Research Article GU J Sci 38(2): 1021-1040 (2025) DOI: 10.35378/gujs.1549073

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Unit Gamma-Lindley Distribution: Properties, Estimation, Regression Analysis, and Practical Applications

Kadir KARAKAYA * , Sule SAGLAM

Selçuk University, Faculty of Science, Department of Statistics, 42130, Konya, Türkiye

Highlights

- Performance analysis using the Monte Carlo simulation.
- A novel unit distribution is introduced.
- A new regression analysis is proposed as an alternative to the Beta regression model.

Article Info

Received: 12 Sep 2024
Accepted: 13 Feb 2025

Keywords

Beta regression analysis,
Cost-effectiveness data,
Maximum likelihood estimation,
OECD,
Unit distribution

Abstract

This study proposes the unit Gamma-Lindley distribution, a novel bounded statistical model that extends the flexibility of existing distributions for modeling data on the (0,1) interval. The proposed distribution is characterized, by closed-form expressions derived for its cumulative distribution, probability density, and hazard rate functions. Some statistical properties, including moments, order statistics, Bonferroni, Lorenz curves, entropy, etc. are examined. To estimate the unknown model parameters, several estimation methods are introduced and their performance is assessed through a Monte Carlo simulation experiment based on bias and mean square error criteria. A real data application focusing on firm management cost-effectiveness highlights the practical utility of the model, demonstrating its superior fit compared to current distributions, such as beta and Kumaraswamy. Furthermore, a novel regression model is developed based on the proposed distribution, with parameter estimation performed using the maximum likelihood method. The new regression model provides an alternative for analyzing bounded response variables. The findings contribute to the statistical literature by offering a flexible and comprehensive modeling framework for bounded data, with theoretical advancements and practical applicability.

1. INTRODUCTION

Statistical distributions play a crucial role in modeling data across many fields, including actuarial science, medicine, chemistry, and engineering sciences. Accurate modeling of real-world data enables us to acquire comprehensive insights into the characteristics and dynamics of populations related to the relevant phenomena. Therefore, an increasing number of alternative distributions provides a more flexible choice in modeling. Despite much data in the (0,1) interval, such as mortality rates, recovery rates, and education measure data, the number of bounded distributions is limited. The popular bounded distributions are Kumaraswamy and beta. The Kumaraswamy and beta distributions may be insufficient for accurately modeling proportional data. This potential insufficientness has prompted statisticians to derive new bounded distributions. Hence, existing lifetime distributions are transformed to the interval (0, 1), i.e., united. The advantage of these transformed unit distributions is that they enhance the flexibility of the baseline distribution within the unit range without asking for the addition of new parameters. The study aims to propose a new unit alternative distribution for modeling rates and proportions as an alternative to current ones, such as beta, Kumaraswamy, and unit Weibull distribution [1]. Several unit distributions investigated in recent years include those proposed by [2-12]. The beta regression model is a popular regression analysis suitable for scenarios where the response variable falls within the range of 0 to 1. In recent years, several alternative regression models are introduced as alternatives to the beta regression model. Examples of these include [13-15]. In this paper, a new regression model is built on the proposed distribution. The probability density function (pdf) and cumulative distribution function (cdf) of the new

*Corresponding author, e-mail: kkarakaya@selcuk.edu.tr

regression model are derived, and model parameters are estimated via the maximum likelihood method. The superior performance of the new regression model over existing models is illustrated by using educational attainment data from Organisation for Economic Co-operation and Development (OECD) countries. In this study, the Gamma-Lindley (GL) distribution, as introduced by [16], is united. This study makes many original contributions to the statistical modeling of unit data. New closed-form expressions are obtained for distribution functions, and several estimators are examined and evaluated through simulations, and a new regression model is introduced based on the new distribution. The results demonstrate the advantages of the proposed methods over existing approaches, such as beta and Kumaraswamy models, in real-world applications. The rest of the paper is organized as follows: section 2 introduces the new model, and some mathematical properties are studied. The various estimators are studied to estimate the two unknown parameters of the new distribution in section 3. In section 4, comprehensive Monte Carlo simulations are carried out to compare the performance of the estimators. A new regression model is introduced in section 5. In section 6, two practical examples related to cost-effectiveness in firms' risk management and educational attainment for OECD countries are analyzed to demonstrate the usability of the proposed distribution and regression model, respectively. The paper concludes with the concluding remarks in section 7.

2. THE BOUNDED VERSION OF GAMMA-LINDLEY DISTRIBUTION

Let X be a random variable (rv) following GL distribution introduced by [16], with cdf and pdf presented

$$F_{GL}\left(x;\alpha,\beta\right)=\frac{x^{\alpha}}{\left(\beta+x\right)^{\alpha}}+\frac{\beta\alpha x^{\alpha}}{\left(1+\beta\right)\left(\beta+x\right)^{1+\alpha}},$$

and

$$f_{GL}\left(x;\alpha,\beta\right)=\frac{\alpha\beta^{2}}{1+\beta}\frac{\left(1+\alpha+\beta+x\right)x^{\alpha-1}}{\left(\beta+x\right)^{2+\alpha}},x>0,\alpha>0,\beta>0, \tag{1}$$

respectively, where α,β is the scale parameters. Let the rv X has pdf in Equation (1), then the cdf and pdf of the rv $Y=\frac{X}{1-X}$ are acquired

$$F_{UGL}\left(y;\alpha,\beta\right)=\frac{\left(\frac{y}{1-y}\right)^{\alpha}}{\left(\beta+\frac{y}{1-y}\right)^{\alpha}}+\frac{\beta\alpha\left(\frac{y}{1-y}\right)^{\alpha}}{\left(1+\beta\right)\left(\beta+\frac{y}{1-y}\right)^{1+\alpha}}, \tag{2}$$

and

$$f_{UGL}\left(y;\alpha,\beta\right)=\frac{\alpha\beta^{2}\left(1+\left(1-y\right)\left(\alpha+\beta\right)\right)y^{\alpha-1}\left(\beta-\beta y+y\right)^{-\alpha}}{\left(1+\beta\right)\left(\beta y-y-\beta\right)^{2}}, \tag{3}$$

respectively, where $y\in\left(0,1\right)$ and $\alpha>0,\beta>0$ is the model parameters of the proposed distribution. The distribution with pdf in Equation (3), is called the unit Gamma-Lindley (UGL), and it is presented by $UGL\left(\alpha,\beta\right)$. The hazard rate function (hrf) of the UGL distribution is given as

$$h_{UGL} = \frac{\alpha\beta^2\left(1+(1-y)(\alpha+\beta)\right)y^{\alpha-1}(\beta-\beta y+y)^{-\alpha}}{(1+\beta)(\beta y-y-\beta)^2\left(1-\left(\frac{\left(\frac{y}{1-y}\right)^{\alpha}}{\left(\beta+\frac{y}{1-y}\right)^{\alpha}}+\frac{\beta\alpha\left(\frac{y}{1-y}\right)^{\alpha}}{(1+\beta)\left(\beta+\frac{y}{1-y}\right)^{1+\alpha}}\right)\right)}.$$

