

# UNIT BURR-HATKE DISTRIBUTION WITH A NEW QUANTILE REGRESSION MODEL

SULE SAĞLAM<sup>1</sup>, KADİR KARAKAYA<sup>1</sup>

---

*Manuscript received: 15.03.2022; Accepted paper: 29.08.2022;*

*Published online: 30.09.2022.*

**Abstract.** In this study, a new distribution is introduced. The Burr-Hatke distribution is considered the baseline distribution. Since the baseline distribution has one parameter, the new unit distribution also has one parameter. Some distributional properties such as moments, coefficients of skewness and kurtosis, stochastic ordering, etc. of the new distribution are studied. Five estimators such as maximum likelihood, least squares, weighted least squares, Anderson-Darling, and Cramer-von Mises are examined to estimate the unknown parameter of the new model. The performances of the estimators are analyzed according to the bias and mean square error criteria calculated by Monte Carlo simulation. Two numerical data analyses are performed. A new quantile regression model is also introduced based on the new distribution as an alternative to beta and Kumaraswamy regression. A Monte Carlo simulation is also conducted for the new regression model.

**Keywords:** Unit distribution; Burr-Hatke distribution; maximum likelihood estimation; Monte Carlo simulation; quantile regression.

## 1. INTRODUCTION

Statistical distributions are very important tools used in modeling real data related to physical, medical, biological, technological, etc. Modeling real data will provide us with extensive information about the population of the relevant phenomenon. Therefore, every distribution which available or will be introduced has vital importance. In the last two decades, many distributions and distribution families have been derived to serve modeling purposes. Some of these distributions can be given as [1-4].

There are few distributions with domain on (0,1) in the literature that model proportional data such as mortality rate, recovery rate, rate of education measures, etc. The well-known unit distributions are beta and Kumaraswamy distributions. Beta and Kumaraswamy distributions can be insufficient for modeling such proportional data. This situation led to the need to propose new unit distributions. Therefore, current lifetime distributions are transformed to the range (0,1), i.e. are united. Some unit distributions studied in recent years can be given as follows [5-10]. In this study, the Burr-Hatke distribution is united. Some studies on the Burr-Hatke distribution can be given as [11-13].

Beta regression [14] is a popular model for explaining the response variable, which has a range (0,1). As an alternative to beta regression, new quantile-based regression models are constructed. Some of these regression models can be reported as [8,9, 15-17]. In this paper, a new quantile regression model is introduced based on the new model. The rest of the paper is organized as follows: In Section 2, a new model is introduced and some mathematical

---

<sup>1</sup> Selçuk University, Department of Statistics, 42100 Konya, Turkey. E-mail: [sulesaglam75@gmail.com](mailto:sulesaglam75@gmail.com); [kkarakaya@selcuk.edu.tr](mailto:kkarakaya@selcuk.edu.tr).

properties are studied. The various estimation methods for estimating one parameter of the new distribution are examined in Section 3. In Section 4, extensive Monte Carlo simulations are conducted to compare the performance of the estimation methods. Two numerical analyses are studied based on simulated data in Section 5. In Section 6, a new quantile regression model is introduced and a simulation experiment is conducted based on the maximum likelihood method. The article ends with the conclusion part in Section 7.

## 2. UNIT BURR-HATKE DISTRIBUTION

Let  $Y$  random variable follow the Burr-Hatke distribution introduced by [18] with cumulative distribution function (cdf) and probability density function (pdf) given, respectively, by,

$$F_{BH}(y; \beta) = 1 - \frac{\exp(-\beta y)}{y+1}, \quad y > 0, \quad (1)$$

and

$$f_{BH}(y; \beta) = \frac{1 + \beta(y+1)}{(y+1)^2} \exp(-\beta y), \quad y > 0, \quad (2)$$

where  $\beta$  is the scale parameter. If the random variable  $Y$  has the pdf given in Equation (2), then the cdf and pdf of the random variable  $X = \exp(-Y)$  are obtained, respectively, by,

$$F_{UBH}(x; \beta) = \frac{x^\beta}{1 - \log(x)} \quad (3)$$

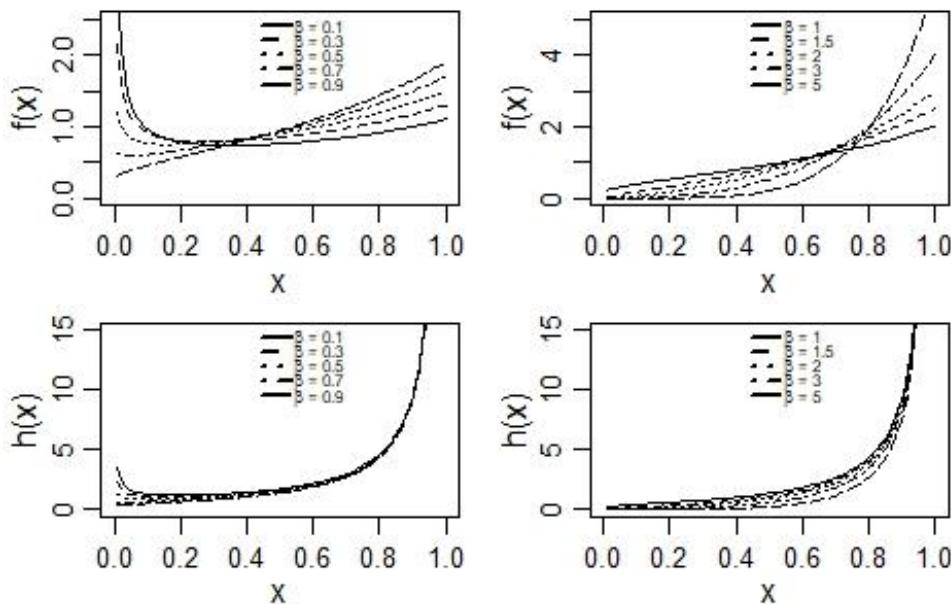
and

$$f_{UBH}(x; \beta) = \frac{x^{\beta-1} \{1 - \beta(\log(x) - 1)\}}{\{1 - \log(x)\}^2}, \quad (4)$$

where  $0 < x < 1$  and  $\beta > 0$  is the model parameter. The new distribution is called as unit Burr-Hatke (UBH) distribution and is denoted by  $UBH(\beta)$ . The hazard rate function (hrf) of the UBH distribution is obtained as

$$h_{UBH}(x; \beta) = \frac{x^{\beta-1} \{1 - \beta(\log(x) - 1)\}}{\{1 - \log(x)\} \{1 - \log(x) - x^\beta\}}. \quad (5)$$

The plots of the pdf and hrf are given for some choices of  $\beta$  in Fig. 1. From Fig. 1, one can see that the pdf of the UBH distribution is decreasing and increasing, and the hrf of the UBH distribution is increasing and bath-tube shaped.

