The Kumaraswamy Lindley Regression Model with Application on the Egyptian Stock Exchange

Samy Abdelmoezz[1]*, Salah M. Mohamed[2]*

Abstract

We introduce and study the Kumaraswamy Lindely Distribution (KLD) model, which has increasing, decreasing, upside-down bathtub and bathtub shaped hazard functions.. We perform a Monte Carlo simulation study to assess the finite sample behavior of the maximum likelihood estimates of the parameters. We define a new regression model based on the new distribution. The new regression was applied to data from the Egyptian stock exchange in the period of (2015-2019). Finally we studies some properties of regression Residual analysis The martingale residual, Deviance component residual.

Keywords: Kumaraswamy Lindely distribution, Kumaraswamy Lindely regression Model, Link funtions, Residual analysis, martingale residual.

I. Introduction

There are many researchers who have been exposed to the Kumaraswamy Lindley distribution to the proposed Kumrsawamy Lindely distribution (KLD). Such as Cordeiro and Castro [8] have proposed a new family of generalized distributions. Elbatal, et al. [10] presented a new generalized Lindley distribution. Çakmakyapan and Kadılar [6] proposed a new customer lifetime duration distribution for the Kumaraswamy Lindley distribution. Broderick, et al. [5] devoted a generalized class of Kumaraswamy Lindley distribution with applications to lifetime data. Mahmoud, et al. [15] analyzed a log-beta log-logistic regression model. Cordeiro, et al. [9] presented the Kumaraswamy normal linear regression model with applications. Cakmakyapan, et.al. [7] presented the Kumaraswamy Marshall-Olkin log-logistic distribution with application. Vigas, et al. [22] presented the Poisson-Weibull regression model. Nofal, et al. [17] presented the transmuted Geometric-Weibull distribution and its regression model. Rocha, et al. [19] presented a Negative Binomial Kumaraswamy-G cure rate regression model. Handique, et al. [14] presented Marshall-Olkin-Kumaraswamy-G family of distributions. Eissa, et al. [11] presented exponentiated Kumaraswamy-Weibull distribution with application to real data. Shawki, et. al. [21] proposed Kumaraswamy Sushila distribution. Altun, et al. [3] presented a new generalization of generalized half-Normal distribution. Abed, et al. [1] proposed a new mixture statistical distribution Exponential – Kumaraswamy. Fachini-Gomes, et al. [12] presented the Bivariate Kumaraswamy Weibull regression model. Arshad, et al. [4] presented the gamma kumarsawmy- G distribution, theory, inference and applications. Mdlongwaa, et al. [16] presented Kumaraswamy log-logistic Weibull distribution, model theory and application to

_____

Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research (FSSR), Cairo University, Egypt.
Email address: samy_ez2010@yahoo.com[1], mahdym62@gmail.com[2]

Samy Abdelmoezz, Salah M. Mohamed

lifetime and survival data. Pumi, et al. [18] presented Kumaraswamy regression model with Aranda-Ordaz link function. Safari, et al. [20] presented robust reliability estimation for Lindley distribution, a probability integral transform statistical approach. Hafez, et al. [13] presented a study on Lindley Distribution accelerated life tests, application and numerical simulation. AlgarniI[2] devoted a new generalized Lindley distribution properties, estimation and applications.

This paper presents a new regression model. The paper is organized as: Section II presents the methods and algorithms. Section III presents the regression model of the proposed model. Section IV presents the numerical study. Finally, Section VI presents the conclusions.

2. Preliminaries

In this part of the research, we provide some information about the Kumasawamy and Lindley, because the model depends on contribute between the distributions So, the next subsection introduces some informations about two distributions.

2.1. Lindley Distribution

The Lindley distribution was introduced in 1958, but it was used as an alternative to the exponential distribution, where the Lindley distribution was used to study many characteristics such as data modeling and other characteristics In this section, the definition and properties of Lindley distribution are provided. Equation (1) presents the pdf of the Lindley distribution with parameter θ :

$$g(x) = \frac{\theta^2}{\theta + 1}(1 + x)e^{-\theta x} \qquad "x > 0, \theta > 0" \qquad (1)$$

The cdf of (1) is:

$$G(x) = 1 - e^{-\theta x}\left(1 + \frac{x\,\theta}{\theta + 1}\right) \qquad "x > 0, \theta > 0" \qquad (2)$$

