estimation was proposed by Swain et al. (1988) to estimate the parameters of the Beta distribution. It is based on the minimization of the sum of the square of differences of theoretical cumulative distribution function and empirical distribution function. Suppose $F_{\theta_1,\theta_2}(y_{i\,:\,n})$ denotes the distribution function of the XgRR distribution and $y_1 < y_2 < \cdots < y_n$ be the n ordered random sample. The ordinary least square estimates (OLSEs) are obtained by minimizing

$$OLS(\theta_1, \theta_2) = \sum_{i=1}^{n} \left[F_{\theta_1,\theta_2}(y_{i\,:\,n}) - \frac{i}{n+1} \right]^2.$$

Now, using (1) we have

$$OLS(\theta_1, \theta_2) = \sum_{i=1}^{n} \left\{ \left[1 - \frac{\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1} \begin{pmatrix} 1 + \theta_1 - \theta_1 \log\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right] \\ +\frac{1}{2}\theta_1^2 \left\{ \log\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right] \right\}^2 \end{pmatrix} \right] - \frac{i}{n+1} \right\}^2.$$

Then, least square estimators (LSE) of the parameters are obtained by simultaneously solving the following non-linear equations:

$$\sum_{i=1}^{n} \left\{ \left[1 - \frac{\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1} \begin{pmatrix} 1 + \theta_1 - \theta_1 \log\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right] \\ +\frac{1}{2}\theta_1^2 \left\{ \log\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right] \right\}^2 \end{pmatrix} \right] - \frac{i}{n+1} \right\} \xi_{\theta_1|\theta_1,\theta_2}(y_i) = 0,$$

and

$$\sum_{i=1}^{n} \left\{ \left[1 - \frac{\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1} \begin{pmatrix} 1 + \theta_1 - \theta_1 \log\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right] \\ +\frac{1}{2}\theta_1^2 \left\{ \log\left[1 - e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right] \right\}^2 \end{pmatrix} \right] - \frac{i}{n+1} \right\} \xi_{\theta_2|\theta_1,\theta_2}(y_i) = 0,$$

where $\xi_{\theta_1|\theta_1,\theta_2}(y_i)$ and $\xi_{\theta_2|\theta_1,\theta_2}(y_i)$ are the values of the first derivatives with respect to parameters of XgRR distribution. The weighted least squares estimates (WLSE) are obtained by minimizing the given form of equation with respect to the parameters

$$WLS(\theta_1, \theta_2) = \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F_{\theta_1,\theta_2}(y_i) - \frac{i}{n+1} \right]^2.$$

The WLSE of the parameters are obtained by solving the following non-linear equations;

$$\sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i(n-i+1)}\left\{\left[1-\frac{\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1}\left(\begin{array}{c}1+\theta_1-\theta_1\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\\ +\frac{1}{2}\theta_1^2\left\{\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\right\}^2\end{array}\right)\right] - \frac{i}{n+1}\right\}\xi_{\theta_1|\theta_1,\theta_2}(y_i)=0,$$

and

$$\sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i(n-i+1)}\left\{\left[1-\frac{\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1}\left(\begin{array}{c}1+\theta_1-\theta_1\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\\ +\frac{1}{2}\theta_1^2\left\{\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\right\}^2\end{array}\right)\right] - \frac{i}{n+1}\right\}\xi_{\theta_2|\theta_1,\theta_2}(y_i)=0,$$

where $\xi_{\theta_1|\theta_1,\theta_2}(y_i)$ and $\xi_{\theta_2|\theta_1,\theta_2}(y_i)$ are the values of first derivatives of the CDF of XgRR distribution.

4.3. Method of Cramer-Von-Mises estimation

The Cramer-Von-Mises estimation (CVME) method of the parameters is based on the theory of minimum distance estimation. It was proposed by MacDonald (1971) and justified that the bias of the estimator is smaller than the other minimum distance estimators. Thus, The Crammer-Von-Mises estimates of the parameter θ_1 and θ_2 are obtained by minimizing the following expression with respect to the parameters θ_1 and θ_2 respectively.

$$CVM(\theta_1,\theta_2)=\frac{1}{12n}+\sum_{i=1}^{n}\left[F_{\theta_1,\theta_2}(y_i)-\frac{2i-1}{2n}\right]^2,$$

and

$$CVM(\theta_1,\theta_2)=\sum_{i=1}^{n}\left[1-\frac{\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1}\left(\begin{array}{c}1+\theta_1-\theta_1\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\\ +\frac{1}{2}\theta_1^2\left\{\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\right\}^2\end{array}\right)-\frac{2i-1}{2n}\right]^2.$$