Figure 1. The pdf and hrf plots for some choices of  $\beta$ 

## 2.1. MOMENTS

The  $r$ th moment of the UBH distribution is obtained as

$$\begin{aligned}
 E(X^r) &= \int_0^1 x^r f(x) dx \\
 &= \int_0^1 \frac{x^{r+\beta-1} \{1 - \beta(\log(x) - 1)\}}{\{1 - \log(x)\}^2} \\
 &= 1 - r \exp(\beta + r) E_i(1, \beta + r),
 \end{aligned} \tag{6}$$

where  $E_i$  is called exponential integrals and given as follows:

$$E_i(a, z) = z^{a-1} \Gamma(1-a, z). \tag{7}$$

The  $E_i$  function given in Equation (7) is easily calculated via the Maple program. Also, an R function that calculated the moments and coefficients of skewness (S), and kurtosis (K) of the UBH distribution according to the parameter of  $\beta$  is presented in the Appendix. By taking  $r = 1, 2, 3$  and  $4$  in the Equation (6), the first four moments can be obtained as follows, respectively,

$$E(X) = 1 - \exp(\beta + 1) E_i(1, \beta + 1), \tag{8}$$

$$E(X^2) = 1 - 2 \exp(\beta + 2) E_i(1, \beta + 2), \tag{9}$$

$$E(X^3) = 1 - 3\exp(\beta + 3)E_i(1, \beta + 3), \quad (10)$$

and

$$E(X^4) = 1 - 4\exp(\beta + 4)E_i(1, \beta + 4). \quad (11)$$

Using Equations (8) and (9) the variance of the UBH distribution is obtained as

$$\text{Var}(X) = 1 - 2\exp(\beta + 2)E_i(1, \beta + 2) - \{1 - \exp(\beta + 1)E_i(1, \beta + 1)\}^2. \quad (12)$$

The S and K for UBH distribution can be obtained by, respectively, using Equations (8)-(12) as follows:

$$S = \frac{E(X^3) - 3E(X)E(X^2) + 2\{E(X)\}^3}{\{\text{Var}(X)\}^{3/2}} \quad (13)$$

and

$$K = \frac{E(X^4) - 4E(X)E(X^3) + 6E(X^2)\{E(X)\}^2 - 3\{E(X)\}^4}{\{\text{Var}(X)\}^2}. \quad (14)$$

The expected value, variance, S, and K are reported in Table 1 for some values of  $\beta$ . It is concluded from Table 1 that the expected value increases as  $\beta$  increases. It is also seen that the variance, S and K decrease as  $\beta$  increases. According to the S, the distribution is skewed to right for small values of  $\beta$ , and skewed to left for large values of  $\beta$ . It is seen that the UBH distribution is platykurtic in all parameter cases according to the K values.

**Table 1. Expected value, variance, S and K for some choices of parameter**

$\beta$	$E(X)$	$\text{Var}(X)$	$S$	$K$
0.1	0.4413	0.1093	0.1089	-6.2404
0.3	0.5029	0.0986	-0.1025	-20.2122
0.5	0.5517	0.0885	-0.2640	-46.3242
0.7	0.5913	0.0795	-0.3936	-89.4778
0.9	0.6242	0.0715	-0.5010	-155.1814
1	0.6387	0.0679	-0.5484	-198.3769
1.5	0.6965	0.0533	-0.7401	-551.7593
2	0.7379	0.0428	-0.8812	-1223.1180
3	0.7937	0.0293	-1.0789	-4102.0320
5	0.8547	0.0161	-1.3112	-21359.8300

The quantile function of the UBH distribution is given as

$$Q(u; \beta) = \exp\left(-\frac{\text{LambertW}\left(\frac{\beta \exp(\beta)}{u}\right) - \beta}{\beta}\right), u \in (0, 1), \quad (15)$$

where  $LambertW$  is Lambert W function and calculated as  $LambertW(x) \times \exp(LambertW(x)) = x$ . If  $u=0.25, 0.5$  and  $0.75$  are chosen in Equation (15), the 1st quartile, median and 3rd quartile are obtained, respectively. Also, Equation (15) can be used to generate random numbers from the UBH distribution.

## 2.2. STOCHASTIC ORDERING

The stochastic order quantifies the concept that one random variable is bigger than another. In this subsection, different types of stochastic orders for the UBH distribution were analyzed. The following theorem illustrates that the UBH random variables can be ordered concerning the likelihood ratio.

**Theorem 1.** Let  $X \sim UBH(\beta_1)$  and  $Y \sim UBH(\beta_2)$ . If  $\beta_1 < \beta_2$  then  $X$  is smaller than  $Y$  in the likelihood ratio order.

*Proof:* For any  $x \in (0,1)$ , the derivate of the logarithm of the ratio of densities is given by

$$\frac{d \log(g(x))}{dx} = \frac{\overbrace{(\log(x)-1)}^{<0} \overbrace{\{\beta_1 \beta_2 \log(x) - \beta_1(1+\beta_2) - \beta_2\}}^{<0} (\beta_1 - \beta_2)}{\underbrace{x(\beta_2 \log(x) - 1 - \beta_2)}_{<0} \underbrace{(\beta_1 \log(x) - 1 - \beta_1)}_{<0}} < 0 \quad (16)$$

for  $\beta_1 < \beta_2$  and the proof is completed.

**Corollary 1.** It is also concluded from [19] that the random variable  $X$  is also smaller than  $Y$  in the stochastic orders, hazard ratio, and mean residual life when the  $\beta_1 < \beta_2$ .

## 2.3. ORDER STATISTICS

Let  $X_1, X_2, \dots, X_n$  be random sample from UBH distribution, and  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  represent the corresponding order statistics. Then, the pdf of the  $r$ th order statistic,  $X_{(r)}$ , is obtained by

$$f(x_{(r)}, \beta) = \frac{x^{\beta-1} \{1 - \beta(\log(x)-1)\}}{\{1-\log(x)\}^2 B(r, n-r+1)} \sum_{i=0}^{n-r} (-1)^i \binom{n-r}{i} \left\{ \frac{x^\beta}{1-\log(x)} \right\}^{r+i-1}, \quad (17)$$

where  $n$  is the sample of size,  $B(\cdot, \cdot)$  is the beta function and  $r = 1, 2, \dots, n$ . It is clear that for  $r=1$  and  $r=n$ , the pdf of the  $X_{(1)} = \min(X_1, X_2, \dots, X_n)$  and  $X_{(n)} = \max(X_1, X_2, \dots, X_n)$  are obtained, respectively.