Generalized distributions are very important in the field of probability distributions, as these generalized distributions contain many mathematical properties that make the distribution more flexible. Generalized distributions depend on two things, namely the cumulative distribution function(CDF) and the probability density function of the distribution (pdf) g(x) defined the K-G family by the cdf and pdf given by:

$$F(x) = 1 - [1 - G(x)^a]^b \, a > 0, \qquad\qquad b > 0, \qquad (3)$$

$$f(x) = ab\, g(x)\, G(x)^{a-1}[1 - G(x)^a]^{b-1}, \quad a > 0, \quad b > 0 \quad (4)$$

where $(x) = \frac{dG(x)}{dx}$, (a) and (b) are two additional positive shape parameters. Clearly, for (a = b = 1), we obtain the baseline distribution. The additional parameters (a) and (b) aim to given skewness and tail weight of the generated distribution.

2.2. Kumaraswamy Lindley Distribution (KLD)

Let G(x, θ) be the cdf of the Lindley distribution given by (2). The cdf of Kw-Lindley distribution can be defined by substituting (2) into equation (3). Hence the cdf of Kw-Lindley is

$$F(x) = 1 - \left[1 - \left(1 - e^{-\theta x}\left(1 + \frac{\theta x}{\theta + 1}\right)\right)^a\right]^b \qquad (5)$$

Samy Abdelmoezz, Salah M. Mohamed

By substituting equation (1) and (2) into equation (4), we get the probability density function for the Kumaraswmy Lindley distribution.

$$f(x) = ab \left(\frac{\theta^2}{\theta + 1} (1 + x)e^{-\theta x} \right) \left(1 - e^{-\theta x} \left(1 + \frac{\theta x}{\theta + 1} \right) \right)^{a-1}$$

$$\left[1 - \left(1 - e^{-\theta x} \left(1 + \frac{\theta x}{\theta + 1} \right) \right)^a \right]^{b-1} \qquad (6)$$
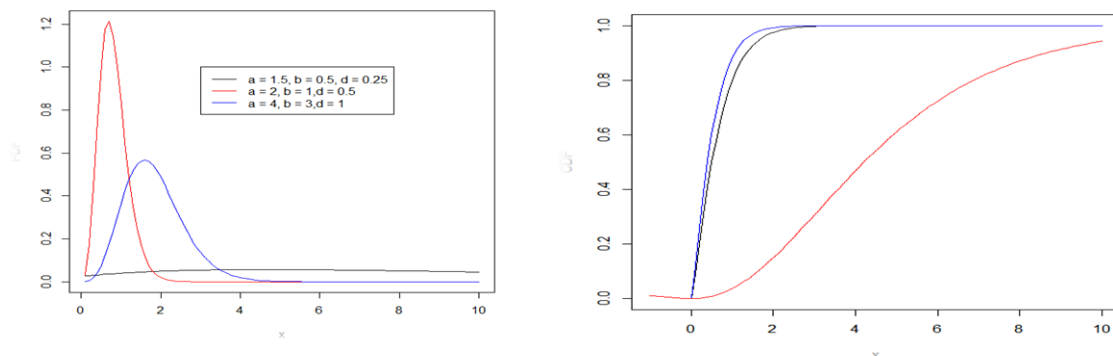

Figure 1. pdf(leght) and cdf(right) plot of(KWL) Kumaraswamy LIndley distribution
The previous figure shows a plot of the probability density function and the cumulative distribution function for the Kumaraswamy Lindley distribution at the different values of the parameters.
Note that the Kumaraswamy Lindley is very flexible distribution that approaches to different distributions for some special values of the parameters which is given below.

Table 1. Special cases of the Kumaraswamy Lindley Distribution

a	B	$f(x)$	Distribution
1	1	$\frac{\theta^2}{\theta + 1} (1 + x)e^{-\theta x}$	Lindley
-	1	$a \left(\frac{\theta^2}{\theta + 1} (1 + x) e^{-\theta x} \cdot (1 - e^{-\theta x} (1 + \frac{\theta x}{\theta+1})) \right)^{a-1}$	GLD

3. Results and Discussion

In this section, we suggest appropriate link function to make a regression model that contributes with the KLD.