The CVME of the parameters is obtained by solving the following non-linear equations

$$\sum_{i=1}^{n}\left[1-\frac{\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1}\begin{pmatrix}1+\theta_1-\theta_1\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\\ +\frac{1}{2}\theta_1^2\left\{\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\right\}^2\end{pmatrix}-\frac{2i-1}{2n}\right]\xi_{\theta_1|\theta_1,\theta_2}(y_i)=0,$$

and

$$\sum_{i=1}^{n}\left[1-\frac{\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]^{\theta_1}}{1+\theta_1}\begin{pmatrix}1+\theta_1-\theta_1\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\\ +\frac{1}{2}\theta_1^2\left\{\log\left[1-e^{-\left(\frac{\theta_2}{y_i}\right)^2}\right]\right\}^2\end{pmatrix}-\frac{2i-1}{2n}\right]\xi_{\theta_2|\theta_1,\theta_2}(y_i)=0,$$

where $\xi_{\theta_1|\theta_1,\theta_2}(y_i)$ and $\xi_{\theta_2|\theta_1,\theta_2}(y_i)$ are the values of the first derivatives of the CDF of XgRR distribution with respect to θ_1 and θ_2 respectively.

4.4. Bootstrapping method

Bootstrapping method is a powerful statistical technique. It is especially useful when the sample size that we are working with is small. Under the usual circumstances, sample sizes of less than 40 cannot be dealt with by assuming a normal or a t distributions. Bootstrap techniques work quite well with samples that have less than 40 elements. The reason for this is that bootstrapping involves resampling. These kinds of techniques assume nothing about the distribution of our data. Bootstrapping has become more popular as computing resources have become more readily available.

5. Numerical results for comparing estimation methods

In this Section, a Monte Carlo simulation study is conducted for comparing the performance of the different estimators of the unknown parameters of the XgRR distribution. The performance of the different estimators proposed in the previous Section is evaluated in terms of their mean squared errors (MSEs). All the computations in this section are done by Mathcad program Version 15.0. We generate 1000 samples of the XgRR distribution, where n = (20,50,100,200,500) and θ_1 and θ_2 are chosen as follows:

	I	II	III
θ_2	2.0	0.9	1.2
θ_1	1.5	0.3	0.6

The average values (AVs) of estimates and the corresponding MSEs of MLEs, LSEs, WLSEs, CVM, MPSD and Bootstrap method are obtained and reported in Tables 1, 2, 3, 4 and 5. We observe that all the estimates show the property of consistency, i.e. the MSEs decrease as the sample size increases.

Table 1. AVs and the corresponding MSEs (in parentheses) for n=20

Parameters	MLE	LS	WLS	CVM	Bootstrap
θ_2=2.0	2.06686	2.07031	2.09047	2.08087	3.14584
	(0.21967)	(0.20541)	(0.24017)	(0.23470)	(4.25306)
θ_1=1.5	1.52659	1.52061	1.51782	1.51839	1.19740
	(0.02907)	(0.03600)	(0.03636)	(0.04756)	(1.19181)
θ_2=0.9	0.99944	0.93704	0.93355	0.92649	0.77474
	(0.35047)	(0.04215)	(0.03902)	(0.03634)	(0.03767)
θ_1=0.3	0.31287	0.28547	0.23809	0.23838	0.43615
	(0.00451)	(1.15021)	(5.42295)	(5.30798)	(0.03477)
θ_2=1.2	1.24013	1.25164	1.24611	1.23804	1.03133
	(0.07944)	(0.08231)	(0.07525)	(0.07098)	(0.08032)
θ_1=0.6	0.62025	0.60920	0.61635	0.61905	0.79832
	(0.00956)	(0.04099)	(0.01483)	(0.01549)	(0.07196)

Table 2. AVs and the corresponding MSEs (in parentheses) for n=50.

Parameters	MLE	LS	WLS	CVM	Bootstrap
θ_2=2.0	2.03571	2.03548	2.04436	2.04309	1.89752
	(0.07429)	(0.08252)	(0.07604)	(0.07831)	(0.07750)
θ_1=1.5	1.50659	1.50738	1.50208	1.50268	1.56873
	(0.01035)	(0.01499)	(0.01267)	(0.01405)	(0.02128)
θ_2=0.9	0.97998	0.91650	0.91229	0.90955	1.01084
	(0.11707)	(0.01234)	(0.01270)	(0.01250)	(0.02619)
θ_1=0.3	0.29945	0.30253	0.30513	0.30620	0.27442
	(0.00206)	(0.00193)	(0.00195)	(0.00223)	(0.00152)
θ_2=1.2	1.22066	1.21680	1.20984	1.20756	1.19130
	(0.02568)	(0.02424)	(0.02090)	(0.02104)	(0.01494)
θ_1=0.6	0.60589	0.60694	0.60833	0.60960	0.61487
	(0.00319)	(0.00528)	(0.00421)	(0.00478)	(0.00325)