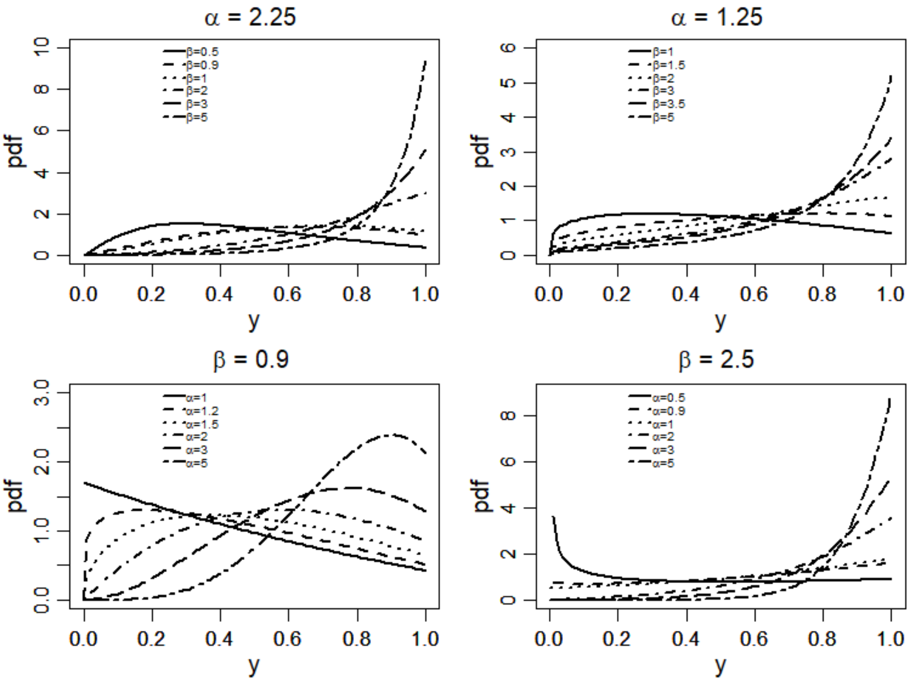

Figure 1. *The pdf of UGL distribution for some parameter choices*

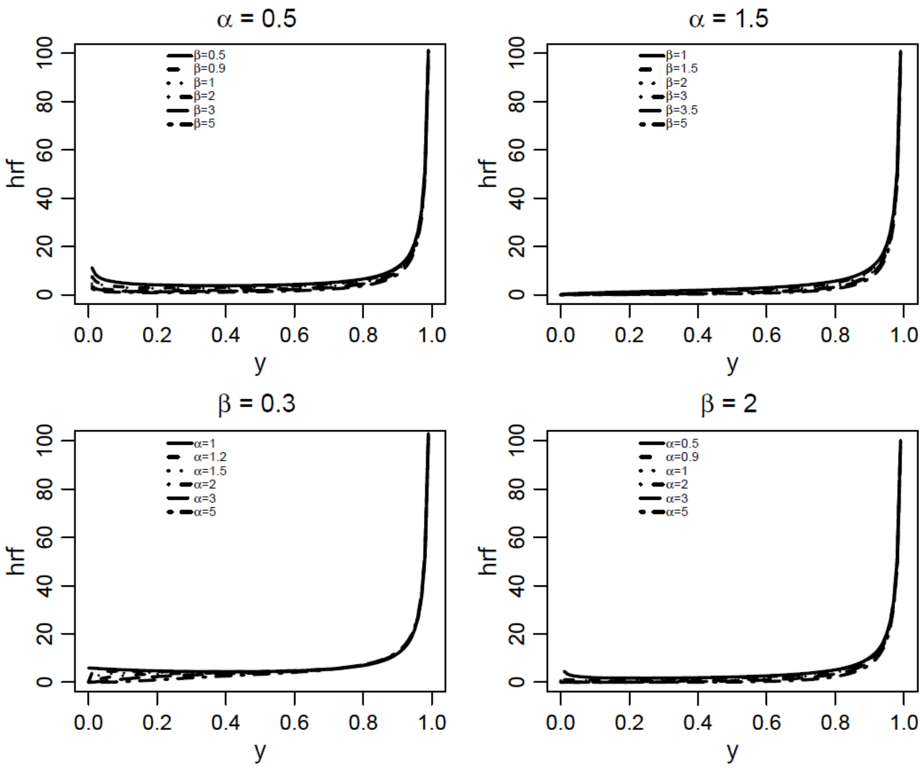

Figure 2. *The hrf of UGL distribution for some parameter choices*

The plots of the pdf and hrf are given in Figures 1 and 2. Figure 1 illustrates that the pdf of the UGL distribution exhibits patterns of increase, decrease, and unimodal shape. From Figure 2, one can see that the hrf of UGL has increased and the bath-tube shape.

2.1. Moments

Let Y be a $UGL(\alpha,\beta)$ rv with pdf in Equation (3). The $r-th$ moment of Y is obtained as

$$\begin{aligned}
\mu_r &= \int_0^1 y^r f(y)\,dy \\
&= \frac{1}{2}\int_{-1}^{1}\left(\frac{z+1}{2}\right)^r f\left(\frac{z+1}{2}\right)dz \\
&= \frac{\alpha\beta^2}{2(1+\beta)}\int_{-1}^{1}\left(\frac{z+1}{2}\right)^{r+\alpha-1}\left(\frac{\left(1+\left(1-\left(\frac{z+1}{2}\right)\right)(\alpha+\beta)\right)\left(\beta-\beta\left(\frac{z+1}{2}\right)+\left(\frac{z+1}{2}\right)\right)^{-\alpha}}{\left(\beta\left(\frac{z+1}{2}\right)-\left(\frac{z+1}{2}\right)-\beta\right)^2}\right)dz \\
&\cong \frac{\alpha\beta^2}{2(1+\beta)}\sum_{\ell=1}^{N}\varpi_\ell\left(\frac{z_\ell+1}{2}\right)^{r+\alpha-1}\left(\frac{\left(1+\left(1-\left(\frac{z_\ell+1}{2}\right)\right)(\alpha+\beta)\right)\left(\beta-\beta\left(\frac{z_\ell+1}{2}\right)+\left(\frac{z_\ell+1}{2}\right)\right)^{-\alpha}}{\left(\beta\left(\frac{z_\ell+1}{2}\right)-\left(\frac{z_\ell+1}{2}\right)-\beta\right)^2}\right),
\end{aligned}$$

where z_ℓ and ϖ_ℓ are the zeros and the corresponding Christoffel numbers of the Legendre-Gauss quadrature formula on the $(-1,1)$, respectively, see [17]. The ϖ_ℓ is given by

$$\varpi_\ell = \frac{2}{(1-z_\ell)^2\left[L'_{N+1}(z_\ell)\right]^2}, \tag{4}$$

where

$$L'_{N+1}(z_\ell) = \frac{dL_{N+1}(z)}{dz} \tag{5}$$

at $z = z_\ell$ and $L_{N+1}(\cdot)$ is the Legendre polynomial of degree N.

2.2. Bonferroni and Lorenz Curves

The Bonferroni and Lorenz curves, introduced by [18], find applications in the fields of economics and insurance. Now, we present the Bonferroni and Lorenz curves of the $UGL(\alpha,\beta)$ distribution. Let the rv Y follow the $UGL(\alpha,\beta)$ distribution with the pdf given by Equation (3). The Bonferroni and Lorenz curves are provided

$$BC(\xi)=\frac{q^2}{4\xi\mu_1}\sum_{\ell=0}^{n}\varpi_\ell\,(z_\ell+1)f\left(\frac{q}{2}(z_\ell+1)\right)$$
$$=\frac{\alpha\beta^2q^2}{4(1+\beta)\xi\mu_1}\sum_{\ell=0}^{n}\varpi_\ell\,(z_\ell+1)$$
$$\times\left\{\frac{\left[1+\left(1-\left(\frac{q}{2}(z_\ell+1)\right)\right)(\alpha+\beta)\right]\left[\frac{q}{2}(z_\ell+1)\right]^{\alpha-1}}{\beta\left[\frac{q}{2}(z_\ell+1)\right]-\left[\frac{q}{2}(z_\ell+1)\right]-\beta}\times\frac{\left(\beta-\beta\left[\frac{q}{2}(z_\ell+1)\right]+\left[\frac{q}{2}(z_\ell+1)\right]\right)^{-\alpha}}{\beta\left(\frac{q}{2}(z_\ell+1)\right)-\left(\frac{q}{2}(z_\ell+1)\right)-\beta}\right\}$$

and

$$LC(\xi)=\frac{q^2}{4\mu_1}\sum_{\ell=0}^{n}\varpi_\ell\,(z_\ell+1)f\left(\frac{q}{2}(z_\ell+1)\right)$$
$$=\frac{\alpha\beta^2q^2}{4(1+\beta)\mu_1}\sum_{\ell=0}^{n}\varpi_\ell\,(z_\ell+1)$$
$$\times\left\{\frac{\left[1+\left(1-\left(\frac{q}{2}(z_\ell+1)\right)\right)(\alpha+\beta)\right]\left[\frac{q}{2}(z_\ell+1)\right]^{\alpha-1}}{\beta\left[\frac{q}{2}(z_\ell+1)\right]-\left[\frac{q}{2}(z_\ell+1)\right]-\beta}\times\frac{\left(\beta-\beta\left[\frac{q}{2}(z_\ell+1)\right]+\left[\frac{q}{2}(z_\ell+1)\right]\right)^{-\alpha}}{\beta\left(\frac{q}{2}(z_\ell+1)\right)-\left(\frac{q}{2}(z_\ell+1)\right)-\beta}\right\}$$

respectively, where ϖ_ℓ is given in Equation (4) and $q=F^1(\xi)$.

2.3. Order Statistics

In this subsection, we explore some results concerning the order statistics of the UGL distribution. Let $Y_1,Y_2,\ldots,Y_n$ be a random sample from the UGL distribution and $Y_{(1)},Y_{(2)},\ldots,Y_{(n)}$ indicates the related order statistics. The cdf and pdf of the $Y_{(r)}$ are presented for general form, respectively, by

$$F_{Y_{(r)}}(y;\alpha,\beta)=\sum_{i=r}^{n}\binom{n}{i}F(y;\alpha,\beta)^i\{1-F(y;\alpha,\beta)\}^{n-i}$$
$$=\sum_{i=r}^{n}\sum_{j=0}^{n-i}(-1)^j\binom{n}{i}\binom{n-i}{j}F(y;\alpha,\beta)^{i+j},$$

and

$$f_{Y_{(r)}}(y;\alpha,\beta)=\frac{1}{B(r,n-r+1)}F(y;\alpha,\beta)^{r-1}\{1-F(y;\alpha,\beta)\}^{n-r}\,f(y;\alpha,\beta)$$
$$=\frac{1}{B(r,n-r+1)}\sum_{i=0}^{n-r}(-1)^i\binom{n-r}{i}F(y;\alpha,\beta)^{r+i-1}\,f(y;\alpha,\beta),$$