3.1 Estimation of Parameters' Regression Model

Let X be a random variable having the (KWL) density function .The random variable Y = log(X) defined the log- Log kumrsawamy Lindely distribution (LKLD) distribution. The density function of Y , re-parameterized in terms $\theta = e^{-\mu}$ of, can be expressed
As

$$f(y) = \frac{abe^{-2\mu}}{1+e^{-\mu}} (1 + e^y) \left[1 - e^{-e^{y-\mu}} \left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}} \right) \right]^{a-1} \cdot \left[1 - e^{-e^{y-\mu}} \left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}} \right) \right]^{b-1} e^{-e^{y-\mu}}$$

Jurnal Matematika, Statistika & Komputasi
Samy Abdelmoezz, Salah M. Mohamed

$\theta_i = \exp(-\mu_i)$, $\mu_i = x'\beta$, $\beta' = (\beta_0 \quad \beta_1)$, $x' = (1 \quad x_i)$, $-\infty \le \mu_i \le \infty$.

$y_j = \log(x_i)$, $x_i = \exp(y_i)$. $\theta = e^{-\mu}$

the survival function corresponding for the Kumaraswamy Lindley distribution is

$$S(x) = 1 - F(x) \quad \text{implies to} \quad s(x) = 1 - \left\{1 - \left[1 - \left(G(x)\right)^a\right]^b\right\}$$

But the commutative function of Lindley distribution is

$$G(x) = 1 - e^{-\theta x}\left(1 + \frac{\theta x}{\theta+1}\right)$$

Substitute Lindley's cumulative distribution function in the survival function corresponding for the Kumaraswamy Lindley distribution we obtain

$$s(x) = \left[1 - \left(1 - e^{-\theta x}\left(1 + \frac{\theta x}{\theta+1}\right)\right)^a\right]^b$$

Compensating for Y = log(X),. θ = e$^{-\mu}$.the survival function becomes

$$s(x) = \left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^a\right]^b$$

The linear regression model tends to provide good data and estimations for the model and by constructing a linear regression model in which the response variable is y, and the independent variable is x_i and depends on the Kumaraswamy Lindley distribution as follows

$$y_i = \beta x^T + z_i \quad \text{where } i = 1,2,3, ...$$

where the random error z_i has density function (KWL) with unknown parameters $\theta > 0$.,

$a > 0, b > 0$ The parameter $\mu_i = x^T\beta \in R$ is the location parameter of yi where the parameter $\mu_i = (\mu_i, \mu_2, ……., \mu_n)^T X_i = (1, x_1, ….x_n)^T$ the log-likelihood function for the vector parameters $⊡ = (\beta^T, \theta, a, b)$ can expression as following

$$L(⊡) = \sum_{i=1}^{n} \delta_i \log f(y_i) + \sum_{i=1}^{n}(1 - \delta_i)\log\left(S(y_i)\right)$$

Taking the logarithm of the probability density function of the Kumaraswamy Lindley distribution to simplify, we get

$$\log f(y) = \log ab + \log e^{-2\mu} - \log(1 + e^{-\mu}) + \log(1 + e^y) + (a-1)\log\left[1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1 + e^{-\mu}}\right)\right]$$

$$+(b-1)\log\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1 + e^{-\mu}}\right)\right)^a\right] - e^{y-\mu}$$

Taking the logarithm of the survival function of the Kumaraswamy Lindley distribution to simplify, we get

$$\log s(y_i) = b \log\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^a\right]$$

Substituting the logarithm value of the probability density function and the survival function of the Kumaraswamy Lindley distribution in the following equation, we get

Jurnal Matematika, Statistika & Komputasi

Samy Abdelmoezz, Salah M. Mohamed

$$L = \sum_{i=1}^{n} \delta_i \, \log f(y_i) + \sum_{i=1}^{n}(1 - \delta_i) \, \log\left(s(y_i)\right)$$

$$L = \sum_{i=1}^{n} \delta_I \log ab + \delta_I \log e^{-2\mu} + \delta_I \log(1 + e^y) - \delta_I \log(1 + e^{-\mu})$$

$$+(a-1)\delta_i \sum_{i=1}^{n} \log\left[1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1 + e^{-\mu}}\right)\right]$$

$$-\delta_i e^{y-\mu} \quad + (b-1)\delta_i \sum_{i=1}^{n} \log\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1 + e^{-\mu}}\right)\right)^a\right]$$

$$+\sum_{i=1}^{n}(1 - \delta_i)b \, \log\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1 + e^{-\mu}}\right)\right)^a\right]$$

let $\sum_{i=1}^{n} \delta_i \, \log f(y_i) = A$ and $B = \sum_{i=1}^{n}(\delta_i - 1) \, \log s(y_i)$

then   the likelihood becomes L = A + B

$$\frac{\partial L}{\partial a} = \frac{\partial A}{\partial a} + \frac{\partial B}{\partial a} = 0, \quad \frac{\partial L}{\partial b} = \frac{\partial A}{\partial b} + \frac{\partial B}{\partial b} = 0, \quad \frac{\partial L}{\partial \beta} = \frac{\partial A}{\partial \beta} + \frac{\partial B}{\partial \beta} = 0.$$