Table 3. AVs and the corresponding MSEs (in parentheses) for n=100

Parameters	MLE	LS	WLS	CVM	Bootstrap
θ_2=2.0	2.01577	2.01094	2.01555	2.01478	1.87189
	(0.03717)	(0.03506)	(0.03362)	(0.03489)	(0.03867)
θ_1=1.5	1.50394	1.50564	1.50324	1.50379	1.56382
	(0.00515)	(0.00693)	(0.00605)	(0.00673)	(0.00922)
θ_2=0.9	0.96579	0.90754	0.90982	0.90921	0.93491
	(0.05381)	(0.00584)	(0.00524)	(0.00541)	(0.01081)
θ_1=0.3	0.29496	0.30160	0.30052	0.30070	0.29517
	(0.00134)	(0.00096)	(0.00075)	(0.00092)	(0.00114)
θ_2=1.2	1.20965	1.21049	1.20931	1.20922	1.14055
	(0.01303)	(0.01178)	(0.01082)	(0.01108)	(0.01160)
θ_1=0.6	0.60315	0.60230	0.60269	0.60268	0.57792
	(0.00155)	(0.00245)	(0.00207)	(0.00238)	(0.00253)

Table 4. AVs and the corresponding MSEs (in parentheses) for n=200

Parameters	MLE	LS	WLS	CVM	Bootstrap
θ_2=2.0	2.00332	2.01138	2.00587	2.00564	1.84919
	(0.01841)	(0.01710)	(0.01740)	(0.01811)	(0.03575)
θ_1=1.5	1.50390	1.50005	1.50278	1.50295	1.57880
	(0.00263)	(0.00329)	(0.00313)	(0.00350)	(0.00910)
θ_2=0.9	0.95153	0.90173	0.90503	0.90449	0.98419
	(0.02540)	(0.00267)	(0.00264)	(0.00270)	(0.01048)
θ_1=0.3	0.29346	0.30154	0.30022	0.30040	0.27433
	(0.00087)	(0.00046)	(0.00038)	(0.00046)	(0.00091)
θ_2=1.2	1.20178	1.19981	1.20636	1.20563	1.25721
	(0.00634)	(0.00491)	(0.00513)	(0.00527)	(0.00807)
θ_1=0.6	0.60278	0.60331	0.60044	0.60074	0.57895
	(0.00079)	(0.00111)	(0.00099)	(0.00115)	(0.00119)

Table 5. AVs and the corresponding MSEs (in parentheses) for n=500.

Parameters	MLE	LS	WLS	CVM	Bootstrap
θ_2=2.0	1.99647	1.99838	1.99702	1.99672	1.99960
	(0.00668)	(0.00654)	(0.00629)	(0.00666)	(0.00620)
θ_1=1.5	1.50344	1.50277	1.50334	1.50356	1.50239
	(0.00100)	(0.00133)	(0.00121)	(0.00137)	(0.00136)
θ_2=0.9	0.94897	0.89934	0.89870	0.89860	0.90144
	(0.01044)	(0.00104)	(0.00629)	(0.00106)	(0.00105)
θ_1=0.3	0.28999	0.30117	0.30139	0.30151	0.30039
	(0.00049)	(0.00019)	(0.00016)	(0.00019)	(0.00018)
θ_2=1.2	1.19801	1.20037	1.20245	1.20247	1.22109
	(0.00229)	(0.00198)	(0.00196)	(0.00202)	(0.00292)
θ_1=0.6	0.60204	0.60112	0.60018	0.60015	0.59211
	(0.00030)	(0.00044)	(0.00038)	(0.00044)	(0.00049)