where $r=1,2,\ldots,n$ and $B(.;.)$ is the classical beta function. The cdf and pdf of the $Y_{(r)}$ order statistic of the UGL distribution is also obtained by

$$F_{Y_{(r)}}\left(y;\alpha,\beta\right)=\sum_{i=r}^{n}\sum_{j=0}^{n-i}\left(-1\right)^{j}\binom{n}{i}\binom{n-i}{j}$$
$$\times\left\{\frac{\left(\frac{y}{1-y}\right)^{\alpha}}{\left(\beta+\frac{y}{1-y}\right)^{\alpha}}+\frac{\beta\alpha\left(\frac{y}{1-y}\right)^{\alpha}}{\left(1+\beta\right)\left(\beta+\frac{y}{1-y}\right)^{1+\alpha}}\right\}^{i+j},$$

and

$$f_{X_{(r)}}\left(y;\alpha,\beta\right)=\frac{\alpha\beta^{2}\left(1+\left(1-y\right)\left(\alpha+\beta\right)\right)y^{\alpha-1}\left(\beta-\beta y+y\right)^{-\alpha}}{\left(1+\beta\right)\left(\beta y-y-\beta\right)^{2}}$$
$$\times\frac{1}{B\left(r,n-r+1\right)}\sum_{i=0}^{n-r}\left(-1\right)^{i}\binom{n-r}{i}\left\{\frac{\left(\frac{y}{1-y}\right)^{\alpha}}{\left(\beta+\frac{y}{1-y}\right)^{\alpha}}+\frac{\beta\alpha\left(\frac{y}{1-y}\right)^{\alpha}}{\left(1+\beta\right)\left(\beta+\frac{y}{1-y}\right)^{1+\alpha}}\right\}^{r+i-1}.$$

When $r=1$ and $r=n$, the cdf and pdf of $Y_{(1)}=\min\{Y_1,Y_2,\ldots,Y_n\}$ and $Y_{(n)}=\max\{Y_1,Y_2,\ldots,Y_n\}$ statistics can be easily achieved, respectively.

2.4. Data Generation for UGL Distribution

This subsection introduces an acceptance-rejection (AR) algorithm for generating data from the UGL distribution. For a comprehensive explanation of the AR algorithm, refer to [19]. The steps of the AR algorithm for generating data from the UGL distribution are outlined in Algorithm 1.

Algorithm 1

Step 1. Generating a sample random variable X from the standard uniform distribution with the following pdf

$$h\left(x\right)=1,\ 0<x<1,$$

Step 2. A sample is generated from a uniform distribution independent from X, say $U\sim U\left(0,1\right)$.

Step 3. If

$$U<\frac{f\left(X\right)}{k\times h\left(X\right)}$$

then $X=Y$ (Accept), otherwise go to **Step 1** (Reject), where f is the pdf in Equation (3) and

$$k=\max_{y\in R^{+}}\frac{f\left(x\right)}{h\left(x\right)}$$

The output of Algorithm 1 gives a sample for the variable X from the UGL distribution.

2.5. Mode

Mode is the frequently occurring value in a set of data. For the UGL, the corresponding mode can be obtained by Equation $f^{'}(y)=0$. That is obtained by,

$$\frac{2\beta^2\left((1-\beta)y+\beta\right)^{-\alpha}\alpha\left(\left(y+\frac{\alpha}{2}-\frac{1}{2}\right)(y-1)\beta^2+\left((\alpha-1)y^2+\left(\frac{\alpha^2}{2}-\frac{3\alpha}{2}\right)y+\frac{1}{2}-\frac{\alpha^2}{2}\right)\beta-y\left(y\alpha-\frac{3\alpha}{2}-\frac{3}{2}\right)\right)y(\alpha-1)}{y\left(\beta(y-1)-y\right)^3}=0,$$

where, $y\in(0,1)$ and $\alpha>0,\beta>0$.

2.6. Entropy

Entropy quantifies the uncertainty associated with the distribution of a random variable Y. The Tsallis entropy is commonly used to quantify the randomness of a random variable. For an in-depth discussion on the roles of various entropy measures in applied sciences, including the Tsallis entropy, see [20]. The entropy of a random variable Y is defined as:

$$T_\tau=\frac{1}{\tau-1}\left[1-\int_0^1 f(y)^\tau\,dy\right]$$

$$=\frac{1}{\tau-1}\left[1-\int_0^1\left(\frac{\alpha\beta^2\left(1+(1-y)(\alpha+\beta)\right)y^{\alpha-1}(\beta-\beta y+y)^{-\alpha}}{(1+\beta)(\beta y-y-\beta)^2}\right)^\tau dy\right]$$

$$=\frac{1}{\tau-1}\left[1-\left(\frac{1}{2}\int_{-1}^1\left(\frac{z+1}{2}\right)^{r\tau}f\left(\frac{z+1}{2}\right)dz\right)\right]$$

$$=\frac{1}{\tau-1}\left[1-\left(\frac{\alpha^\tau\beta^{2\tau}}{2(1+\beta)^\tau}\int_{-1}^1\left(\frac{z+1}{2}\right)^{r\tau+\alpha-1}\left(\frac{\left(1+\left(1-\left(\frac{z+1}{2}\right)\right)(\alpha+\beta)\right)\left(\beta-\beta\left(\frac{z+1}{2}\right)+\left(\frac{z+1}{2}\right)\right)^{-\alpha}}{\left(\beta\left(\frac{z+1}{2}\right)-\left(\frac{z+1}{2}\right)-\beta\right)^2}\right)dz\right)\right]$$

$$\cong\frac{1}{\tau-1}\left[1-\left(\frac{\alpha^\tau\beta^{2\tau}}{2(1+\beta)^\tau}\sum_{\ell=1}^N\varpi_\ell\left(\frac{z_\ell+1}{2}\right)^{r\tau+\alpha-1}\left(\frac{\left(1+\left(1-\left(\frac{z_\ell+1}{2}\right)\right)(\alpha+\beta)\right)\left(\beta-\beta\left(\frac{z_\ell+1}{2}\right)+\left(\frac{z_\ell+1}{2}\right)\right)^{-\alpha}}{\left(\beta\left(\frac{z_\ell+1}{2}\right)-\left(\frac{z_\ell+1}{2}\right)-\beta\right)^2}\right)\right)\right],$$

where ϖ_ℓ and $L_{N+1}^{'}(z_\ell)$ are presented in Equations (4)-(5), respectively.

2.7. Quantile Function

It may be noted that $F(y)$ in Equation (2) is continuous and strictly increasing, so the quantile function of Y is defined as follows:

$$Q_Y(u)=y_u=F_Y^{-1}(u),\ u\in[0,1].$$

3. DIFFERENT POINT ESTIMATION METHODS

In this section, several estimators for estimating the parameters of the UGL distribution are proposed. We discuss the estimators, such as maximum likelihood (ML), least squares (LS), weighted least squares (WLS), Cramer-von Mises (CvM), and Anderson-Darling (AD), and compare the performances of these estimators via a Monte Carlo simulation. Let $Y_1,Y_2,\ldots,Y_n$ denote random samples from the $UGL(\alpha,\beta)$ distribution, and let $y_1,y_2,\ldots,y_n$ represent the observed values of the sample. Further, let

$Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ denote the order statistics based on the sample $y_1, y_2, \ldots, y_n$, with the realizations $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$. The likelihood and log-likelihood functions can be expressed, respectively, as

$$L(\Xi) = \prod_{i=1}^{n} \frac{\alpha\beta^2 \left(1+(1-y_i)(\alpha+\beta)\right) y_i^{\alpha-1} \left(\beta - \beta y_i + y_i\right)^{-\alpha}}{(1+\beta)\left(\beta y_i - y_i - \beta\right)^2}$$

and

$$\ell(\Xi) = n\log\left(\alpha\beta^2\right)\sum_{i=1}^{n}\log\left(1+(1-y_i)(\alpha+\beta)\right) + (\alpha-1)\sum_{i=1}^{n}\log(y_i)$$
$$-2(1+\beta)\sum_{i=1}^{n}\log\left(\beta y_i - y_i - \beta\right) - \alpha\sum_{i=1}^{n}\log\left(\beta - \beta y_i + y_i\right),$$

where $\Xi = (\alpha, \beta)$. Then the ML of $\Xi = \left(\hat{\alpha}, \hat{\beta}\right)$ for Ξ given as

$$\Xi_1 = \arg\max \ell(\Xi)$$

$$(\alpha, \beta) \in (0, \infty) \times (0, \infty).$$

Let us address the following functions to derive other estimators

$$Q_{LS}(\Xi) = \sum_{i=1}^{n}\left(\left(\frac{\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}} + \frac{\beta\alpha\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{(1+\beta)\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{1+\alpha}}\right) - \frac{i}{n+1}\right)^2, \quad (6)$$

$$Q_{WLS}(\Xi) = \sum_{i=1}^{n}\frac{(n+2)(n+1)^2}{i(n-i+1)}\left(\left(\frac{\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}} + \frac{\beta\alpha\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{(1+\beta)\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{1+\alpha}}\right) - \frac{i}{n+1}\right)^2, \quad (7)$$

$$Q_{AD}(\Xi) = -n - \sum_{i=1}^{n}\frac{2i-1}{n}\left[\log\left\{\frac{\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}} + \frac{\beta\alpha\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{(1+\beta)\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{1+\alpha}}\right\} + \log\left\{1-\left(\frac{\left(\frac{y_{(n+i-1)}}{1-y_{(n+i-1)}}\right)^{\alpha}}{\left(\beta+\frac{y_{(n+i-1)}}{1-y_{(n+i-1)}}\right)^{\alpha}} + \frac{\beta\alpha\left(\frac{y_{(n+i-1)}}{1-y_{(n+i-1)}}\right)^{\alpha}}{(1+\beta)\left(\beta+\frac{y_{(n+i-1)}}{1-y_{(n+i-1)}}\right)^{1+\alpha}}\right)\right\}\right], \quad (8)$$

and

$$Q_{CvM}(\Xi)=\frac{1}{12n}+\sum_{i=1}^{n}\left[\left(\frac{\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}+\frac{\beta\alpha\left(\frac{y_{(i)}}{1-y_{(i)}}\right)^{\alpha}}{(1+\beta)\left(\beta+\frac{y_{(i)}}{1-y_{(i)}}\right)^{1+\alpha}}\right)-\frac{2i-1}{2n}\right]^{2}. \quad (9)$$

