A New Regression Model for Transmuted Weibull with Application

Mohamed S. Hamed[1,4], Mostafa Abd-elwahab Hamed[3*], Salah Mahdy Ramadan[2] and Amal Mohamed Abdel Fattah[2]

[1]Department of Business Administration, Gulf Colleges, KSA.

[2]Department of Applied Statistics and Econometrics, Faculty of Graduates Studies for Statistical Research (FSSR), Cairo University, Egypt.

[3]Department of Statistics, Faculty of commerce, Aswan University, Egypt.

[4]Department of Statistics & Mathematics and Insurance, Faculties of Commerce, Benha University, Egypt.

***Corresponding Author**

DOI: https://doi.org/10.51584/IJRIAS.2024.908070

Received: 26 August 2024; Accepted: 31 August 2024; Published: 23 September 2024

ABSTRACT

In this research paper, a new regression model for the transmuted Weibull distribution was proposed, and the parameters of the new model were estimated by using the Maximum likelihood method, the different types of residuals were calculated for the new model, such as Cox and Snell Residual, Pearson Residuals, Deviance component Residual, and the martingale residual. The proposed model was applied to a set of real data representing bad debit rate, and the proposed model was also compared with some other models such as the Kumaraswamy Lindley regression model and the LG Weibull regression model.

Keywords: Weibull distribution, transmutation map, regression approach, survival analysis, and residual analysis.

INTRODUCTION

Regression analysis is statistical modeling that can be used in financing, marketing, investing and other fields that utilized to infer the strength and direction of relationships between dependent variable and one or more independent variables. The dependent variable Y is also known as response variable or outcome, and the independent variables X_k where $k = (1, 2, \dots, p)$ as predictors, explanatory variables, or covariates. Regression analysis, more precisely, attempts identifying the mathematical formula that explains Y in terms of X as, $Y = f(x)$. **Sarstedt. M., & Mooi. E. (2014)**. Transmutation maps are generally a handy tool for creating a new distribution, especially, survival ones. Transmutation maps, defined by **Shaw and Buckley (2009),** are the functional composition of one distribution's inverse cumulative distribution (quantile) function and one distribution's cumulative distribution function. **Recently,** some studies involving the type of quadratic rank transmutation map (QRTM) and its application such as survival analysis. by applying QRTM **Aryal and Toskes (2011)** proposed Transmuted Weibull distribution, **Aryal and Tsokos (2009)** presented a generalization of the extreme value distribution applied this new distribution to analyze snowfall data at Midway Airport in the US state of Illinois, **Merovci, F. (2013a)** introduced transmuted Lindley distribution and applied the proposed distribution to an uncensored data of random sample from 128 bladder cancer patients, **Merovci, F. (2013b)** generalized the Rayleigh distribution and applied to Real data of nicotine measurements obtained from several brands of cigarettes for the year 1998. **Louzada and Granzotto, (2016)** proposed the transmuted log-logistic model and the regression one in the study of the time until the first calve of a polled Tabapua race. **Dey. S, et al. (2021)** stated that with increasing the variability of applications, the standard distributions were not sufficient for modeling complicated phenomena, therefore the need to generalize standard distributions is required. It is necessary to emphasis transmutation map as generalization tool for obtaining new distributions that are more flexible, the new generalized distributions have the capacity for modeling complicated phenomena more accurately. These generalizations made a considerable advancement in the direction of constriction flexile distributions in order to facilitate good modeling of lifetime data. Recently some generalizations of Weibull distribution including exponentiated Weibull, modified Weibull, extended Weibull and transmuted Weibull that can be derived by using QRTM. The generalizations and modifications of Weibull distribution have been approved by **Cordeiro, G. M. et al (2011), Elbatal. I and Aryal. G. (2013), Aryal and Toskes (2013), Eltehiwy, et al. (2013), Merovice, et al. (2013), Afify. A. Z, et al. (2014), Al-Kadim, K. A. et al. (2017), Afify. A. Z, et al. (2017), Ahmed. Z. Afify, et al. (2018), Nofal, Z. M., et al. (2018), Khan. (2019).**

In this paper we introduced a New Regression Model for Transmuted Weibull with application. The background of transmuted weibull distribution will be presented at section (3) are divided into three subsections including the log transmuted regression weibull, inference about the parameters of regression model by using maximum likelihood estimation, and finally subsection is residual analysis. At section (3) the numerical experiment will be presented and classified to several subsections involving simulation study in order to assess the performance of T-W regression model, applications of T-W to real data set, goodness of fit to ensure that real data more fitted to the proposed model, and then comparing the T-W regression model with kumaraswamy Lindley regression model and MG weibull regression model based on some goodness of fit criteria, and also the global influence will be presented to shows the sensitive analysis for T-W regression model.