6. Modified Right-Censored Test for Validation

6.1. The N.R.R statistic test

Many goodness-of-fit tests are used to indicate whether or not it is reasonable to assume that a random sample comes from a specific distribution. For this purpose, researchers proposed many different goodness-of-fit tests. For the complete data, Nikulin (1973a, 1973b and 1973c) and Rao and Robson (1974) separately proposed a statistic known today as the N.R.R statistic. This statistical test is a natural modification of the Pearson statistic. To test the hypothesis H_0 we have

$$H_0: P\{T_i \leq t\} = F(t,\zeta)|_{(t\in R,\ \zeta=(\zeta_1,\zeta_2,\cdots,\zeta_s)^T)},$$

where ζ represents the vector of unknown parameters. Nikulin (1973a, 1973b and 1973c) and Rao and Robson (1974) proposed the N.R.R statistic defined as follows: Observations $T_1, T_2, \cdots, T_n$ are grouped in r subintervals and $\nu_j = (\nu_1, \nu_2, \cdots, \nu_r)^T$ is the vector of frequencies, where ν_j is frequency of ith group and $\sum_{j=1}^{r} \nu_j = n.$ The tests are based on the following Pearson's statistic

$$Y^2\left(\hat{\underline{\zeta}}_n\right) = \chi_n^2\left(\hat{\underline{\zeta}}_n\right) + n^{-1}\ell^T\left(\hat{\underline{\zeta}}_n\right)\left(\mathbf{I}\left(\hat{\underline{\zeta}}_n\right) - \mathbf{J}\left(\hat{\underline{\zeta}}_n\right)\right)^{-1}\ell\left(\hat{\underline{\zeta}}_n\right),$$

where

$$\chi_n^2(\zeta) = \left(\frac{\nu_1 - np_1(\zeta)}{\sqrt{np_1(\zeta)}}, \frac{\nu_2 - np_2(\zeta)}{\sqrt{np_2(\zeta)}}, \cdots, \frac{\nu_r - np_r(\zeta)}{\sqrt{np_r(\zeta)}}\right)^T,$$

and $p(\zeta)$ is the vector of probabilities and ζ is the vector of parameters which can be known (simple hypothesis) or unknown (composite hypothesis). The Y^2 statistic follows a chi-square distribution with $(r-1)$ degrees of freedom (for more details see Nikulin (1973a, 1973b and 1973c)).

6.2. Application to right-censored real data

To test the null hypothesis H_0, we use the N.R.R statistic. We compute the maximum likelihood estimators

$$\widehat{\theta_1} = 0.95473 \text{ and } \widehat{\theta_2} = 1.24885.$$

We then deduce the value of $Y^2 = 11.05847$. The critical value is

$$\chi_{0.05}^2(6-1) = 11.0705.$$

Then, the N.R.R Y^2 statistic value is less than the critical value, we say that taxes revenue data can be fitted by the XgRR model. The modified chi-squared test for composite hypothesis for complete samples was first considered by Nikulin (1973a, 1973b and 1973c), Rao and Robson (1974). Several goodness-of-fit tests have been suggested by the statisticians for censored data. Bagdonavicius and Nikulin $(2011a, b)$ proposed a modification of the N.R.R statistic that takes into account random right censorship and based on the maximum likelihood estimators on the initial data, also follows a limiting Chi-square distribution. In this Section we develop the approach proposed by Bagdonavicius and Nikulin $(2011a, b)$ to confirm the adequacy of XgRR model when the parameters are unknown and data are censored. Let us consider the composite hypothesis

$$H_0 \;:\; F(t) \in F_0 = F_0(t, \zeta)|_{(t\in R^1,\ \zeta\in\Psi\subset R^s)},$$

where ζ is an unknown m-dimensional parameter and F_0 is a differentiable and completely specified cdf with the support $(0, \infty)$. Let us consider a finite time interval, say, $[0, \tau]$, where τ is the maximum time of the study, and divide it into $k > s$ smaller intervals $I_j = (a_{j-1}, a_j]$, where

$$0 < a_0 < a_1 \ldots < a_{k-1} < a_k < +\infty.$$

In this case the estimated $\hat{a}_{\boldsymbol{k}}$ is given by

$$\hat{a}_{\boldsymbol{k}} = \Lambda^{-1}\left\{\frac{1}{n-i+1}\left[E_{\boldsymbol{j}} - \sum_{l=1}^{i-1}\Lambda\left(\mathrm{T}_{(l)},\underline{\hat{\zeta}}\right)\right],\underline{\hat{\zeta}}\right\}, \hat{a}_{\boldsymbol{k}} = \mathrm{t}_{(n)} | \boldsymbol{j} = 1,2,\dots,k$$

where $\underline{\hat{\zeta}}$ is the maximum likelihood estimator of the parameter ζ, Λ^{-1} is the inverse of cumulative hazard function Λ , $T_{(i)}$ is the i^{th} element in the ordered statistics $(T_{(1)},\dots,T_{(n)})$ and