Then the LS, WLS, AD, and CvM are obtained by minimizing Equations (6)-(9), respectively, as

$$\Xi_2 = \arg\min Q_{LS}(\Xi)$$

$$(\alpha,\beta)\in(0,\infty)\times(0,\infty)$$

$$\Xi_3 = \arg\min Q_{WLS}(\Xi)$$

$$(\alpha,\beta)\in(0,\infty)\times(0,\infty)$$

$$\Xi_4 = \arg\min Q_{AD}(\Xi)$$

$$(\alpha,\beta)\in(0,\infty)\times(0,\infty)$$

$$\Xi_5 = \arg\min Q_{CvM}(\Xi)$$

$$(\alpha,\beta)\in(0,\infty)\times(0,\infty).$$

Some numerical methods, such as Nelder-Mead or BFGS, can solve all maximizing and minimization problems. In the following, a simulation study is conducted to evaluate the performance of these estimators.

4. SIMULATION EXPERIMENT FOR POINT ESTIMATES

In this section, a Monte Carlo simulation experiment is conducted with 5000 runs. The bias and mean square errors (MSE) of ML, LS, WLS, AD, and CvM of unknown parameters of the UGL distribution are estimated. Six different parameter settings are considered, and the sample sizes are chosen as $n = 25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900$ and 1000. The simulation results are reported in Tables 1-2. It is observed from these tables that the bias and MSE of all estimators tend to decrease towards zero as the sample size increases. In cases with small sample sizes, the AD performs the best based on the bias criterion, while the ML is the best according to the MSE criterion. Despite the convergence of bias and MSE for all estimators as the sample size increases, AD remains superior in bias, and ML excels in MSE.

Table 1. *The bias values for* α *and* β *parameters*

			α					β				
n	α	β	ML	LS	WLS	AD	CvM	ML	LS	WLS	AD	CvM
25	1	0.75	0.2247	0.1312	0.1531	0.1382	0.3438	0.0048	0.1731	0.1287	0.0907	0.0313
50			0.0970	0.0448	0.0561	0.0595	0.1281	-0.0035	0.0799	0.0500	0.0388	0.0138
100			0.0482	0.0234	0.0307	0.0315	0.0615	-0.0035	0.0358	0.0191	0.0168	0.0039
200			0.0208	0.0084	0.0136	0.0127	0.0266	-0.0005	0.0196	0.0093	0.0098	0.0038
300			0.0145	0.0066	0.0098	0.0093	0.0186	-0.0005	0.0120	0.0057	0.0061	0.0015
400			0.0114	0.0051	0.0082	0.0073	0.0140	-0.0008	0.0088	0.0035	0.0042	0.0010
500			0.0090	0.0028	0.0058	0.0052	0.0099	0.0005	0.0091	0.0044	0.0050	0.0028
600			0.0071	0.0023	0.0046	0.0041	0.0082	0.0007	0.0078	0.0039	0.0044	0.0026
700			0.0059	0.0030	0.0045	0.0040	0.0081	0.0004	0.0053	0.0025	0.0030	0.0009
800			0.0060	0.0031	0.0045	0.0040	0.0075	-0.0003	0.0045	0.0018	0.0023	0.0006
900			0.0034	0.0001	0.0017	0.0013	0.0040	0.0013	0.0059	0.0033	0.0037	0.0024
1000			0.0037	0.0009	0.0021	0.0018	0.0044	0.0000	0.0040	0.0019	0.0022	0.0009
25	2	1.5	0.6988	0.3069	0.6605	0.3819	0.8881	0.0514	0.6483	0.4792	0.3750	0.2484
50			0.3531	0.2425	0.3073	0.2392	0.5323	0.0243	0.2943	0.1971	0.1696	0.1099
100			0.1880	0.1305	0.1422	0.1290	0.2635	-0.0052	0.1306	0.0736	0.0712	0.0430
200			0.0896	0.0596	0.0684	0.0617	0.1204	-0.0026	0.0636	0.0324	0.0353	0.0208
300			0.0579	0.0345	0.0428	0.0380	0.0738	-0.0037	0.0429	0.0195	0.0227	0.0146
400			0.0416	0.0263	0.0321	0.0274	0.0554	-0.0017	0.0319	0.0143	0.0176	0.0108
500			0.0403	0.0272	0.0335	0.0294	0.0505	-0.0050	0.0237	0.0078	0.0107	0.0068
600			0.0270	0.0225	0.0245	0.0211	0.0418	0.0006	0.0192	0.0086	0.0110	0.0052
700			0.0232	0.0191	0.0209	0.0182	0.0356	-0.0002	0.0158	0.0068	0.0090	0.0038
800			0.0146	0.0099	0.0125	0.0096	0.0242	0.0048	0.0200	0.0111	0.0133	0.0095
900			0.0183	0.0114	0.0144	0.0123	0.0240	0.0000	0.0151	0.0069	0.0086	0.0058
1000			0.0139	0.0079	0.0113	0.0092	0.0193	0.0006	0.0144	0.0063	0.0081	0.0060
25	1.5	2	0.4785	0.2463	0.4124	0.2726	0.6733	0.0704	0.8044	0.6223	0.4672	0.3002
50			0.2419	0.1505	0.1756	0.1496	0.3482	0.0180	0.3633	0.2441	0.2060	0.1291
100			0.1128	0.0801	0.0844	0.0773	0.1672	0.0033	0.1614	0.0943	0.0909	0.0503
200			0.0544	0.0358	0.0404	0.0365	0.0759	-0.0004	0.0798	0.0412	0.0451	0.0255
300			0.0330	0.0219	0.0267	0.0235	0.0479	0.0044	0.0557	0.0280	0.0317	0.0198
400			0.0276	0.0133	0.0186	0.0166	0.0325	-0.0034	0.0424	0.0196	0.0224	0.0156
500			0.0182	0.0156	0.0173	0.0150	0.0310	0.0042	0.0299	0.0140	0.0171	0.0085
600			0.0170	0.0111	0.0132	0.0116	0.0238	0.0006	0.0263	0.0131	0.0155	0.0085
700			0.0157	0.0106	0.0127	0.0111	0.0215	-0.0004	0.0223	0.0100	0.0125	0.0070
800			0.0130	0.0071	0.0099	0.0085	0.0166	-0.0007	0.0203	0.0088	0.0110	0.0070
900			0.0091	0.0051	0.0069	0.0055	0.0135	0.0029	0.0206	0.0108	0.0130	0.0088
1000			0.0072	0.0027	0.0046	0.0035	0.0102	0.0052	0.0219	0.0129	0.0147	0.0112