### 3. POINT ESTIMATION

In this section, five estimators such as maximum likelihood, least squares, weighted least squares, Anderson-Darling, and Cramer-von Mises are studied to estimate the unknown parameter  $\beta$  of UBH distribution. Let  $X_1, X_2, \dots, X_n$  be a random sample from the  $UBH(\beta)$  distribution and  $x_1, x_2, \dots, x_n$  is observed values of the sample. Let  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  be the order statistics based on sample  $X_1, X_2, \dots, X_n$  with the realization  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ . Then, the likelihood and log-likelihood function can be written by

$$L(\beta) = \prod_{i=1}^n \frac{x_i^{\beta-1} \{1 - \beta(\log(x_i) - 1)\}}{\{1 - \log(x_i)\}^2} \quad (18)$$

and

$$\ell(\beta) = (\beta - 1) \sum_{i=1}^n \log(x_i) + \sum_{i=1}^n \log\{1 - \beta(\log(x_i) - 1)\} - 2 \sum_{i=1}^n \log\{1 - \log(x_i)\} \quad (19)$$

The maximum likelihood estimates (MLEs) of  $\beta$  is given by

$$\hat{\beta} = \arg \max_{\beta} \ell(\beta). \quad (20)$$

Let us tackle the following four functions to obtain the other estimators:

$$LS(\beta) = \sum_{i=1}^n \left( \frac{x_{(i)}^{\beta}}{1 - \log(x_{(i)})} - \frac{i}{n+1} \right)^2, \quad (21)$$

$$WLS(\beta) = \sum_{i=1}^n \frac{(n+2)(n+1)^2}{i(n-i+1)} \left( \frac{x_{(i)}^{\beta}}{1 - \log(x_{(i)})} - \frac{i}{n+1} \right)^2, \quad (22)$$

$$AD(\beta) = -n - \sum_{i=1}^n \frac{2i-1}{n} \left[ \log \left\{ \frac{x_{(i)}^{\beta}}{1 - \log(x_{(i)})} \right\} + \log \left\{ 1 - \frac{x_{(n+i-1)}^{\beta}}{1 - \log(x_{(n+i-1)})} \right\} \right] \quad (23)$$

and

$$CvM(\beta) = \frac{1}{12n} + \sum_{i=1}^n \left[ \frac{x_{(i)}^{\beta}}{1 - \log(x_{(i)})} - \frac{2i-1}{2n} \right]^2. \quad (24)$$

The least-square estimates (LSEs), weighted least square estimate (WLSEs), Anderson-Darling estimate (ADEs) and Cramer-von Mises estimate (CvMEs) are obtained by minimizing Equations (21)-(24), respectively. One can find many research that used these estimators such as [20] and [21].

#### 4. SIMULATION STUDY

In the simulation study, the bias and mean square errors (MSEs) of MLEs, LSEs, WLSEs, ADEs and CvMEs for the new model parameter are simulated with 5000 trials. The results are presented in Table 2 for the true parameter vectors  $\beta = 0.5, 0.9, 1, 1.5, 2.5, 3$  and 5. The sample sizes are selected as  $n = 25, 50, 100, 150, 175$  and 200. It is concluded from Table 2 that the bias and MSEs of all estimators decrease to zero when the sample size increases. In the small sample cases, ADEs is the best method according to the bias criterion and MLEs is the best according to MSEs criterion. Although the bias and MSEs values of all estimators converge as the sample size increases, it is still the best ADEs in bias and the best MLEs in MSEs.