$$E_{\boldsymbol{j}} = (n+1-i)\Lambda\left(\hat{a}_{(j)},\underline{\hat{\zeta}}\right) + \sum_{\boldsymbol{l}=1}^{i-1}\Lambda\left(\mathrm{T}_{(l)},\underline{\hat{\zeta}}\right),$$

and a_j are random data functions such as the k intervals have equal expected numbers of failures e_j . Usually in real application we fix k. The test statistic for H_0 is given in Goual et al. (2020) and Goual and Yousof (2019). The survival times in days are for the $n = 51$ patients. The data are: 7, 34, 42, 63, 64, 74*, 83, 84, 91, 108, 112, 129, 133, 133, 139, 140, 140, 146, 149, 154, 157, 160, 160, 165, 173, 176, 185*, 218, 225, 241, 248, 273, 277, 279*, 297, 319*, 405, 417, 420, 440, 523*, 523, 583, 594, 1101, 1116*, 1146, 1226*, 1349*, 1412*, 1417. (* censored). We suppose that these data are distributed according to the XgRR distribution, we transform the survival times in months (1 month = 30.438 days), so the maximum likelihood estimates of the parameter vector ζ are

$$\zeta = (5.00248, 1.378452)^T$$

We choose $r = 7$ as the number of classes. The elements of the test statistic Y_n^2 is presented as follows, we find $Y_n^2 = 14.000154$ and the critical value $\chi^2_{0.05}(7) = 14.00924$. Comparing the critical value and the statistic test Y_n^2, we can say that Arm-A head and neck cancer data can be adjusted by the XgRR model.

7. Concluding remarks

In this article, a new reciprocal Rayleigh extension called the Xgamma reciprocal Rayleigh model is defined and studied. Relevant statistical properties such as raw moments, incomplete moments and moment generating function are derived. After a quick study for their properties, different non-Bayesian estimation methods under uncensored schemes are considered and described such as the maximum likelihood estimation method, ordinary least square estimation method, weighted least square estimation method, Cramér–von-Mises estimation method and Bootstrapping method. The performances of the proposed estimation methods are investigated through a simulation study. Many bivariate and multivariate type models have been also derived based on Farlie Gumbel Morgenstern copula, Clayton copula, Renyi's entropy copula and Ali–Mikhail–Haq copula. A modified right-censored test for validation is applied to a right-censored real data set.

As a potential future work, we can use and apply the well-known Bagdonavičius-Nikulin goodness-of-fit test and the modified version of the Bagdonavičius-Nikulin goodness-of-fit test to our new XgRR model and many other useful lifetime models (see Goual et al. (2019), Ibrahim et al. (2020), Yadav et al. (2020) and Mansour et al. (2020a-f) for more details). The reciprocal Rayleigh distribution can be extended using some new G families such as presented by Alizadeh et al. (2020) and El-Morshedy et al. (2021). Some useful real-life data sets can be cited from Elgohari and Yousof (2020a and 2020b).

References

Alizadeh, M., Ghosh, I., Yousof, H. M., Rasekhi, M. and Hamedani G. G., (2017). The generalized odd generalized exponential family of distributions: properties, characterizations and applications, J. Data Sci., Vol. 15, pp. 443–466.

Alizadeh, M., Jamal, F., Yousof, H. M., Khanahmadi, M. and Hamedani, G. G., (2020). Flexible Weibull generated family of distributions: characterizations, mathematical properties and applications. University Politehnica of Bucharest Scientific Bulletin-Series A-Applied Mathematics and Physics, Vol. 82, pp. 145–150.

Alizadeh, M., Rasekhi, M., Yousof, H. M. and Hamedani G. G., (2018). The transmuted Weibull G family of distributions. Hacettepe Journal of Mathematics and Statistics, Vol. 47, pp. 1–20.

Alizadeh, M., Yousof, H. M. Rasekhi, M. and Altun, E., (2019). The odd log-logistic Poisson-G Family of distributions, Journal of Mathematical Extensions, Vol. 12, pp. 81–104.

Aryal, G. R., Yousof, H. M., (2017). The exponentiated generalized-G Poisson family of distributions. Economic Quality Control, 32, pp. 1–17.

Bagdonavicius, V., Nikulin, M., (2011a). Chi-squared Goodness-of-fit Test for Right Censored Data. International Journal of Applied Mathematics and Statistics, Vol. 24, pp. 30–50.

Bagdonavicius, V., Nikulin, M., (2011b). Chi-squared tests for general composite hypotheses from censored samples. Comptes Rendus de l'académie des Sciences de Paris, Mathématiques, Vol. 349, pp. 219–223.