Table 1. *Contunied*

			α					β				
n	α	β	ML	LS	WLS	AD	CvM	ML	LS	WLS	AD	CvM
25	1.25	2	0.3867	0.2179	0.2870	0.2188	0.5614	0.0541	0.7037	0.5376	0.4004	0.2325
50			0.1801	0.1175	0.1254	0.1191	0.2635	-0.0025	0.3005	0.1889	0.1590	0.0834
100			0.0803	0.0417	0.0517	0.0507	0.1039	-0.0005	0.1592	0.0896	0.0835	0.0551
200			0.0398	0.0228	0.0285	0.0268	0.0522	-0.0002	0.0749	0.0378	0.0398	0.0241
300			0.0204	0.0083	0.0134	0.0116	0.0274	0.0089	0.0587	0.0329	0.0360	0.0249
400			0.0206	0.0122	0.0154	0.0141	0.0265	-0.0035	0.0322	0.0136	0.0160	0.0071
500			0.0124	0.0084	0.0109	0.0094	0.0197	0.0025	0.0273	0.0122	0.0150	0.0072
600			0.0111	0.0054	0.0079	0.0068	0.0148	0.0012	0.0264	0.0127	0.0148	0.0097
700			0.0114	0.0087	0.0102	0.0092	0.0168	-0.0005	0.0161	0.0060	0.0081	0.0019
800			0.0109	0.0075	0.0092	0.0086	0.0146	-0.0002	0.0157	0.0065	0.0078	0.0032
900			0.0070	0.0033	0.0049	0.0041	0.0095	0.0031	0.0197	0.0106	0.0123	0.0086
1000			0.0068	0.0048	0.0060	0.0052	0.0105	0.0016	0.0136	0.0063	0.0077	0.0036
25	2.5	4	1.6507	0.3397	1.8831	0.9183	1.0719	0.3641	1.9205	2.2856	1.7456	0.8820
50			0.8882	0.5512	1.0067	0.6504	1.0289	0.0141	1.0919	0.7986	0.7137	0.4806
100			0.4253	0.4062	0.4137	0.3402	0.6606	-0.0304	0.5767	0.3149	0.3206	0.2506
200			0.1886	0.1821	0.1662	0.1424	0.2985	-0.0074	0.2805	0.1478	0.1653	0.1182
300			0.1139	0.1354	0.1172	0.0967	0.2095	0.0038	0.1832	0.0933	0.1109	0.0760
400			0.0869	0.1029	0.0899	0.0758	0.1567	-0.0062	0.1195	0.0506	0.0684	0.0396
500			0.0738	0.0712	0.0688	0.0592	0.1131	-0.0129	0.1122	0.0451	0.0584	0.0482
600			0.0602	0.0726	0.0651	0.0552	0.1075	-0.0024	0.0790	0.0303	0.0438	0.0259
700			0.0545	0.0572	0.0523	0.0454	0.0867	-0.0145	0.0639	0.0218	0.0321	0.0185
800			0.0466	0.0440	0.0434	0.0367	0.0695	-0.0097	0.0688	0.0263	0.0359	0.0290
900			0.0361	0.0437	0.0376	0.0315	0.0664	-0.0013	0.0528	0.0219	0.0314	0.0174
1000			0.0377	0.0398	0.0371	0.0321	0.0601	-0.0089	0.0452	0.0148	0.0227	0.0134
25	4	5	4.9023	0.2057	5.0886	2.0289	1.1241	0.5935	2.2629	4.2298	2.8032	1.0928
50			2.4317	0.3903	2.8675	1.5411	1.0232	0.2696	1.7143	1.8167	1.6307	0.9164
100			1.3474	0.7658	1.6478	1.1830	1.2096	-0.0515	0.9658	0.6030	0.6132	0.4865
200			0.5778	0.5878	0.6453	0.5296	0.8539	0.0041	0.6265	0.3240	0.3645	0.3521
300			0.3795	0.4588	0.4219	0.3525	0.6383	-0.0182	0.3992	0.1900	0.2293	0.2141
400			0.2674	0.3453	0.2937	0.2487	0.4788	-0.0138	0.2848	0.1293	0.1610	0.1455
500			0.1951	0.2639	0.2218	0.1868	0.3686	0.0010	0.2295	0.1040	0.1300	0.1176
600			0.1706	0.2191	0.1761	0.1487	0.3046	-0.0046	0.2045	0.0970	0.1214	0.1110
700			0.1330	0.1835	0.1486	0.1246	0.2560	0.0090	0.1890	0.0925	0.1156	0.1088
800			0.1253	0.1741	0.1380	0.1166	0.2372	0.0038	0.1579	0.0774	0.0977	0.0877
900			0.0983	0.1594	0.1243	0.1034	0.2150	0.0095	0.1130	0.0488	0.0693	0.0509
1000			0.1012	0.1282	0.1062	0.0892	0.1775	-0.0054	0.1302	0.0586	0.0765	0.0741

Table 2. *The MSE values for α and β parameters*

			α					β				
n	α	β	ML	LS	WLS	AD	CvM	ML	LS	WLS	AD	CvM
25	1	0.75	0.4297	0.5754	0.9461	0.3565	1.0594	0.1563	0.3494	0.2679	0.2376	0.2323
50			0.1057	0.1596	0.1582	0.1106	0.2206	0.0692	0.1219	0.0954	0.0887	0.0995
100			0.0401	0.0573	0.0466	0.0434	0.0675	0.0321	0.0486	0.0395	0.0384	0.0440
200			0.0171	0.0255	0.0203	0.0194	0.0276	0.0161	0.0235	0.0190	0.0187	0.0223
300			0.0106	0.0160	0.0126	0.0123	0.0169	0.0104	0.0145	0.0119	0.0119	0.0140
400			0.0083	0.0123	0.0097	0.0095	0.0128	0.0082	0.0111	0.0092	0.0091	0.0108
500			0.0063	0.0093	0.0074	0.0073	0.0096	0.0063	0.0086	0.0071	0.0071	0.0084
600			0.0054	0.0082	0.0064	0.0063	0.0084	0.0055	0.0076	0.0063	0.0063	0.0075
700			0.0044	0.0068	0.0053	0.0052	0.0069	0.0044	0.0061	0.0051	0.0050	0.0060
800			0.0038	0.0058	0.0045	0.0045	0.0060	0.0039	0.0053	0.0044	0.0044	0.0053
900			0.0034	0.0051	0.0039	0.0039	0.0052	0.0034	0.0046	0.0038	0.0039	0.0046
1000			0.0031	0.0045	0.0036	0.0035	0.0046	0.0031	0.0042	0.0035	0.0035	0.0042
25	2	1.5	3.3429	3.2902	15.1646	2.8230	5.5874	0.9129	2.9693	2.1326	1.8524	1.8547
50			1.2725	2.0704	3.9096	1.4734	2.9107	0.4148	0.9980	0.7092	0.6507	0.7687
100			0.4329	0.7376	0.5762	0.4840	0.9347	0.1910	0.3814	0.2799	0.2697	0.3345
200			0.1599	0.2818	0.2115	0.1952	0.3183	0.0966	0.1699	0.1277	0.1263	0.1589
300			0.0962	0.1699	0.1243	0.1185	0.1842	0.0622	0.1085	0.0808	0.0811	0.1037
400			0.0702	0.1268	0.0923	0.0884	0.1347	0.0464	0.0793	0.0592	0.0592	0.0766
500			0.0569	0.1042	0.0755	0.0727	0.1095	0.0377	0.0657	0.0489	0.0489	0.0640
600			0.0453	0.0833	0.0601	0.0580	0.0869	0.0321	0.0538	0.0410	0.0409	0.0527
700			0.0389	0.0700	0.0516	0.0505	0.0726	0.0273	0.0458	0.0351	0.0352	0.0450
800			0.0318	0.0593	0.0421	0.0412	0.0610	0.0229	0.0399	0.0297	0.0298	0.0391
900			0.0292	0.0524	0.0380	0.0374	0.0539	0.0213	0.0354	0.0270	0.0270	0.0349
1000			0.0256	0.0481	0.0344	0.0339	0.0492	0.0186	0.0317	0.0239	0.0240	0.0313
25	1.5	2	1.7073	1.9418	8.4897	1.5200	3.4496	1.6727	4.8103	3.8506	3.2248	3.1353
50			0.5502	0.8672	1.4699	0.5221	1.3004	0.7134	1.6408	1.1852	1.1005	1.2739
100			0.1670	0.3116	0.2328	0.1889	0.3950	0.3295	0.6294	0.4625	0.4512	0.5534
200			0.0705	0.1222	0.0897	0.0840	0.1370	0.1566	0.2771	0.2055	0.2042	0.2595
300			0.0433	0.0774	0.0574	0.0540	0.0835	0.1090	0.1815	0.1384	0.1377	0.1736
400			0.0312	0.0536	0.0395	0.0381	0.0567	0.0815	0.1344	0.1025	0.1026	0.1299
500			0.0249	0.0451	0.0327	0.0315	0.0472	0.0654	0.1070	0.0818	0.0814	0.1043
600			0.0198	0.0354	0.0259	0.0252	0.0368	0.0526	0.0859	0.0666	0.0664	0.0840
700			0.0176	0.0313	0.0227	0.0223	0.0323	0.0454	0.0749	0.0569	0.0571	0.0735
800			0.0143	0.0253	0.0184	0.0181	0.0260	0.0379	0.0620	0.0475	0.0476	0.0610
900			0.0129	0.0235	0.0169	0.0166	0.0240	0.0354	0.0581	0.0445	0.0446	0.0572
1000			0.0114	0.0203	0.0147	0.0145	0.0207	0.0312	0.0502	0.0388	0.0389	0.0494