**Table 2. Average biases and MSEs of all estimates.**

n	$\beta$	Bias					MSEs				
		MLEs	LSEs	WLSEs	ADEs	CvMEs	MLEs	LSEs	WLSEs	ADEs	CvMEs
25	0.5	0.0765	0.0591	0.0556	0.0497	0.0718	0.0669	0.1085	0.0950	0.0819	0.1111
		0.0368	0.0248	0.0245	0.0216	0.0312	0.0276	0.0438	0.0374	0.0346	0.0444
		0.0159	0.0092	0.0099	0.0080	0.0125	0.0123	0.0196	0.0165	0.0159	0.0197
		0.0120	0.0074	0.0080	0.0068	0.0096	0.0077	0.0130	0.0108	0.0105	0.0130
		0.0088	0.0053	0.0057	0.0043	0.0071	0.0067	0.0114	0.0094	0.0091	0.0114
		0.0068	0.0031	0.0036	0.0024	0.0047	0.0055	0.0096	0.0079	0.0077	0.0097
50	0.9	0.0794	0.0554	0.0510	0.0449	0.0708	0.1173	0.1901	0.1680	0.1445	0.1934
		0.0413	0.0275	0.0262	0.0239	0.0354	0.0573	0.0840	0.0736	0.0695	0.0849
		0.0179	0.0119	0.0121	0.0098	0.0158	0.0250	0.0376	0.0323	0.0313	0.0378
		0.0111	0.0065	0.0065	0.0052	0.0092	0.0162	0.0247	0.0211	0.0207	0.0248
		0.0140	0.0093	0.0098	0.0085	0.0116	0.0146	0.0219	0.0188	0.0185	0.0220
		0.0098	0.0072	0.0075	0.0064	0.0091	0.0118	0.0178	0.0152	0.0150	0.0178
100	1	0.0994	0.0709	0.0668	0.0618	0.0872	0.1498	0.2212	0.1977	0.1749	0.2258
		0.0421	0.0248	0.0248	0.0218	0.0330	0.0630	0.0937	0.0821	0.0774	0.0947
		0.0266	0.0178	0.0185	0.0165	0.0220	0.0303	0.0449	0.0390	0.0379	0.0451
		0.0125	0.0060	0.0071	0.0056	0.0088	0.0188	0.0284	0.0245	0.0240	0.0285
		0.0137	0.0113	0.0110	0.0098	0.0136	0.0167	0.0243	0.0210	0.0206	0.0244
		0.0144	0.0103	0.0108	0.0094	0.0124	0.0139	0.0211	0.0181	0.0178	0.0212
150	1.5	0.0992	0.0684	0.0619	0.0555	0.0878	0.2425	0.3436	0.3073	0.2771	0.3495
		0.0502	0.0320	0.0306	0.0281	0.0419	0.1083	0.1555	0.1374	0.1303	0.1571
		0.0254	0.0151	0.0155	0.0137	0.0201	0.0520	0.0733	0.0645	0.0630	0.0737
		0.0186	0.0113	0.0123	0.0107	0.0147	0.0337	0.0481	0.0419	0.0413	0.0482
		0.0115	0.0039	0.0054	0.0041	0.0068	0.0291	0.0410	0.0361	0.0357	0.0411
		0.0143	0.0136	0.0129	0.0116	0.0161	0.0253	0.0362	0.0316	0.0312	0.0364
175	2.5	0.1674	0.1115	0.1041	0.1005	0.1381	0.5634	0.7407	0.6735	0.6196	0.7543
		0.0928	0.0611	0.0600	0.0567	0.0746	0.2435	0.3203	0.2865	0.2772	0.3238
		0.0410	0.0233	0.0237	0.0216	0.0301	0.1170	0.1611	0.1428	0.1399	0.1619
		0.0227	0.0166	0.0163	0.0144	0.0211	0.0752	0.1027	0.0906	0.0894	0.1030
		0.0196	0.0127	0.0131	0.0112	0.0166	0.0632	0.0860	0.0765	0.0753	0.0862
		0.0176	0.0151	0.0144	0.0129	0.0185	0.0564	0.0761	0.0674	0.0667	0.0763
200	3	0.1855	0.1452	0.1331	0.1252	0.1755	0.7618	1.0344	0.9421	0.8481	1.0519
		0.0954	0.0626	0.0589	0.0566	0.0779	0.3280	0.4207	0.3798	0.3665	0.4251
		0.0482	0.0325	0.0328	0.0304	0.0402	0.1505	0.2041	0.1819	0.1775	0.2053
		0.0234	0.0152	0.0159	0.0137	0.0204	0.0997	0.1382	0.1225	0.1207	0.1387
		0.0290	0.0229	0.0227	0.0208	0.0273	0.0871	0.1175	0.1047	0.1035	0.1179
		0.0211	0.0154	0.0150	0.0133	0.0192	0.0735	0.1022	0.0901	0.0892	0.1025
25	5	0.2500	0.1878	0.1688	0.1636	0.2325	1.7076	2.1918	2.0018	1.8621	2.2309
		0.1142	0.0790	0.0736	0.0689	0.1015	0.7363	0.9726	0.8838	0.8501	0.9821
		0.0667	0.0422	0.0427	0.0398	0.0536	0.3618	0.4695	0.4202	0.4145	0.4719
		0.0435	0.0288	0.0300	0.0263	0.0364	0.2353	0.3215	0.2856	0.2817	0.3226
		0.0441	0.0327	0.0324	0.0295	0.0392	0.2040	0.2713	0.2418	0.2392	0.2722
		0.0275	0.0157	0.0172	0.0145	0.0214	0.1793	0.2395	0.2133	0.2111	0.2400

## 5. NUMERICAL DATA ANALYSIS

This section provides two numerical applications for illustrating the modeling ability of the UBH distribution. We compare the fits of the UBH distribution with Kumaraswamy (KM), Unit Weibull (UW) [10], Beta(B) and Log-Lindley (LL) [22] distributions since they are used for modeling bounded data. The pdf of the KM, UW, B and LL distributions given, respectively, by

- **Kumaraswamy Distribution:**

$$f(y; \alpha, \beta) = \alpha\beta y^{(\alpha-1)}(1-y^\alpha)^{(\beta-1)}$$

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

- **Unit Weibull Distribution:**

$$f(y; \alpha, \beta) = \frac{\alpha\beta(-\log(y))^{(\beta-1)} \exp(-\alpha(-\log(y)))^\beta}{y}$$

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

- **Beta Distribution:**

$$f(y; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1}$$

where  $y \in (0,1)$  and  $\alpha > 0, \beta > 0$  and  $B(\alpha, \beta)$  denotes the classical beta function.

- **Log-Lindley Distribution:**

$$f(y; \alpha, \beta) = \frac{\alpha^2(\beta-\log(y))y^{(\alpha-1)}}{1+\beta\alpha}$$

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

The MLEs of parameters, log-likelihood value  $(\hat{\ell})$ , Akaike's information criteria (AIC), Bayesian information criterion (BIC), consistent AIC (CAIC), Hannan-Quinn information criterion (HQIC), p-values of the Kolmogorov-Smirnov statistics (KS pval), Anderson-Darling statistics (AD pval), Cramer von Mises statistic (CvM pval) are presented in Tables 3 and 4. The best model have smaller values of the AIC, BIC, CAIC and HQIC, and the larger values of  $\hat{\ell}$ . Tables 3 and 4 indicate that the UBH distribution is the best model for all criteria for both simulated data sets.

### 5.1. NUMERICAL DATA ANALYSIS I

The first example is generated from the UBH distribution. The data consist of 50 samples. The data are: 0.74240, 0.94342, 0.39573, 0.44630, 0.00008, 0.92311, 0.91661, 0.33440, 0.21177, 0.85058, 0.26264, 0.84503, 0.04222, 0.28067, 0.70795, 0.68991, 0.20695, 0.29767, 0.32418, 0.41980, 0.95322, 0.41824, 0.38850, 0.59224, 0.18536, 0.86103, 0.63010, 0.70765, 0.76696, 0.81375, 0.39155, 0.73277, 0.14208, 0.22898, 0.52555, 0.48655, 0.42048, 0.31397, 0.27340, 0.67456, 0.42820, 0.00034, 0.62733, 0.06998, 0.61898, 0.91235, 0.24682, 0.76829, 0.10191, 0.75293.