Brito, E., Cordeiro, G. M., Yousof, H. M., Alizadeh, M. and Silva, G. O., (2017). Topp-Leone Odd Log-Logistic Family of Distributions, Journal of Statistical Computation and Simulation, Vol. 87, pp. 3040–3058.

Cordeiro, G. M., Altun, E., Korkmaz, M. C., Pescim, R. R., Afify, A. Z. and Yousof, H. M., (2020). The xgamma Family: Censored Regression Modelling and Applications. Revista Colombiana de Estadística, Vol. 18, pp. 593–612.

Cordeiro, G. M., Yousof, H. M., Ramires, T. G. and Ortega, E. M. M., (2018). The Burr XII system of densities: properties, regression model and applications. Journal of Statistical Computation and Simulation, Vol. 88, pp. 432–456.

Chakraborty, S., Handique, L., Altun, E. and Yousof, H. M., (2018). A new statistical model for extreme values: mathematical properties and applications. International Journal of Open Problems in Computer Science and Mathematics, Vol. 12, pp. 1–18.

Elgohari, H., Yousof, H. M., (2020a). A Generalization of Lomax Distribution with Properties, Copula and Real Data Applications. *Pakistan Journal of Statistics and Operation Research*, Vol. 16, pp. 697–711.

Elgohari, H., Yousof, H. M., (2020b). New Extension of Weibull Distribution: Copula, Mathematical Properties and Data Modeling. *Statistics, Optimization & Information Computing*, Vol. 8, pp. 972–993.

El-Morshedy, M., Alshammari, F. S., Hamed, Y. S., Eliwa, M. S., Yousof, H. M., (2021). A new family of continuous probability distributions. Entropy, Vol. 23, 194. https://doi.org/10.3390/e23020194

Fréchet, M., (1927). Sur la loi de probabilité de lécart maximum. Ann. Soc. Pol. Math., Vol. 6, pp. 93–116.

Gradshteyn, I. S., Ryzhik, I. M., (2002). Table of integrals, series, and products. San Diego, CA: Academic Press.

Goual, H., Yousof, H. M. and Ali, M. M., (2019). Validation of the odd Lindley exponentiated exponential by a modified goodness of fit test with applications to censored and complete data. Pakistan Journal of Statistics and Operation Research, Vol. 15, pp. 745–771.

Goual, H., Yousof, H. M., (2019). Validation of Burr XII inverse Rayleigh model via a modified chi-squared goodness-of-fit test. Journal of Applied Statistics, Vol. 47, pp. 1–32.

Goual, H., Yousof, H. M., and Ali, M. M., (2020). Lomax inverse Weibull model: properties, applications, and a modified Chi-squared goodness-of-fit test for validation. Journal of Nonlinear Sciences & Applications (JNSA), Vol. 13, pp. 330–353.

Gusmao, F. R. S., Ortega, E. M. M. and Cordeiro, G. M., (2011). The generalized inverse Weibull distribution. Statistical Papers, Vol. 52, pp. 591–619.

Haq, M. A., Yousof, H. M. and Hashmi, S., (2017). A New Five-Parameter Fréchet Model for Extreme Values. Pakistan Journal of Statistics and Operation Research, Vol. 13(3), pp. 617–632.

Hamedani G. G., Altun, E, Korkmaz, M. C., Yousof, H. M. and Butt, N. S., (2018). A new extended G family of continuous distributions with mathematical properties, characterizations and regression modeling. Pakistan Journal of Statistics and Operation Research, Vol. 14, pp. 737–758.

Hamedani G. G., Rasekhi, M., Najibi, S. M., Yousof, H. M. and Alizadeh, M., (2019). Type II general exponential class of distributions. Pakistan Journal of Statistics and Operation Research , Vol. 15, pp. 503–523.

Hamedani G. G., Yousof, H. M., Rasekhi, M., Alizadeh, M. and Najibi, S. M., (2017). Type I general exponential class of distributions. Pakistan Journal of Statistics and Operation Research, Vol. 14, pp. 39–55.

Ibrahim, M., (2019). A new extended Fréchet distribution: properties and estimation. Pakistan Journal of Statistics and Operation Research, Vol. 15, pp. 773–796.

Ibrahim, M., (2020a). The compound Poisson Rayleigh Burr XII distribution: properties and applications. Journal of Applied Probability and Statistics, Vol. 15, pp. 73–97.

Ibrahim, M., (2020b). The generalized odd Log-logistic Nadarajah Haghighi distribution: statistical properties and different methods of estimation. Journal of Applied Probability and Statistics, Vol. 15, pp. 61–84.