Table 2. *Contunied*

			α					β				
n	α	β	ML	LS	WLS	AD	CvM	ML	LS	WLS	AD	CvM
25	1.25	2	1.0619	1.3018	2.9529	0.8140	2.3880	1.5850	4.2485	3.3815	2.7909	2.7744
50			0.3062	0.4835	0.4163	0.3064	0.7097	0.6280	1.3745	0.9856	0.9234	1.0794
100			0.0955	0.1553	0.1203	0.1064	0.1901	0.3055	0.5636	0.4176	0.4069	0.4958
200			0.0418	0.0688	0.0519	0.0492	0.0762	0.1516	0.2531	0.1925	0.1912	0.2374
300			0.0241	0.0402	0.0304	0.0292	0.0429	0.0972	0.1566	0.1212	0.1212	0.1493
400			0.0182	0.0304	0.0227	0.0221	0.0320	0.0729	0.1114	0.0878	0.0874	0.1080
500			0.0136	0.0238	0.0178	0.0172	0.0248	0.0576	0.0906	0.0712	0.0710	0.0883
600			0.0115	0.0198	0.0146	0.0143	0.0205	0.0490	0.0780	0.0606	0.0604	0.0763
700			0.0096	0.0162	0.0119	0.0117	0.0168	0.0409	0.0624	0.0492	0.0491	0.0614
800			0.0087	0.0143	0.0108	0.0107	0.0147	0.0373	0.0555	0.0445	0.0444	0.0546
900			0.0075	0.0132	0.0096	0.0094	0.0135	0.0323	0.0501	0.0391	0.0390	0.0493
1000			0.0064	0.0112	0.0082	0.0081	0.0114	0.0289	0.0438	0.0347	0.0345	0.0433
25	2.5	4	18.5782	4.4395	71.9526	11.3363	7.4128	11.9828	19.0966	37.0716	27.6431	13.5818
50			5.7239	5.1960	21.4893	6.0972	7.3752	4.8213	11.4205	9.8272	9.2041	8.8636
100			1.6451	3.0909	2.9626	2.0307	3.9604	2.2766	5.7218	3.7982	3.7774	4.9035
200			0.4937	1.0781	0.7234	0.6178	1.2835	1.1064	2.3734	1.6362	1.6352	2.1858
300			0.2840	0.7510	0.5058	0.4091	0.8426	0.7459	1.5809	1.0901	1.0885	1.4973
400			0.1972	0.4552	0.2972	0.2733	0.4996	0.5562	1.1537	0.7948	0.7960	1.1103
500			0.1560	0.3521	0.2325	0.2184	0.3786	0.4446	0.9153	0.6340	0.6375	0.8854
600			0.1240	0.2860	0.1865	0.1747	0.3048	0.3690	0.7568	0.5226	0.5216	0.7377
700			0.1019	0.2268	0.1477	0.1411	0.2395	0.3119	0.6275	0.4372	0.4378	0.6141
800			0.0864	0.1965	0.1302	0.1242	0.2057	0.2657	0.5448	0.3783	0.3775	0.5336
900			0.0773	0.1748	0.1133	0.1091	0.1823	0.2504	0.5004	0.3463	0.3473	0.4919
1000			0.0705	0.1540	0.1020	0.0988	0.1600	0.2221	0.4382	0.3076	0.3079	0.4316
25	4	5	106.6929	6.6061	254.3130	34.6963	9.8095	31.3769	26.0113	122.4950	60.3000	18.3926
50			36.7571	7.1169	116.8464	25.4971	9.5351	13.6862	19.7807	33.3856	29.3658	15.0285
100			12.8277	8.9829	34.8834	15.214	10.9649	5.8766	12.2858	10.9129	10.8425	10.3872
200			3.5843	6.5001	7.1514	5.2699	7.6006	3.0028	7.5649	4.9933	5.0688	6.8010
300			1.8626	4.3405	3.4378	2.7599	4.9170	2.0021	5.0564	3.2884	3.3571	4.7201
400			1.1797	3.0534	2.0186	1.7662	3.4214	1.4552	3.4784	2.2848	2.2964	3.3013
500			0.8322	2.2720	1.4545	1.2772	2.5024	1.1569	2.6895	1.7775	1.7796	2.5783
600			0.6857	1.8422	1.1225	1.0215	1.9953	0.9718	2.2692	1.4808	1.4917	2.1887
700			0.5692	1.5219	0.9459	0.8757	1.6350	0.8539	2.0708	1.3561	1.3683	2.0072
800			0.4883	1.3571	0.8249	0.7664	1.4467	0.7333	1.7327	1.1440	1.1478	1.6860
900			0.4045	1.0938	0.6693	0.6229	1.1604	0.6387	1.4576	0.9588	0.9596	1.4254
1000			0.3724	0.9991	0.6149	0.5797	1.0515	0.5877	1.3841	0.9102	0.9175	1.3538

5. A NOVEL REGRESSION ANALYSIS

In this section, we introduce a novel regression analysis based on the UGL distribution. Let $\log(\alpha_i) = z_i^T\boldsymbol{\gamma}$, where $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \ldots, \gamma_p)^T \in \mathbb{R}^p$ and $z_i = (1, z_{i1}, z_{i2}, \ldots, z_{ip})$ and $i = 1, 2, \ldots, n.$ The reparameterization version of the cdf and pdf are presented

$$F_{RUGL}(y_i;\boldsymbol{\psi}) = \frac{\left(\frac{y_i}{1-y_i}\right)^{\exp\left(z_i^T\gamma\right)}}{\left(\beta+\frac{y_i}{1-y_i}\right)^{\alpha}} + \frac{\beta\left(\exp\left(z_i^T\boldsymbol{\gamma}\right)\right)\left(\frac{y_i}{1-y_i}\right)^{\exp\left(z_i^T\gamma\right)}}{(1+\beta)\left(\beta+\frac{y_i}{1-y_i}\right)^{1+e\,\mathrm{xp}\left(z_i^T\gamma\right)}},$$

and

$$f_{RUGL}(y_i;\boldsymbol{\psi}) = \frac{e\,\mathrm{xp}\left(z_i^T\boldsymbol{\gamma}\right)\beta^2\left(1+(1-y_i)\left(e\,\mathrm{xp}\left(z_i^T\boldsymbol{\gamma}\right)+\beta\right)\right)y_i^{e\,\mathrm{xp}\left(z_i^T\gamma\right)-1}\left(\beta-\beta y_i+y_i\right)^{-e\,\mathrm{xp}\left(z_i^T\gamma\right)}}{(1+\beta)\left(\beta y_i-y_i-\beta\right)^2},$$

respectively, where $\boldsymbol{\psi} = (\boldsymbol{\gamma}, \beta)$. In the remainder of the manuscript, the rv Y will be denoted by $Y \sim RUGL(\boldsymbol{\psi})$.

5.1. Estimation for Regression Model Parameters

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the $RUGL(\boldsymbol{\psi})$ distribution. The corresponding log-likelihood function is expressed as

$$\begin{aligned}\ell(\boldsymbol{\psi}) &= 2n\log(\beta)\sum_{i=1}^{n} z_i^T\gamma + \sum_{i=1}^{n}\log\left(1+(1-y_i)\left(\exp\left(z_i^T\gamma\right)+\beta\right)\right)\\ &-2\sum_{i=1}^{n}\log\left(\beta-\beta y_i+y_i\right)+\sum_{i=1}^{n}\left(\exp\left(z_i^T\gamma\right)-1\right)\log(y_i) \\ &-\sum_{i=1}^{n}\exp\left(z_i^T\gamma\right)\log\left(\beta-\beta y_i+y_i\right)-n\log(1+\beta)\end{aligned} \quad (10)$$

The ML of ψ, say $\hat{\boldsymbol{\psi}} = \left(\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_p, \hat{\beta}\right)$, is obtained by maximizing $\ell(\boldsymbol{\psi})$ in Equation (10) as follow:

$$\boldsymbol{\psi}_1 = \arg\max \ell(\boldsymbol{\psi}).$$

Under some regularity conditions, the asymptotic distribution of $(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi})$ is a multivariate normal $N_{p+1}\left(0, J^{-1}\right)$ where J represents the expected information matrix. In practice, the observed information matrix is often used instead of J.

6. REAL-LIFE DATA ANALYSES

In this section, two real-life data analyses are carried out to demonstrate the applicability of the proposed distribution and regression model.

6.1. Application of New Distribution without Covariates

In this subsection, we compare the goodness of fit of the UGL distribution with the gamma and Lindley (GL) [16], unit Weibull (UW) [1], Kumaraswamy (KW) [21], unit Burr-XII (UB-XII) [22], and beta (B) distributions. The pdfs of GL, UW, KW, UB-XII, and B are given, respectively, by

$$f_{GL}\left(y;\alpha,\beta\right)=\frac{\alpha\beta^{2}}{1+\beta}\frac{\left(1+\alpha+\beta+y\right)y^{\alpha-1}}{\left(\beta+y\right)^{2+\alpha}}, \qquad y>0,\alpha>0,\beta>0,$$

$$f_{UW}\left(y;\alpha,\beta\right)=\frac{\alpha\beta\left(-\log\left(y\right)\right)^{\beta-1}\exp\left(-\alpha\left(-\log\left(y\right)\right)\right)^{\beta}}{y}, \qquad y>0,\alpha>0,\beta>0,$$

$$f_{KW}\left(y;\alpha,\beta\right)=\left(\left(1-y^{\alpha}\right)^{\beta-1}\beta y^{\alpha-1}\alpha\right), \qquad y>0,\alpha>0,\beta>0,$$

$$f_{UB-XII}\left(y;\alpha,\beta\right)=\alpha\beta y^{-1}\left(-\log\left(y\right)\right)^{\beta-1}\left(1+\left(-\log\left(y\right)\right)^{\beta}\right)^{-\alpha-1}, \qquad y>0,\alpha>0,\beta>0,$$

and

$$f_{B}\left(y;\alpha,\beta\right)=\frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}y^{\alpha-1}\left(1-y\right)^{\beta-1} \qquad y>0,\alpha>0,\beta>0.$$

The real data set is sourced from firms' risk management cost-effectiveness, which can be accessed on the personal web page of Professor E. Frees (Wisconsin School of Business). The data is characterized as the total property and casualty premiums and uninsured losses as a percentage of the total assets, resulting in a bound on (0,1). Detailed information about this data can be found in [23] and [24]. Some descriptive statistics about this data can be given as the mean is 0.1097, variance is 0.0261, skewness is 3.7938, kurtosis is 16.1223, and the median is 0.0608. The ℓ, Akaike's information criteria (AIC), Bayesian information criterion (BIC), consistent AIC (CAIC), Hannan-Quinn information criterion (HQIC), Kolmogorov-Smirnov (KS) statistics, and KS p-values are calculated, and they are presented in Table 3. It can be seen from Table 3 and Figures 3, 4 and 5 that the UGL distribution emerges as the best model according to all criteria.