But $A = n \ln ab \, \delta_i - 2\sum_{i=1}^{n} \delta_i \mu + \delta_i \sum_{i=1}^{n} \ln(1 + e^{y_i}) - \sum_{i=1}^{n} \delta_i \, \ln(1 + e^{-\mu})$

$$-\sum_{i=1}^{n} \delta_i e^{y-\mu} + (a-1)\sum_{i=1}^{n} \delta_i \ln\left[1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1 + e^{-\mu}}\right)\right]$$

$$+ (b-1) \ln\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1 + e^{-\mu}}\right)\right)^a\right]$$

And $B = \sum_{i=1}^{n}(1 - \delta_i)b \ln\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^a\right]$ (9)

To estimate a, we differentiate the equation with respect to a,

$$\frac{\partial A}{\partial a} = \frac{n}{a}\delta_i + \sum_{i=1}^{n} \delta_i \ln\left[1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right] -$$

$$\frac{(b-1)\delta_i \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^a \ln\left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)}{\left[1 - \left[1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right]^a\right]} = 0 \quad (10)$$

$$\frac{\partial B}{\partial a} = \frac{\sum_{i=1}^{n}(i - \delta_i) \, b \left[-\left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^a (1) \ln\left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)\right]}{\left[1 - \left[1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right]^a\right]} = 0 \quad (11)$$

To estimate b, we differentiate the equation with respect to b,

$$\frac{\partial A}{\partial b} = \frac{n\delta_i}{b} + \sum_{i=1}^{n} \delta_i \ln\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^a\right] \quad (12)$$

$$\frac{\partial B}{\partial b} = \sum_{i=1}^{n}(1 - \delta_i) \ln\left[1 - \left(1 - e^{-e^{y-\mu}}\left(1 + \frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^a\right] \quad (13)$$

To estimate β, we differentiate the equation with respect to β, $\quad \frac{\partial A}{\partial \beta} + \frac{\partial B}{\partial \beta} = 0$

Samy Abdelmoezz, Salah M. Mohamed

$$\frac{\partial B}{\partial \beta} = \sum_{i=1}^{n}(1-\delta_i)\ b\ \frac{(-)\,a\left(1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^{a-1}\cdot\left[-e^{-e^{y-\mu}}\cdot\left(\frac{(1+e^{-\mu})e^{y-\mu}(-x^T)-e^{y-\mu}(-x^t e^{-\mu})}{(1+e^{-\mu})2}\right)\right]}{\left[1-\left[1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right]^a\right]}$$

$$+\frac{\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)(-)\,e^{-e^{y-\mu}}(-e^{y-\mu})(-x^T)}{\left[1-\left[1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right]^a\right]}$$

$$\frac{\partial A}{\partial \beta} = e^{y-\mu}x^T\ -2\delta_i x^T - \sum_{i=1}^{n}\delta_i\frac{-x^T e^{-\mu}}{1+e^{-\mu}}+(a-1)$$

$$\sum_{i=1}^{n}\frac{(-\delta_i)e^{-e^{y-\mu}}\left(\frac{(1+e^{-\mu})e^{y-\mu}(-x^T)-e^{y-\mu}(e^{-\mu})(-x^T)}{1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)}\right)+\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)(e^{-e^{y-\mu}})(-e^{y-\mu}(-x^T))}{\left[1-\left[1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right]^a\right]}$$
$$+$$

$$+(b-1)\sum_{i=1}^{n}\delta_i\frac{(-a)\left(1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right)^{a-1}\left[-e^{-e^{y-\mu}}\left[\frac{(1+e^{-\mu})\,e^{y-\mu}(-x^T)-(e^{y-\mu})(-x^T e^{-\mu})}{(1+e^{-\mu})2}\right]+\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)(-e^{-e^{y-\mu}})(-e^{y-\mu})(-x^T)\right]}{\left[1-\left[1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{1+e^{-\mu}}\right)\right]^a\right]} = 0 \quad (14)$$

3.2. Residual analysis

In order to verify departures from the error assumptions as well as the presence of outliers, we consider two types of residuals: the martingale residual and the deviance component residual