Ibrahim, M., Altun, E. and Yousof, H. M., (2020). A new distribution for modeling lifetime data with different methods of estimation and censored regression modeling. Statistics, Optimization & Information Computing, Vol. 8, pp. 610–630.

Ibrahim, M., Yadav, A. S., Yousof, H. M., Goual, H. and Hamedani, G. G., (2019). A new extension of Lindley distribution: modified validation test, characterizations and different methods of estimation, Communications for Statistical Applications and Methods, Vol. 26, pp. 473–495.

Jahanshahi, S. M. A., Yousof, H. M. and Sharma, V. K., (2019). The Burr X Fréchet Model for Extreme Values: Mathematical Properties, Classical Inference and Bayesian Analysis. Pakistan Journal of Statistics and Operation Research, Vol. 15, pp. 797–818.

Johnson, N. L., Kemp, A. W. and Kotz, S., (2005). Univariate discrete distributions, 3rd edn. Wiley, Hoboken.

Korkmaz, M. C., Altun, E., Yousof, H. M. and HAMEDANI G. G., (2019). The odd power Lindley generator of probability distributions: properties, characterizations and regression modeling, International Journal of Statistics and Probability, Vol. 8, pp. 70–89.

Korkmaz, M. C., Yousof, H. M. and Ali, M. M., (2017). Some theoretical and computational aspects of the odd Lindley Fréchet distribution, Journal of Statisticians: Statistics and Actuarial Sciences, Vol. 2, pp. 129–140.

Korkmaz, M. C., Yousof, H. M. and Hamedani G. G., (2018). The exponential Lindley odd log-logistic G family: properties, characterizations and applications. Journal of Statistical Theory and Applications, Vol. 17, pp. 554–571.

Krishna, E., Jose, K. K., Alice, T. and Risti, M. M., (2013). The Marshall-Olkin Fréchet distribution. Communications in Statistics-Theory and Methods, Vol. 42, pp. 4091–4107

Mahmoud, M. R., Mandouh, R. M., (2013). On the transmuted Fréchet distribution. Journal of Applied Sciences Research, Vol. 9, pp. 5553–5561.

Mansour, M. M., Ibrahim, M., Aidi, K., Shafique Butt, N., Ali, M. M., Yousof, H. M. and Hamed, M. S., (2020a). A New Log-Logistic Lifetime Model with Mathematical Properties, Copula, Modified Goodness-of-Fit Test for Validation and Real Data Modeling. Mathematics, Vol. 8, p. 1508.

Mansour, M. M., Butt, N. S., Ansari, S. I., Yousof, H. M., Ali, M. M. and Ibrahim, M., (2020b). A new exponentiated Weibull distribution's extension: copula, mathematical properties and applications. Contributions to Mathematics, Vol. 1 (2020), pp. 57–66, doi: 10.47443/cm.2020.0018

Mansour, M., Korkmaz, M. Ç., Ali, M. M., Yousof, H. M., Ansari, S. I. and Ibrahim, M., (2020c). A generalization of the exponentiated Weibull model with properties, Copula and application. Eurasian Bulletin of Mathematics, Vol. 3, pp. 84–102.

Mansour, M., Rasekhi, M., Ibrahim, M., Aidi, K., Yousof, H. M. and Elrazik, E. A., (2020d). A New Parametric Life Distribution with Modified Bagdonavičius–Nikulin Goodness-of-Fit Test for Censored Validation, Properties, Applications, and Different Estimation Methods. Entropy, Vol. 22, p. 592.

Mansour, M., Yousof, H. M., Shehata, W. A. and Ibrahim, M., (2020e). A new two parameter Burr XII distribution: properties, copula, different estimation methods and modeling acute bone cancer data. Journal of Nonlinear Science and Applications, Vol.13, pp. 223–238.

Mansour, M. M., Butt, N. S., Yousof, H. M., Ansari, S. I. and Ibrahim, M., (2020f). A Generalization of Reciprocal Exponential Model: Clayton Copula, Statistical Properties and Modeling Skewed and Symmetric Real Data Sets. Pakistan Journal of Statistics and Operation Research, Vol. 16, pp. 373–386.

Merovci, F., Alizadeh, M., Yousof, H. M. and Hamedani G. G., (2017). The exponentiated transmuted-G family of distributions: theory and applications, Communications in Statistics-Theory and Methods, Vol. 46, pp. 10800–10822.

Merovci, F., Yousof, H. M. and Hamedani, G. G., (2020). The Poisson Topp Leone Generator of Distributions for Lifetime Data: Theory, Characterizations and Applications. Pakistan Journal of Statistics and Operation Research, Vol. 16, pp. 343–355.