Table 3. *Data analysis results for cost-effectiveness in firms' risk management data*

	GL	UGL	UW	KW	UB-XII	B
ℓ	92.0825	93.0552	88.1005	78.6539	46.5066	76.1175
AIC	-180.1650	-182.1105	-172.2010	-153.3079	-89.0132	-148.2350
BIC	-175.5841	-177.5296	-167.6201	-148.7269	-84.4323	-143.6541
CAIC	-179.9936	-181.9390	-172.0296	-153.1364	-88.8418	-148.0636
HQIC	-178.3394	-180.2849	-170.3754	-151.4823	-87.1876	-146.4095
KS	0.0818	0.0705	0.0931	0.1535	0.3641	0.1805
KS p-values	0.7128	0.8617	0.5516	0.0642	0.0000	0.0171
α	1.5909	1.3823	0.0653	0.6648	0.3482	0.6125
β	0.0749	0.0979	2.3527	3.4407	2.8408	3.7979

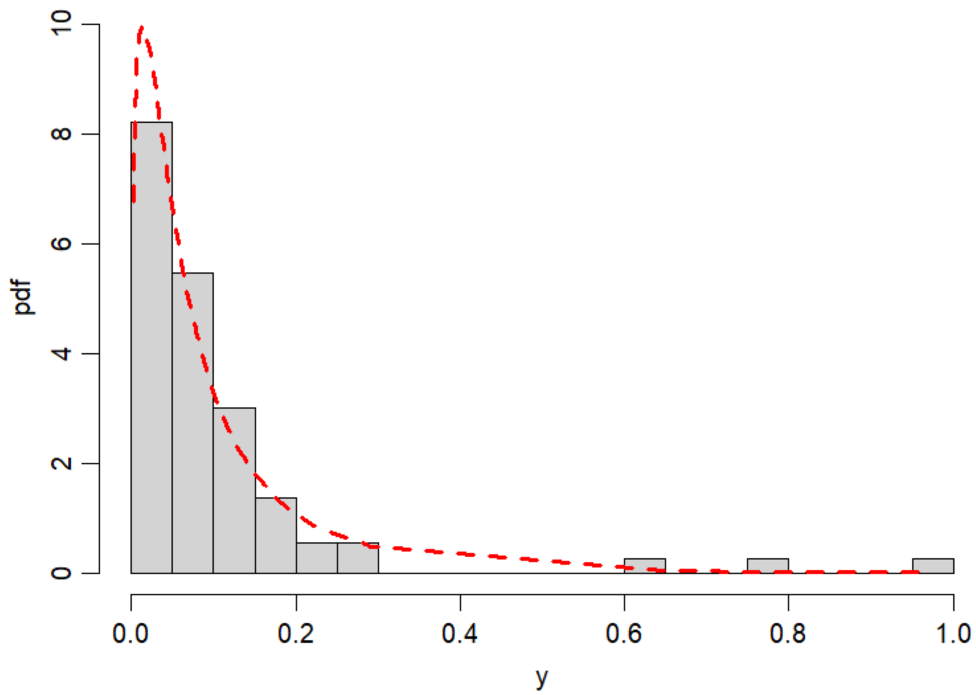

Figure 3. *The fitted pdf plot for UGL distribution on cost-effectiveness in firms' risk management data*

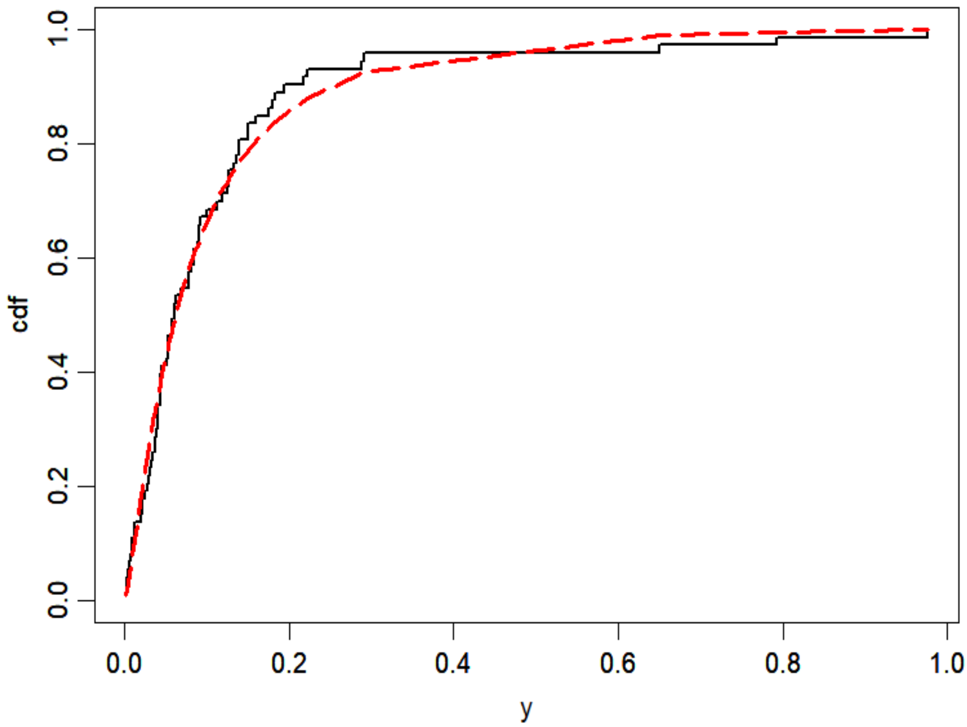

Figure 4. *The fitted cdf plot for UGL distribution on cost-effectiveness in firms' risk management data*

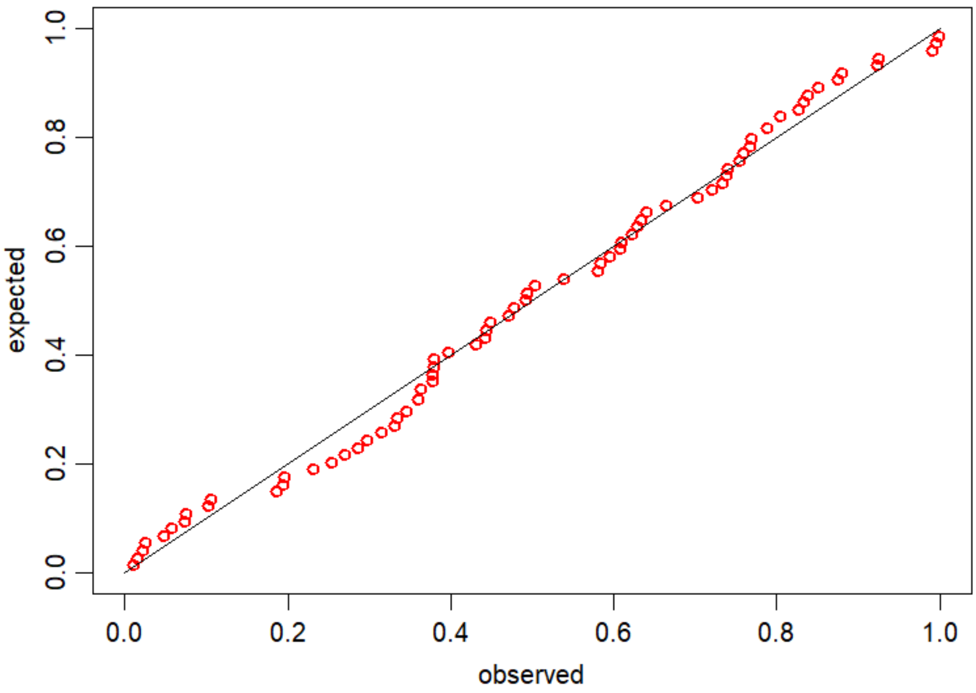

Figure 5. *The fitted probability-probability plot for UGL distribution on cost-effectiveness in firms' risk management data*

6.2. Application of New Regression Model with Covariates

In this subsection, a real data application is conducted to observe the usability and superiority of the new regression model. The data for the application consists of a percentage of educational attainment for OECD countries, as well as components such as percentage of voter turnout, homicide rate, and life satisfaction. It can also be reached through data at https://stats.oecd.org/. Detailed information about this dataset can be obtained from [14]. The beta regression (BR), Kumaraswamy regression (KR), and log-extended exponential geometric regression (LEEG) models [25] are considered for comparison purposes. This application aims to relate the percentage of education level values of OECD countries (variable y) with voter participation percentage (variable x_1), homicide rate (variable x_2), and life satisfaction (variable x_3). For all models, ML, standard error (se), $\hat{\ell}$, and AIC are calculated. The results are presented in Table 4. Table 4 shows that the RUGL model has the best modeling capacity and can be used as an alternative to the beta and Kumaraswamy regression models in the literature.