3.2.1. The martingale residual

The martingale residual is much used in counting process (Fleming and Harrington, 1991). These residuals are asymmetric and take maximum values +1 and minimum values$-\infty$. The martingale residual is defined by

$$r_{Mi} = \delta_i - \left(\int_0^y h(u)\ du\right) \quad (15)$$

where
$$\delta_i = \begin{bmatrix}0 & \text{i th observation is censored} \\ 1 & \text{i th observation is uncensored}\end{bmatrix}$$

As we known

$$\int_0^y h(u)\ du = \ln[s(y_i)] \quad (16)$$

The martingale residual can be reduces to

$$r_{Mi} = \delta_i + \ln[s(y_i)]$$

The martingale residual for the log Kumaraswamy Lindely model takes the from

$$r_{Mi} = 1 + \ln\left[1-\left(1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{e^{-\mu}}\right)\right)^a\right]^b \quad (17)$$

3.3.2 Deviance component residual

This residue was suggested to make the martingale residual more symmetric around zero. The deviance component for the parametric regression model is given

$$r_{Di} = sign\,(r_{Mi})\,\{-2\,[r_{Mi}+\ln(1-r_{Mi})]\}^{1/2} \quad (18)$$

Where r_{Mi} is the martingale residual.

sign () function is a function that drives the (+1) values if the argument is positive and

Samy Abdelmoezz, Salah M. Mohamed

(-1) is negative. The deviance component residual for the KLL model is given by

$$r_{Di} = sign\,(r_{Mi}) \left\{ \begin{array}{l} -2\left[\ln\left[1-\left(1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{e^{-\mu}}\right)\right)^a\right]^b\right] \\ +log\left[1-\ln\left[1-\left(1-e^{-e^{y-\mu}}\left(1+\frac{e^{y-\mu}}{e^{-\mu}}\right)\right)^a\right]^b\right] \end{array} \right\}^{1/2}$$

4. Global influences

One of the most famous ways to assess the sensitivity of the model is the global influence. Starting from case deletion (see, cook, 1977). The case – deletion is a common approach to study the effected of dropping the (*i-th*) observation from the data. First we obtain the estimates of parameters and get the log-like of these estimates. Then we estimate the parameters after cancelling one observation from the dataset. The aim is comparing between the log-likelihood in two cases. This measure can be written as:

Cook measure = 2(Log-like before eliminating obs. - Log-like after eliminating obs.) To explain the effect of cancelling one observation on the estimation of parameters, we generate the same data with $n = 99$ observations. The

Table 2. Estimation of these parameters in the two cases is:

Case	$\hat{a}$	$\hat{b}$	$\hat{\theta}$	LOGL
Estimates with 100 obs.	1.498504	0.2209227	0.4840113	-384.3581
Estimates with 99 obs.	1.499413	0.2219762	0.4888745	-381.0098

Cook = 2(LOGLB-LOGLA) = -6.696586

We have noted that the cook is positive negative value, this means that the log-likelihood in small values (negative) is less than it is large values. This is happened because the values of estimates become less.

1. Simulation study

In this section of the research, We perform a Monte Carlo simulation study to assess the finite sample behavior of the MLEs of d, a and b. The results are obtained from 1000 Monte Carlo simulations carried out using the R statistical software. In each replication, a random sample of size n is drawn from the(KLD) distribution and the parameters are estimated by maximum likelihood. The random variable X is generated using the inversion method. We consider there setups with the following values for the parameters: a=1.5,b=0.5, d= 0.5 step(1), a=2, b=0.5, d=1and (step2), a=1.5, b=2.5, d=0.5. After that, the parameters were estimated in the three cases, as well as some statistical measures were calculated, such as: AKaike information criteria(AIC) Bayesian information criteria(BIC) and Quinn information criteria(HQIC) *t* is known that when the value of the following statistical measures AKaike information criteria (AIC) Bayesin information criteria (BIC), and AKaike information criteria (AIC) Hannan-Quinn information criteria (HQIC) is small, the model is better by extrapolating the following table 3. We determine that When the sample size increases, as well as the increase in the value of the parameters, we notice that the value of the AKaike Information Standards (AIC) Bayesin Information Standards (BIC) and the AKaike Information Standards (AIC) Hannan-Quinn Information Standards (HQIC) value decreases, so we recommend this when using a model Kumarasamy Lindley which is used with small data.