**Table 3. Data analysis results for numerical analysis I**

Model	$\hat{\ell}$	AIC	BIC	CAIC	HQIC	KS pval	AD pval	CvM pval
UBH	2.2555	-2.5110	-0.5990	-2.4277	-1.7829	0.7072	0.4623	0.5031
KM	0.5903	2.8193	6.6434	3.0747	4.2756	0.4619	0.4145	0.4900
UW	1.1859	1.6280	5.4520	1.8833	3.0842	0.5652	0.4996	0.6206
B	0.5812	2.8375	6.6616	3.0929	4.2938	0.4663	0.4120	0.4858
LL	0.5699	2.8601	6.6842	3.1154	4.3163	0.4679	0.3953	0.4611

### 5.2. NUMERICAL DATA ANALYSIS II

The second example is generated from the UBH distribution. The data consist of 20 samples. The data are: 0.39588, 0.75345, 0.09370, 0.46893, 0.06053, 0.83617, 0.97784, 0.64811, 0.37029, 0.68505, 0.32765, 0.25660, 0.62508, 0.80431, 0.66203, 0.00217, 0.82395, 0.90785, 0.33660, 0.62289.

**Table 4. Data analysis results for numerical analysis II**

Model	$\hat{\ell}$	AIC	BIC	CAIC	HQIC	KS pval	AD pval	CvM pval
UBH	0.7469	0.5061	1.5018	0.7283	0.70051	0.9428	0.9579	0.9241
KM	0.0880	3.8238	5.8152	4.5297	4.2125	0.6041	0.8200	0.7370
UW	0.3238	3.3524	5.3438	4.0582	3.7411	0.8019	0.9028	0.8555
B	0.0822	3.8354	5.8269	4.5413	4.2242	0.5994	0.8195	0.7377
LL	0.0123	3.9753	5.9668	4.6812	4.3641	0.4600	0.7344	0.6355

## 6. A NEW QUANTILE REGRESSION

In this section, a new quantile regression model is introduced based on the UBH distribution. The quantile function given in Equation (15) is used to construct the quantile regression. If  $\beta$  is taken from the  $Q(u; \beta) = \mu$ , then

$$\beta = \frac{\log\left(\frac{1}{u(1-\log(\mu))}\right)}{-\log(\mu)} \quad (25)$$

is obtained. The cdf and pdf in Equations (3) and (4) are re-parameterized, respectively, by

$$G(z; \mu) = \frac{z^{\frac{\log\left(\frac{1}{u(1-\log(\mu))}\right)}{-\log(\mu)}}}{1 - \log(z)} \quad (26)$$

and

$$g(z; \mu) = \frac{z^{\frac{\log\left(\frac{1}{u(1-\log(\mu))}\right)}{-\log(\mu)} - 1} \left\{ 1 + \left[ \frac{\log\left(\frac{1}{u(1-\log(\mu))}\right)}{\log(\mu)} \right] (\log(z) - 1) \right\}}{\{1 - \log(z)\}^2}, \quad (27)$$

where  $\mu \in (0,1)$  represents the quantile parameter, and  $u \in (0,1)$  is given or chosen as 0.5. It is noted that the random variable  $Z$  is denoted by  $Z \sim QUBH(\mu, u)$ . After the QUBH distribution is defined, the step of establishing a new regression model can be started based on the QUBH with the pdf given in Equation (27). Let  $z_1, z_2, \dots, z_n$  such that  $z_i$  is a realization of  $Z \sim QUBH(\mu_i, u)$  for  $i = 1, 2, \dots, n$ , where  $\mu_i$  is unknown parameters, and the  $u$  is known or given. As a result, the proposed quantile regression model is given as follows:

$$g(\mu_i) = \mathbf{x}_i \boldsymbol{\theta}^T \quad (28)$$

where  $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_p)^T$ ,  $\mathbf{x}_i = (\mathbf{1}, x_{i1}, \dots, x_{ip})$  are the unknown regression model parameter vector and the function  $g$  is a link function. In the simulation study,  $u = 0.5$  is evaluated which the covariates are linked to the conditional median of the response variable. We use the following logit-link function because the QUBH distribution has a domain (0,1) interval:

$$g(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right) = \mathbf{x}_i \boldsymbol{\theta}^T, \quad i = 1, 2, \dots, n. \quad (29)$$

### 6.1. ESTIMATION FOR REGRESSION MODEL PARAMETER

In this subsection, the maximum likelihood estimation procedure is discussed for the estimation of the unknown regression parameters. Using Equation (29), it is easily obtained

$$\mu_i = \frac{\exp(\mathbf{x}_i \boldsymbol{\theta}^T)}{1 + \exp(\mathbf{x}_i \boldsymbol{\theta}^T)}. \quad (30)$$

Let  $Z_1, Z_2, \dots, Z_n$  be a random sample of size  $n$  from the  $QUBH(\mu_i, u)$  distribution with realizations  $z_1, z_2, \dots, z_n$ , where the  $\mu_i$  is given in Equation (30) for  $i = 1, 2, \dots, n$ . Then the log-likelihood function is given by

$$\ell(\boldsymbol{\theta}) = \left( \frac{\log\left(\frac{1}{u(1-\log(\mu_i))}\right)}{-\log(\mu_i)} - 1 \right) \sum_{i=1}^n \log(z_i) + \sum_{i=1}^n \log \left\{ 1 + \frac{\log\left(\frac{1}{u(1-\log(\mu_i))}\right)}{\log(\mu_i)} (\log(z_i) - 1) \right\} - 2 \sum_{i=1}^n \log\{1 - \log(z_i)\}, \quad (31)$$

where  $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_p)$  is the unknown parameter vector. The MLEs of  $\boldsymbol{\theta}$ , say  $\hat{\boldsymbol{\theta}} = (\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_p)$ , is achieved by maximizing the  $\ell(\boldsymbol{\theta})$  given in Equation (31) with respect to  $\theta_0, \theta_1, \dots, \theta_p$ . The **optim** function in R can be used to maximize the  $\ell(\boldsymbol{\theta})$  in Equation (31).

Based on some regularity conditions, the asymptotic distribution of  $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$  is multivariate normal  $N_{p+1}(0, J^{-1})$ , where  $J$  is the expected information matrix. It is usually used instead of  $J$  with the observed information matrix in practice. Any software can calculate the observed information matrix. The asymptotic standard errors of estimates based on the observed Fisher information matrix can be calculated in R using the **optim** function.

## 6. 2. QUANTILE REGRESSION SIMULATION EXPERIMENTS

A simulation study is carried out in this subsection to observe the MLEs behaviors of the parameters of the QUBH regression model. The simulation is conducted with  $u = 0.5$ ,  $n = 25, 50, 75, 100, 250, 500$ . It is considered the case where the number of covariates are one, two, or three and the covariates are generated independently from the normal and binomial distribution. The bias and MSEs of the MLEs of the QUBH model and the coverage probabilities (CPs) and mean lengths (MLs) of the MLEs based confidence intervals are given in Tables 5-7. From Tables 5-7, it is concluded that the  $\hat{\boldsymbol{\theta}}$  are unbiased and MSEs of these estimates decrease to 0 for a large sample of size and it indicates that the MLEs are also consistent. The approximate confidence intervals for all model parameters work quite well. The CPs for all model parameters are almost equal to the nominal level 0.95 and the MLs are decrease to zero when the sample of size is increased.