Table 4. *Data analysis results for regression real data*

	RUGL			BR			KR			LEGG		
Parameters	Estimate	SE	p-value	Estimate	SE	p-value	Estimate	SE	p-value	Estimate	SE	p-value
γ_0	4.2725	1.8469	0.0207	0.9615	0.9685	0.3208	1.6318	1.1629	0.1606	0.3269	1.0711	0.7602
γ_1	-2.0758	0.9827	0.0347	-2.9211	1.0176	0.0041	-4.1024	1.3947	0.0033	-4.0914	1.4627	0.0052
γ_2	-0.0674	0.0223	0.0025	-0.0470	0.0178	0.0083	-0.0405	0.0167	0.0153	-0.0476	0.0145	0.0010
γ_3	0.4916	0.1489	0.0010	0.3794	0.1492	0.0110	0.4206	0.2532	0.0967	0.6215	0.1746	<0.0001
β	0.0185	0.0276	<0.0001	11.5900	2.6100	<0.0001	6.2166	1.0844	<0.0001	7.8374	1.7424	<0.0001
$\hat{\ell}$	32.4682			30.9024			29.4339			28.6481		
AIC	-54.9363			-51.8048			-48.8678			-47.2962		

7.RESULTS

This paper introduces a novel bounded distribution with some mathematical properties, including Lorenz and Bonferroni curves. Several estimators are proposed to estimate unknown parameters of the distribution, and their performances are evaluated through Monte Carlo simulation. It is observed in the Monte Carlo simulation that all estimators perform well, particularly in large sample sizes. Furthermore, a new regression model based on the proposed distribution is introduced, and a comprehensive application for the distribution is implemented. As a result of real data analysis, it is determined that new models are highly effective in modeling the data on cost-effectiveness in firms' risk management and OECD education attainment, according to various criteria.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Mazucheli, J., Menezes, A.F.B., Fernandes, L. B., De Oliveira, R.P., and Ghitany, M.E., "The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. Journal of Applied Statistics", 47(6): 954-974, (2020). DOI: 10.1080/02664763.2019.1657813

[2] Bhatti, F. A., Ali, A., Hamedani, G., Korkmaz, M. Ç., and Ahmad, M. "The unit generalized log Burr XII distribution: Properties and applications", AIMS Mathematics. (2021). DOI: 10.3934/math.2021592

[3] Ghitany, M. E., Mazucheli, J., Menezes, A. F. B., and Alqallaf, F., "The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval", Communications in Statistics-Theory and Methods, 48(14): 3423-3438, (2019). DOI: 10.1080/03610926.2018.1476717

[4] Guerra, R. R., Pena-Ramirez, F. A., and Bourguignon, M., "The unit extended Weibull families of distributions and its applications", Journal of Applied Statistics, 48(16): 3174-3192, (2021). DOI: 10.1080/02664763.2020.1796936

[5] Korkmaz, M.Ç., Leiva, V., and Martin-Barreiro, C., "The continuous Bernoulli distribution: Mathematical characterization, fractile regression, computational simulations, and applications" Fractal and Fractional, 7(5): 386, (2023). DOI: https://doi.org/10.3390/fractalfract7050386

[6] Korkmaz, M. Ç., Altun, E., Alizadeh, M., and El-Morshedy, M., "The log exponential-power distribution: Properties, estimations and quantile regression model", Mathematics, 9(21): 2634, (2021). DOI: https://doi.org/10.3390/math9212634

[7] Korkmaz, M. Ç., Chesneau, C., and Korkmaz, Z. S. "The unit folded normal distribution: A new unit probability distribution with the estimation procedures, quantile regression modeling and educational attainment applications", Journal of Reliability and Statistical Studies, 261-298, (2022). DOI: 10.13052/jrss0974-8024.15111

[8] Maya, R., Jodra, P., Irshad, M. R., and Krishna, A.," The unit Muth distribution: Statistical properties and applications", Ricerche di Matematica, 1-24, (2022). DOI: https://doi.org/10.1007/s11587-022-00703-7

[9] Mazucheli J, Menezes A.F., and Dey S.," The unit-Birnbaum-Saunders distribution with applications", Chilean Journal of Statistics, 9(1): 47-57, (2018).

[10] Mazucheli, J., Alves, B., Korkmaz, M. Ç., and Leiva, V., "Vasicek quantile and mean regression models for bounded data: New formulation, mathematical derivations, and numerical applications", Mathematics, 10(9): 1389, (2022). DOI: https://doi.org/10.3390/ math10091389

[11] Mazucheli, J., Korkmaz, M. Ç., Menezes, A. F., and Leiva, V., "The unit generalized half-normal quantile regression model: formulation, estimation, diagnostics, and numerical applications", Soft Computing, 27(1): 279-295, (2023). DOI: https://doi.org/10.1007/s00500-022-07278-3

[12] Mazucheli, J., Alves, B., and Korkmaz, M. Ç., "The Unit-Gompertz Quantile Regression Model for the Bounded Responses", Mathematica Slovaca, 73(4): 1039-1054, (2023). DOI: https://doi.org/10.1515/ms-2023-0077

[13] Altun, E., and Cordeiro, G. M., "The unit-improved second-degree Lindley distribution: inference and regression modeling", Computational Statistics, 35: 259-279, (2020). DOI: https://doi.org/10.1007/s00180-019-00921-y

[14] Korkmaz, M. C¸., Chesneau, C., and Korkmaz, Z. S., "Transmuted unit Rayleigh quantile regression model: Alternative to beta and Kumaraswamy quantile regression models", University Politehnica of Bucharest Scientific Bulletin-Series A-Applied Mathematics and Physics, 83: 149-158, (2021).

[15 Ribeiro, T.F., Cordeiro, G.M., Pena-Ramirez, F.A., and Guerra, R.R., "A new quantile regression for the COVID-19 mortality rates in the United States", Computational and Applied Mathematics, 40: 1-16, (2021). DOI: https://doi.org/10.1007/s40314-021-01553-z

[16] Abdi, M., Asgharzadeh, A., Bakouch, H. S., and Alipour, Z., "A new compound gamma and Lindley distribution with application to failure data", Austrian Journal of Statistics, 48(3): 54-75, (2019). DOI: https://doi.org/10.17713/ajs.v48i3.843

[17] Canuto, C., Hussaini, M. Y., Quarteroni, A., and Zang, T. A., "Spectral methods: evolution to complex geometries and applications to fluid dynamics", Springer Science and Business Media, (2007).

[18] Bonferroni, C., Elmenti di statistica generale [elements of general statistics]. Firenze: Libreria Seber, (1930).

[19] Casella, G., Robert, C. P., and Wells, M. T., "Generalized accept-reject sampling schemes", Lecture notes-monograph series, 342-347, (2004).

[20] Amigó, J. M., Balogh, S. G., and Hernández, S., "A brief review of generalized entropies", Entropy, 20(11): 813, (2018). DOI: https://doi.org/10.3390/e20110813

[21] Kumaraswamy, P., "A generalized probability density function for double bounded random processes", Journal of Hydrology, 46(1-2): 79-88, (1980). DOI: https://doi.org/10.1016/0022-1694(80)90036-0

[22] Korkmaz, M. C¸., and Chesneau, C., "On the unit Burr-XII distribution with the quantile regression modeling and applications", Computational and Applied Mathematics, 40(1): 29, (2021). DOI: https://doi.org/10.1007/s40314-021-01418-5

[23] Abd El-Bar, A., Bakouch, H. S., and Chowdhury, S., "A new trigonometric distribution with bounded support and an application", Revista de la Union Matematica Argentina, 62(2): 459-473, (2021). DOI: https://doi.org/10.33044/revuma.1872

[24] Gomez-Deniz, E., Sordo, M. A., and Calderin-Ojeda, E., “The Log–Lindley distribution as an alternative to the beta regression model with applications in insurance”, Insurance: Mathematics and Economics, 54: 49-57, (2014). DOI: https://doi.org/10.1016/j.insmatheco.2013.10.017

[25] Jodra, P., and Jimenez-Gamero, M. D., “A quantile regression model for bounded responses based on the exponential-geometric distribution”, REVSTAT-Statistical Journal, 18(4): 415-436, (2020).