Table 3. Simulation results of the KLD distribution for several values of parameters

Samy Abdelmoezz, Salah M. Mohamed

a	b	D	N	$\hat{a}$	$\hat{b}$	$\hat{d}$	AIC	BIC	HQIC
1.5	0.5	0.5	50	1.50386	0.21786	0.4734	384.12	374.45	388.56
			100	1.4985	0.22092	0.4842	758.71	745.6904	762.607
			1000	1.4972	0.2387	0.4837	6940.9	6916.388	6916.38
2	0.5	1	50	2.02894	0.30220	1.2183	433.176	423.616	6943.19
			100	2.0200	0.32619	1.1700	760.89	747.887	764.7847
			1000	2.0173	0.35459	1.1636	6937.8	6913.34	6940.157
1.5	2.5	0.5	50	0.25012	18.3174	5.2396	48451.4	48441.8	48455.99
			100	0.29857	18.1652	5.3256	95339	95325.9	95342.9
			1000	0.50024	4.76823	4.7882	215235.4	215210.8	215237.6

2. Numerical study

This paper dealt with some data from the Egyptian Stock Exchange, and these data represent the monthly data from the Egyptian Stock Exchange, during the period starting from February 2015 to February 2019, and it represents the relationship between the impact of the transactions of the finance, tourism, food and health sectors on the Egyptian Stock Exchange indices, where we find that this The data contains one dependent variable, which is the Egyptian stock market indices, and four independent variables, which are food, health, tourism, and finance sectors.
In Table 4, in this table, the arithmetic mean of the data was calculated, as well as the variance, standard deviation, kurtosis and skewness for these data. We note on these data that the skew value is close to zero and thus the data are similar around the origin and contain the kurtosis coefficient negative.

Table 4. Descriptive statistics of the indicator of EDX30

Mean	SD	Var	skewness	kurtosis	Min	Max	Range
11421.27	3513.659	12345797.743	.039	-1.360	5993	18296	12303

And because we will present to the following table, which contains an estimate of the parameters for each of the following distributions, and they are, in order, the Kamraswamy Lindley distribution, the generalized Lindley distribution, and finally the Lindley distribution. Also, the table contains some properties of the parameters, which are (BIC) Bayesian information criterion ,(AIC) Akaike information criterion ,(HQ)Hannan –Quiun information criterion which are, and the principle on which we depend in judging the preference of any distribution is the distribution that has a value less than the previous criteria will be the best.

Table 5. MLES and information criteria for klD,GL,L

distribution	$\hat{a}$	$\hat{b}$	$\hat{d}$	AIC	BIC	HQIC	
KLD	2.8156391	0.3222782	4.296497	8240.118	8256.609	8236.787	
GL	2.7562	1	3.42356	8245.26	8257.42	8238.54	
L	1	1	3.5861	8249.23	8259.24	8239.23	

Regression model

The following data set contains 50 observations representing the relationship between the impact of the transactions of the finance, tourism, food and health sectors on the impact of the Egyptian Stock Exchange during the period from February 2015 to February 2019. We want to calculate the relationship between food and the influencers of the performance of the Egyptian Stock Exchange, as it represents the independent variable in this The relationship between food and the dependent variable is the effect of the stock market performance, and the relationship between the two variables can be put in the form of a linear regression model on the following

$$y_i = \beta_0 + \beta_1 x_i + z_i$$

Where y_i has the (kwL)density for i=1……….50, The following table contains the estimation of parameters for the proposed model, as well as some statistical measures that are used in comparing the different models. This proposed model is fully compared to each of the Lindley Ingression and the Lindley distribution as well as the generalized Lindley distribution model, where it is completely relied on three different measures,(BIC)Bayesian information criterion ,(AIC) Akaike information criterion , (HQ)Hannan –Quiun information.Where the model is better if the value of these measures is less than the value of the same measures for any other model. We note from the table that the value of the scales in relation to the Kumarswamy Lindley model is smaller than the value of the same measures of the Lindley regression model, as well as the generalized Lindley regression model and therefore, the Kumarswamy Lindley regression model is the best.