**Table 5. Bias and MSEs of the MLEs of the QUBH regression model and the CPs and MLs of the MLEs based approximately confidence intervals for three covariates case**

$\theta$	Covar iates	n	$\theta_0$				$\theta_1$				$\theta_2$				$\theta_3$			
			Bias	MSE s	CP	ML	Bias	MSE s	CP	ML	Bias	MSE s	CP	ML	Bias	MSE s	CP	ML
$\theta = (0.8, 0.1, 0.4, 0.1)$	Normal (0,1)	25	0.11 18	0.07 67	0.91 42	0.97 57	- 0.00 11	0.08 68	0.94 66	1.10 34	- 0.01 21	0.08 85	0.94 36	1.09 91	- 0.00 77	0.08 88	0.94 86	1.10 48
		50	0.07 02	0.03 27	0.92 74	0.65 26	- 0.00 17	0.03 30	0.95 02	0.70 36	- 0.01 70	0.03 28	0.95 22	0.70 13	- 0.00 64	0.03 21	0.95 12	0.70 45
		75	0.04 51	0.02 04	0.93 30	0.52 55	- 0.00 61	0.02 05	0.94 38	0.55 35	- 0.01 11	0.01 99	0.94 96	0.54 83	- 0.00 31	0.01 98	0.95 16	0.55 66
		100	0.03 42	0.01 50	0.93 00	0.45 29	- 0.00 01	0.01 48	0.94 40	0.47 35	- 0.01 10	0.01 40	0.95 02	0.46 71	- 0.00 23	0.01 46	0.95 04	0.47 40
		250	0.01 37	0.00 54	0.94 82	0.28 40	0.00 08	0.00 55	0.95 20	0.28 99	- 0.00 85	0.00 52	0.94 78	0.28 37	0.00 06	0.00 54	0.94 76	0.28 91
		500	0.00 71	0.00 27	0.94 58	0.20 02	- 0.00 14	0.00 27	0.95 16	0.20 22	- 0.00 29	0.00 25	0.95 20	0.19 68	- 0.00 10	0.00 28	0.94 74	0.20 21
	Binom (5,0.5)	25	0.08 29	1.05 87	0.94 82	3.87 64	0.00 29	0.05 10	0.95 52	0.86 63	0.00 07	0.05 22	0.94 46	0.87 42	- 0.00 17	0.05 24	0.94 72	0.86 22
		50	0.06 33	0.43 62	0.94 92	2.54 97	- 0.00 26	0.02 15	0.95 06	0.56 51	- 0.00 42	0.02 05	0.95 14	0.56 69	- 0.00 11	0.02 13	0.95 02	0.56 48
		75	0.01 85	0.27 31	0.94 76	2.02 87	0.00 27	0.01 37	0.94 46	0.45 01	- 0.00 23	0.01 36	0.95 26	0.45 14	0.00 21	0.01 33	0.95 28	0.44 94
		100	0.02 23	0.19 10	0.95 26	1.72 88	- 0.00 03	0.00 94	0.95 58	0.38 47	0.00 03	0.00 98	0.94 86	0.38 49	- 0.00 04	0.00 95	0.94 70	0.38 32
		250	0.00 28	0.07 33	0.95 24	1.07 07	0.00 03	0.00 36	0.94 78	0.23 66	0.00 14	0.00 37	0.94 66	0.23 79	0.00 02	0.00 36	0.95 04	0.23 69
		500	0.00 41	0.03 70	0.94 84	0.74 95	0.00 03	0.00 18	0.95 06	0.16 58	0.00 00	0.00 18	0.94 58	0.16 67	- 0.00 05	0.00 18	0.95 04	0.16 60
$\theta = (0.8, 0.1, 0.5, 0.1)$	Normal (0,1)	25	0.12 36	0.07 86	0.91 22	0.96 93	- 0.00 58	0.08 91	0.94 54	1.10 38	- 0.03 30	0.09 00	0.94 38	1.09 66	- 0.01 57	0.09 05	0.94 50	1.10 05
		50	0.07 16	0.03 31	0.92 28	0.64 87	- 0.00 21	0.03 21	0.95 26	0.69 91	- 0.01 72	0.03 27	0.94 64	0.69 64	- 0.00 62	0.03 26	0.94 76	0.70 02
		75	0.04 83	0.02 01	0.93 30	0.52 16	- 0.00 51	0.02 04	0.95 14	0.55 36	- 0.01 36	0.01 86	0.94 94	0.54 50	- 0.00 59	0.01 91	0.95 36	0.55 02
		100	0.03 70	0.01 48	0.92 96	0.44 90	- 0.00 02	0.01 38	0.95 18	0.46 86	- 0.01 53	0.01 37	0.95 34	0.46 16	- 0.00 28	0.01 42	0.95 44	0.46 98
		250	0.01 75	0.00 53	0.94 70	0.28 11	- 0.00 32	0.00 51	0.95 78	0.28 71	- 0.00 81	0.00 51	0.94 58	0.27 89	- 0.00 24	0.00 53	0.94 84	0.28 61
		500	0.00 85	0.00 26	0.94 52	0.19 80	- 0.00 13	0.00 26	0.94 82	0.19 99	- 0.00 46	0.00 23	0.95 74	0.19 28	- 0.00 12	0.00 26	0.94 76	0.19 98
	Binom (5,0.5)	25	0.06 37	1.07 05	0.94 60	3.86 45	0.00 36	0.05 22	0.94 86	0.85 97	- 0.00 02	0.05 36	0.94 48	0.86 26	0.00 31	0.05 18	0.94 24	0.85 89
		50	0.04 21	0.42 78	0.94 84	2.52 41	0.00 13	0.02 09	0.94 74	0.55 92	0.00 11	0.02 14	0.94 92	0.56 57	- 0.00 12	0.02 13	0.95 14	0.56 10
		75	0.02 84	0.26 30	0.95 50	2.00 63	- 0.00 14	0.01 27	0.95 14	0.44 58	0.00 16	0.01 31	0.94 96	0.44 85	0.00 01	0.01 33	0.94 88	0.44 52
		100	0.01 22	0.19 60	0.95 10	1.71 90	0.00 16	0.00 97	0.94 76	0.37 97	0.00 01	0.00 98	0.94 58	0.38 32	0.00 18	0.00 97	0.94 72	0.38 32
		250	0.01 46	0.07 28	0.95 00	1.06 07	- 0.00 05	0.00 36	0.95 10	0.23 45	- 0.00 16	0.00 36	0.95 26	0.23 62	0.00 03	0.00 35	0.95 26	0.23 48
		500	0.00 31	0.03 62	0.95 20	0.74 34	- 0.00 02	0.00 18	0.95 30	0.16 44	0.00 08	0.00 18	0.95 04	0.16 55	- 0.00 02	0.00 17	0.95 58	0.16 44