Nikulin, M. S., (1973a). Chi-square Test for Continuous Distributions with Shift and Scale Parameters. teor. Veroyatn. Primen, Vol. 18**,** pp. 559–568.

Nikulin, M. S., (1973b). Chi-squared test for continuous distributions with shift and scal parameters. Theory of Probability and its Applications, Vol. 18, pp. 559–568.

Nikulin, M. S., (1973c). On a Chi-squared test for continuous distributions. Theory of Probability and its Aplications, Vol. 19, pp. 638–639.

Salah, M., El-Morshedy, M., Eliwa, M. S. and Yousof, H. M., (2020). Expanded Fréchet Model: Mathematical Properties, Copula, Different Estimation Methods, Applications and Validation Testing. Mathematics, Vol. 8, pp.19–49.

Swain, J. Venkatraman, S. and Wilson, J., (1988). Least squares estimation of distribution function in Johnsons translation system, J. Stat. Comput. Simul., Vol. 29, pp. 271–297.

Nascimento, A. D. C., Silva, K. F., Cordeiro, G. M., Alizadeh, M. and Yousof, H. M., (2019). The odd Nadarajah-Haghighi family of distributions: properties and applications. Studia Scientiarum Mathematicarum Hungarica, Vol. 56, pp. 1–26.

Rao, K. C., Robson, D. S., (1974). A Chi-square statistic for goodness-of-fit tests within the Exponential family. Communication in Statistics, Vol. 3, pp.1139–1153.

Yadav, A. S., Goual, H., Alotaibi, R. M. Rezk, H., Ali, M. M. and Yousof, H. M., (2020). Validation of the Topp-Leone-Lomax model via a modified Nikulin-Rao-Robson goodness-of-fit test with different methods of estimation. Symmetry, Vol. 12, pp. 1–26. DOI: 10.3390/sym12010057

Yousof, H. M., Afify, A. Z., Alizadeh, M., Butt, N. S., Hamedani, G. G. and Ali, M. M., (2015). The transmuted exponentiated generalized-G family of distributions. Pakistan Journal of Statistics and Operation Research, Vol. 11, pp. 441–464.

Yousof, H. M., Afify, A. Z., Ebraheim, A. N., Hamedani, G. G. and Butt, N. S., (2016). On six-parameter Fréchet distribution: properties and applications, Pakistan Journal of Statistics and Operation Research, Vol. 12, pp. 281–299.

Yousof, H. M., Alizadeh, M., Jahanshahiand, S. M. A., Ramires, T. G., Ghosh, I. and Hamedani G. G., (2017). The transmuted Topp-Leone G family of distributions: theory, characterizations and applications, Journal of Data Science. Vol. 15, pp. 723–740.

Yousof, H. M., Afify, A. Z., Alizadeh, M., Nadarajah, S., Aryal, G. R. and Hamedani, G. G., (2018a). The Marshall-Olkin generalized-G family of distributions with Applications, STATISTICA, Vol. 78, pp. 273–295.

Yousof, H. M., Altun, E. and Hamedani, G. G., (2018b). A new extension of Frechet distribution with regression models, residual analysis and characterizations. Journal of Data Science, Vol. 16, pp. 743–770.

Yousof, H. M., Afify, A. Z., Alizadeh, M., Hamedani G. G., Jahanshahi, S. M. A. and Ghosh, I., (2018c). The generalized transmuted Poisson-G family of Distributions. Pakistan Journal of Statistics and Operation Research, Vol. 14, pp. 759–779.

Yousof, H. M., Altun, E., Ramires, T. G., Alizadeh, M. and Rasekhi, M., (2018c). A new family of distributions with properties, regression models and applications, Journal of Statistics and Management Systems, Vol. 21, pp. 163–188.

Yousof, H. M., Butt, N. S., Alotaibi, R., Rezk, H. R., Alomani, G. A. and Ibrahim, M., (2019). A new compound Fréchet distribution for modeling breaking stress and strengths data. Pakistan Journal of Statistics and Operation Research, Vol. 15, pp. 1017–1035.

Yousof, H. M., Hamedani, G. G. and Ibrahim, M., (2020). The Two-parameter Xgamma Fréchet Distribution: Characterizations, Copulas, Mathematical Properties and Different Classical Estimation Methods. *Contributions to Mathematics*, Vol. 2 (2020), pp. 32–41.

Yousof, H. M., Rasekhi, M., Altun, E. and Alizadeh, M., (2018d). The extended odd Frechet family of distributions: properties, applications and regression modeling. International Journal of Applied Mathematics and Statistics, Vol. 30, pp. 1–30.