Table 6. Table of estimation of parameters , regression and information criteria for each of , KLD, GL, L

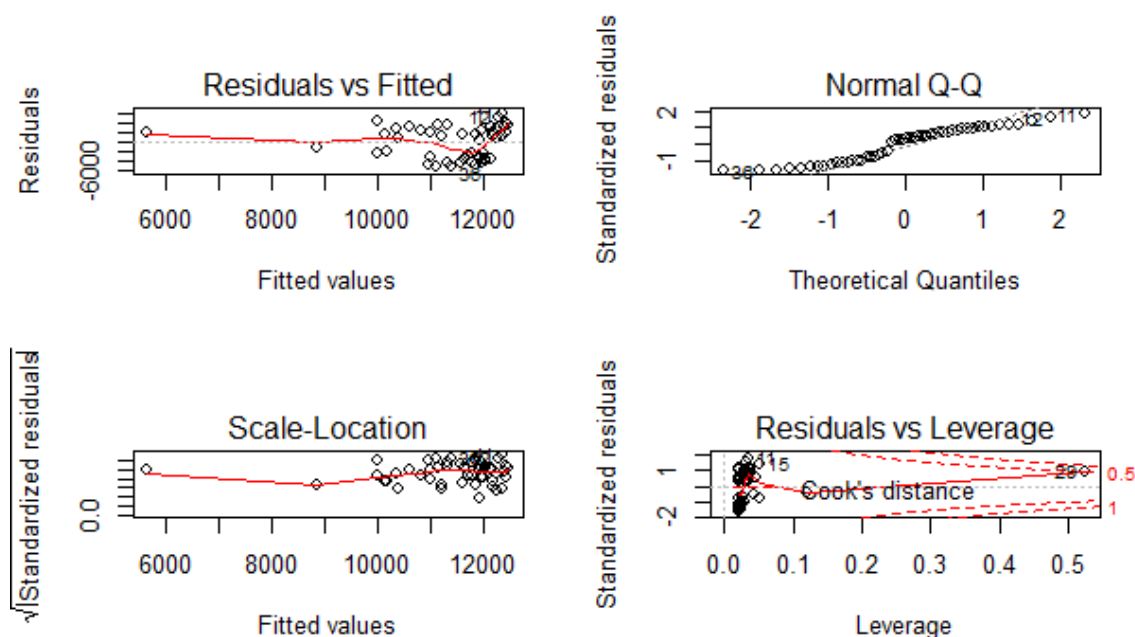| Distribution	$\hat{a}$	$\hat{b}$	$\hat{d}$	$\widehat{\beta_0}$	$\widehat{\beta_1}$	AIC	BIC	HQIC
KLD	2.9985	2.00910	-0.00406	-0.1743	0.0077	2743.7	2760.2	2740.42
GL	1.2356	-	0.0425	-0.1234	0.0052	2745.5	2762.2	2744.42
L	-	-	0.01235	0.4235	-0.4423	2746.42	2854.2	2751.51

Figure 2. Plot regression between indicator and food

4.   Conclusions

In this paper, we introduce the Kumaraswamy Lindley distribution regression model and study some new properties of the regression. Also, we used Cook measure to indicate the changes on estimation parameters when one observation is eliminated from observed data. Furthermore, the appropriate regression model according to appropriate suggest link function is constructed, and its features including the residuals error, deviance, log-likelihood, AIC, BIC,….etc. are obtained. We hope that the proposed model may attract applications.

References

[1].   Abed, M.Q., Alwan, E.B, Kneehr, A.L., 2018.   Proposing a New Mixture Statistical Distribution Exponential – Kumaraswamy. *International Journal of Applied Science and Technology*, Vol. 8, No. 2, pp. 48-56.

[2].   Algarni, A.,ID2021. On a new generalized lindley distribution Properties, estimation and applications. *PLoS ONE* 16(2), https:// doi.org/10.1371/journal.pone.0244328.

[3]. Altun, E., Yousof, H.M., Hamedani, G.G., 2018. A new generalization of generalized half-normal distribution: properties and regression models. *Journal of Statistical Distributions and Applications*, Vol.5, No.4, pp. 2-16.

[4]. Arshad, R.M.I., Chesneau, C. Jamal, F., Mansoor, M., Zubair, M., Nasir, M., 2019. The gamma kumarsawmy-G familey distribution: theory, inference and applications. *HAL* Id: hal-02103938, https://hal.archives-ouvertes.fr/hal-02103938.

[5]. Broderick O. Oluyede, Tiantian Yang , Bernard Omolo, 2015. A generalized class of Kumaraswamy Lindley Distribution with Applications to lifetime data. *Journal of Computations & Modelling*, Vol.5, No.1, pp. 27-70.

[6]. Çakmakyapan, S., Kadılar, G. Ö., 2014. A new customer lifetime duration distribution: The Kumaraswamy Lindley Distribution. *International Journal of Trade, Economics and Finance*, Vol. 5, No. 5, pp. 12-24.