**Table 6. Bias and MSEs of the MLEs of the QUBH regression model and the CPs and MLs of the MLEs based approximately confidence intervals for two covariates case**

$\theta$	Covariates	n	$\theta_0$				$\theta_1$				$\theta_2$			
			Bias	MSEs	CP	ML	Bias	MSEs	CP	ML	Bias	MSEs	CP	ML
$\Theta = (0.4, 0.3, 0.1)$	Normal (0,1)	25	0.1183	0.0749	0.9112	0.9445	-0.0344	0.0847	0.9386	1.0701	-0.0107	0.0815	0.9402	1.0727
		50	0.0631	0.0320	0.9274	0.6482	-0.0248	0.0298	0.9578	0.6858	-0.0073	0.0322	0.9500	0.7021
		75	0.0429	0.0207	0.9312	0.5258	-0.0198	0.0196	0.9502	0.5430	-0.0061	0.0197	0.9488	0.5525
		100	0.0358	0.0147	0.9420	0.4539	-0.0153	0.0136	0.9510	0.4600	-0.0038	0.0151	0.9500	0.4755
		250	0.0146	0.0054	0.9512	0.2860	-0.0087	0.0050	0.9510	0.2784	-0.0017	0.0055	0.9478	0.2902
		500	0.0076	0.0027	0.9502	0.2020	-0.0061	0.0024	0.9548	0.1935	-0.0007	0.0027	0.9508	0.2030
	Binom (5,0,3)	25	0.0724	0.3743	0.9436	2.3156	0.0041	0.0677	0.9496	0.9892	0.0016	0.0688	0.9510	1.0057
		50	0.0336	0.1562	0.9504	1.5540	0.0028	0.0283	0.9546	0.6559	0.0027	0.0292	0.9496	0.6582
		75	0.0237	0.1004	0.9518	1.2469	0.0006	0.0184	0.9498	0.5235	0.0012	0.0179	0.9538	0.5262
		100	0.0152	0.0753	0.9448	1.0648	0.0013	0.0133	0.9490	0.4472	0.0019	0.0136	0.9454	0.4492
		250	0.0094	0.0290	0.9480	0.6632	-0.0015	0.0050	0.9492	0.2767	0.0011	0.0051	0.9476	0.2780
		500	0.0020	0.0145	0.9500	0.4658	0.0003	0.0025	0.9470	0.1938	0.0009	0.0025	0.9448	0.1948
$\Theta = (0.6, 0.1, 0.3)$	Normal (0,1)	25	0.1006	0.0732	0.9158	0.9468	0.0002	0.0821	0.9494	1.0743	-0.0202	0.0807	0.9456	1.0582
		50	0.0505	0.0310	0.9308	0.6512	-0.0023	0.0319	0.9516	0.6997	-0.0144	0.0303	0.9454	0.6876
		75	0.0363	0.0206	0.9326	0.5279	-0.0056	0.0201	0.9488	0.5572	-0.0163	0.0197	0.9508	0.5486
		100	0.0252	0.0143	0.9428	0.4561	-0.0011	0.0143	0.9478	0.4730	-0.0105	0.0143	0.9484	0.4673
		250	0.0131	0.0055	0.9420	0.2873	-0.0005	0.0056	0.9490	0.2919	-0.0051	0.0052	0.9558	0.2858
		500	0.0069	0.0028	0.9490	0.2031	-0.0009	0.0027	0.9500	0.2042	-0.0029	0.0026	0.9528	0.1999
	Binom (5,0,3)	25	0.0715	0.3767	0.9392	2.2918	0.0041	0.0684	0.9464	0.9853	0.0005	0.0660	0.9460	0.9847
		50	0.0223	0.1549	0.9486	1.5324	0.0053	0.0276	0.9526	0.6485	0.0022	0.0276	0.9556	0.6476
		75	0.0240	0.0976	0.9500	1.2267	0.0013	0.0178	0.9496	0.5171	0.0007	0.0176	0.9516	0.5152
		100	0.0175	0.0720	0.9492	1.0519	0.0018	0.0129	0.9524	0.4426	0.0008	0.0127	0.9510	0.4411
		250	0.0055	0.0276	0.9496	0.6537	0.0007	0.0050	0.9474	0.2739	0.0014	0.0049	0.9554	0.2722
		500	0.0014	0.0142	0.9484	0.4596	0.0012	0.0024	0.9472	0.1920	0.0002	0.0024	0.9502	0.1913

**Table 7. Bias and MSEs of the MLEs of the QUBH regression model and the CPs and MLs of the MLEs based approximately confidence intervals for one covariate case**