[7]. Cakmakyapan, S., Ozel, G., El Gebaly, Y.M.H., Hamedaniet, G.G., 2017. The Kumaraswamy Marshall-Olkin Log-Logistic Distribution with Application. *Journal of Statistical Theory and Applications*, Vol. 17, No. 1,pp. 59–76.

[8]. Cordeiro, G. M., de Castro, M., 2011. A new family of generalized distributions. *Journal of statistical computation and simulation*, Vol. 81, No. 7, pp. 883-898.

[9]. Cordeiro, G.M., Machado, E.C., Botter, D.A., & Sandoval, M.C., 2018. The Kumaraswamy normal linear regression model with applications. *Communications in Statistics-Simulation and Computation*, Vol. 47, No. 10, pp. 3062-3082.

[10]. Elbatal, I., Merovci, F., Elgarhy, M., 2013. A new generalized Lindley distribution. *Mathematical theory and Modeling*, Vol. 3, No. 13, pp. 30-47.

[11]. Eissa, F. H., Abdulaziz, R. K., 2017. The exponentiated Kumaraswamy–Weibull distribution with application to real data. *International Journal of Statistics and Probability*, Vol. 6, No. 6, pp. 167-182.

[12]. Fachini-Gomes, J.B., Edwin M. M. Ortega, E.M.M., Cordeiro, G.M., Suzuki, A.K.,ʼ 2018. The Bivariate Kumaraswamy Weibull   regression model: a complete classical and Bayesian analysis. *Communications for Statistical Applications and Methods*, Vol. 25, No. 5, pp. 523–544.

[13]. Hafez, E. H., Riad, F. H., Mubarak, S. A., & Mohamed, M. S., 2020. Study on Lindley distribution accelerated life tests: application and numerical simulation. *Symmetry*, Vo. 12, No. 12, 2080, doi:10.3390/sym12122080.

[14]. Handique, L., Chakraborty, S., Hamedani, G.G., 2017. The Marshall-Olkin-Kumaraswamy-G family of distributions. *Journal of Statistical Theory and Applications*, Vol. 16, No. 4 pp. 427–447.

[15]. Mahmoud, M.R., EL- Sheikhb, A.A, Moradc, N.A, Ahmad, M.A.M., 2015. Log-Beta Log-Logistic Regression Model. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, Vol. 22, No. 2, pp. 389-405.

[16]. Mdlongwa, P., Oluyede, B.O., Amey, A.K.A., Fagbamigbe, A.F., Makubate, B. 2019. Kumaraswamy log-logistic Weibull distribution: model theory and application to lifetime and survival data. *Heliyon* 5 (2019) e01144. doi: 10.1016/j.heliyon.2019.e01144.

[17]. Nofal, Z. M., El Gebaly, Y. M., Altun, E., Alizadeh, M., & Butt, N. S., 2017. The transmuted geometric-Weibull distribution: properties, characterizations and regression models. *Pakistan Journal of Statistics and Operation Research*, Vol. 13, No.2, pp. 395-416.

[18]. Pumi, G., Rauber, C., Bayer, F. M., 2020. Kumaraswamy regression model with Aranda-Ordaz link function. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research, Springer;Sociedad de Estadística e Investigación Operativa*, Vol. 29, No. 4, pp. 1051-1071, https://doi.org/10.1007/s11749-020-00700-8.

[19]. Rocha, R., Tomazella, V., Louzada, F., 2017. Negative Binomial Kumaraswamy-G Cure Rate Regression Model. *Journal of risk and financial management*,Vol. 11, No. 6, pp. 2-14.

[20]. Safari, M.A.M., Masseran, N., Majid, M.H.A., 2020. Robust Reliability Estimation for Lindley Distribution—A Probability Integral Transform Statistical Approach. *Mathematics*,Vol. 8, issue 1634, https://doi.org/10.3390/math8091634.

[21]. Shawki, A.W, Elgarhy, M., 2017. Kumaraswamy SUSHILA Distribution. *International Journal of Scientific Engineering and Science*, Vol. 1, Issue 7, pp. 29-32.

[22]. Vigas, V.P., Silva, G.O., Louzada, F., 2017. The Poisson-Weibull Regression Model. *Journal of Statistics*, Vol. 8, No. 1, pp. 25-51.

Jurnal Matematika, Statistika & Komputasi
Samy Abdelmoezz, Salah M. Mohamed

.