$\theta$	Covariates	n	$\theta_0$				$\theta_1$			
			Bias	MSEs	CP	ML	Bias	MSEs	CP	ML
$\Theta = (0.4, 0.2)$	Normal (0,1)	25	0.0825	0.0631	0.9280	0.9262	-0.0129	0.0733	0.9418	1.0210
		50	0.0415	0.0294	0.9372	0.6494	-0.0130	0.0304	0.9496	0.6830
		75	0.0282	0.0195	0.9386	0.5300	-0.0108	0.0189	0.9516	0.5464
		100	0.0180	0.0138	0.9488	0.4591	-0.0080	0.0140	0.9520	0.4689
		250	0.0079	0.0057	0.9436	0.2904	-0.0038	0.0053	0.9508	0.2895
	Binom (5,0,4)	25	0.0761	0.3012	0.9384	2.0568	-0.0078	0.0580	0.9496	0.9130
		50	0.0417	0.1301	0.9448	1.3974	-0.0042	0.0251	0.9486	0.6128
		75	0.0270	0.0832	0.9466	1.1233	-0.0031	0.0158	0.9506	0.4907
		100	0.0232	0.0620	0.9504	0.9682	-0.0042	0.0116	0.9496	0.4209
		250	0.0075	0.0241	0.9488	0.6059	-0.0002	0.0045	0.9498	0.2626
$\Theta = (0.8, 0.4)$	Normal (0,1)	25	0.0645	0.0611	0.9274	0.9137	-0.0231	0.0697	0.9496	1.0021
		50	0.0332	0.0272	0.9456	0.6382	-0.0063	0.0301	0.9458	0.6649
		75	0.0221	0.0181	0.9490	0.5196	-0.0094	0.0173	0.9526	0.5318
		100	0.0171	0.0137	0.9430	0.4494	-0.0073	0.0135	0.9520	0.4556
		250	0.0084	0.0054	0.9420	0.2836	-0.0027	0.0048	0.9550	0.2801
	Binom (5,0,4)	25	0.0431	0.2633	0.9422	1.9468	0.0024	0.0504	0.9480	0.8503
		50	0.0257	0.1197	0.9444	1.3307	-0.0007	0.0219	0.9508	0.5763
		75	0.0231	0.0745	0.9520	1.0761	-0.0030	0.0141	0.9508	0.4640
		100	0.0157	0.0567	0.9512	0.9258	-0.0012	0.0106	0.9492	0.3991
		250	0.0056	0.0214	0.9532	0.5785	-0.0012	0.0040	0.9532	0.2484
		500	0.0008	0.0108	0.9498	0.4079	0.0007	0.0020	0.9520	0.1750

## 7. CONCLUSIONS

In this study, a new distribution is proposed as an alternative to the Beta and Kumaraswamy distributions. Many mathematical properties of new distribution such as moments, coefficients of skewness and kurtosis, quantile function, stochastic order, order statistic, etc. are investigated. Five methods are examined for the estimation of the unknown parameter of the UBH distribution, and the performances of these estimators are examined by Monte Carlo simulation. From the simulation results, it is seen that all estimators behaved similarly when the sample size increased. Two numerical examples are studied to illustrate the usefulness of the new model. A new quantile regression model is constructed based on the UBH distribution.

## REFERENCES

- [1] Karakaya, K., Kinaci, İ., Kuş, C., Akdoğan, Y., *Hacettepe Journal of Mathematics and Statistics*, **46**(2), 303, 2017.
- [2] Kinaci, İ., Kuş, C., Karakaya, K., Akdoğan, Y., *Cumhuriyet Science Journal*, **40**(2), 378, 2019.
- [3] Alizadeh, M., Jamal, F., Yousof, H.M., Khanahmadi, M., Hamedani, G.G., *Scientific Bulletin-Series A-Applied Mathematics and Physics*, **82**(1), 145, 2021.
- [4] Rasekhi, A., Rasekhi, M., Hamedani, G.G., *Scientific Bulletin-Series A-Applied Mathematics and Physics*, **81**(2), 177, 2020.
- [5] Korkmaz, M.Ç., *Scientific Bulletin-Series A-Applied Mathematics and Physics*, **82**(2), 133, 2021.
- [6] Mazucheli, J., Menezes, Andre Felipe, B., Chakraborty, S., *Journal of Applied Statistics*, **46**(4), 700, 2019.
- [7] Mazucheli, J., Bapat, S.R., Menezes, Andre Felipe, B., *Chilean Journal of Statistics*, **11**(1), 53, 2020.
- [8] Korkmaz, M.Ç., Korkmaz, Z.S., *Journal of Applied Statistics*, <https://doi.org/10.1080/02664763.2021.2001442>, 2021.
- [9] Korkmaz, M.Ç., Chesneau, C., *Computational and Applied Mathematics*, **40**(1), 1, 2021.
- [10] Mazucheli, J., Menezes, A.F.B., Fernandes, L.B., de Oliveira, R.P., Ghitany, M.E., *Journal of Applied Statistics*, **47**(6), 954, 2020.
- [11] El-Morshedy, M., Eliwa, M.S., Altun, E., *IEEE Access*, **8**, 74359, 2020.
- [12] Yadav, A.S., Altun, E., Yousof, H.M., *Annals of Data Science*, **8**(2), 241, 2021.
- [13] El-Morshedy, M., Aljohani, H.M., Eliwa, M.S., Nassar, M., Shakhatreh, M.K., Afify, A.Z., *Mathematics*, **9**(18), 2277, 2021.
- [14] Ferrari, S. Cribari-Neto, F., *Journal of Applied Statistics*, **31**, 799, 2004.
- [15] Korkmaz, M.Ç., Altun, E., Chesneau, C., Yousof, H.M., *Mathematica Slovaca*, **72**(3), 765, 2022.
- [16] Korkmaz, M.Ç., Chesneau, C., Korkmaz, Z.S., *Journal of Reliability and Statistical Studies*, **15**(1), 261, 2022.
- [17] Korkmaz, M.Ç., Altun, E., Alizadeh, M., El-Morshedy, M., *Mathematics*, **9**(21), 2634, 2021.
- [18] Maniu, A.I., Voda, V.G., *Journal of Applied Quantitative Methods*, **3**(1), 215, 2008.
- [19] Shaked, M., Shanthikumar, J.G., *Stochastic Orders and their Applications*, Academic Press, London, England, 1994.
- [20] Taniş, C., *Journal of Science and Arts*, **21**(4), 951, 2021.
- [21] Taniş, C., *Journal of New Theory*, **34**, 72, 2021.
- [22] Deniz, E.G., Sordo, M.A., Calderin-Ojeda, E., *Mathematics and Economics*, **54**, 49, 2014.

## APPENDIX

```

ubh.moment=function(beta)
{
library(pracma) # it is needed for function incgam
Ei=function (a, z) {a^(z-1)*incgam(z,1-a)}
ex1=-exp(1+beta) *Ei(1,1+beta)+1
ex2=-2*exp(beta+2) *Ei(1,beta+2)+1
ex3=-3*exp(beta+3) *Ei(1,beta+3)+1
ex4=-4*exp(beta+4) *Ei(1,beta+4)+1
expected=ex1
variance=ex2-expected^2
skew=(ex3-3*ex1*ex2+2*ex1^3)/((variance)^(3/2))
kurt=(ex4-4*ex1*ex3+6*ex2*((ex1)^2-3*ex1^4))/(variance^2)
result=list("EX"=expected,"VarX"=variance,"Skewness"=skew,"Kurtosis"=kurt)
return(result)
}

```