MATERIAL METHODS

Definition of Transmuted Weibull distribution

The Weibull distribution is playing very important role for modeling lifetime data in many filed for example medicine, biology, engineering and finance, a very small amount of the massive applications of Weibull model. let's random variable x is said to have Weibull distribution with 2 parameters α and β if its probability function (pdf) is given by **(Aryal and Tsokos 2009):**

$$g(x) = \frac{\alpha}{\beta^{\alpha}}\ x^{\alpha-1}\, e^{\ -(\frac{x}{\beta})^{\alpha}} \qquad x > 0\ and\ \alpha, \beta > 0 \qquad \textbf{(1)}$$

The corresponding cumulative distribution function of Weibull distribution (CDF) is given by:

$$G(X) = 1 - e^{\ -(\frac{x}{\beta})^{\alpha}} \qquad x > 0\ and\ \alpha, \beta > 0 \qquad \textbf{(2)}$$

By using methodology of QRTM shown by Shaw and Buckley (2007), **Aryal and Toskes (2011)** proposed transmuted weibull distribution based on weibull distribution as baseline distribution. Let X is said to have transmuted distribution if its cumulative distribution (CDF) is given by:

$$F(x) = (1 + k)\ G(x) - k[G(x)]^2 \qquad \textbf{(3)}$$

Where $k \in [-1,1]$, *F(x)* is CDF of transmuted distribution, And *G(x)* is CDF of parent distribution. When K=0, the Transmuted distribution is the same as base distribution. The corresponding probability density function (pdf) is given by:

$$f(x) = g(x)[(1 + k)\ - 2kG(x)] \qquad \textbf{(4)}$$

Hence, the pdf and CDF of Transmuted weibull distribution is given by respectively:

$$f(x) = \frac{\alpha}{\beta^{\alpha}}\ x^{\alpha-1}\, e^{\ -(\frac{x}{\beta})^{\alpha}} \left[(1 + k)\ - 2k(1 - e^{\ -(\frac{x}{\beta})^{\alpha}})\right] \qquad (5)$$

And

$$F(x) = (1+k)\left(1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}}\right) - k\left[1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}}\right]^2 \quad (6)$$

According to **Aryal and Tsokos (2011);** Transmuted Weibull distribution is extended model for analyzing data that is more complex and it generalizes some of widely used distributions. From equation (5) note that when parameters vary, there are special cases from Transmuted Weibull distribution will be placed as following:

i. When $\alpha = 1$, transmuted exponential distribution (skew exponential) with two parameters β and k will be placed. Shaw, et al. (2009).

ii. By butting $k = 0$, Weibull distribution is generated which is the base distribution.

When $\alpha = 1$ and $k = 1$, the Exponential distribution with parameter $(\frac{\beta}{2})$ is given.

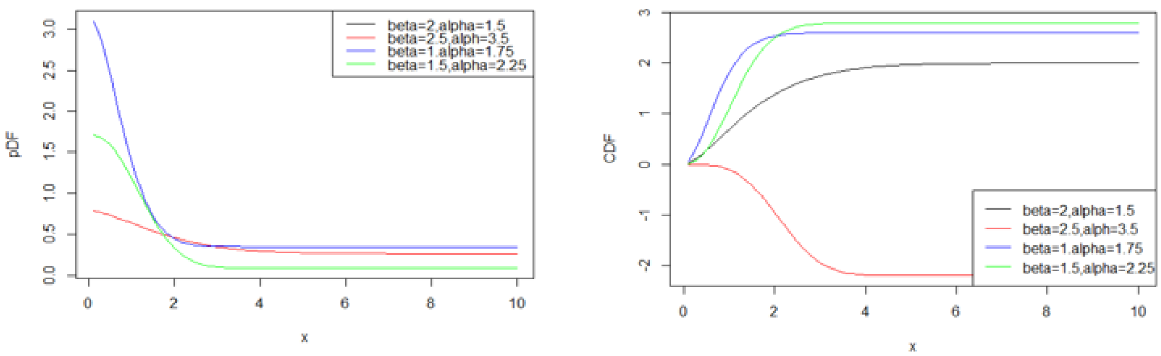

Figure (1). the pdf and CDF of the transmuted weibull distribution at different value with k=0.6

The Log Transmuted Weibull (T-W) Regression Model

The main objective of this paper is to introduce a new application of the Transmuted Weibull distribution in regression modeling. The proposed model utilizes the log- Transmuted Weibull distribution, which is derived from the positive Transmuted Weibull random quantity through a log transformation. The log-location-scale regression models are popular models to analyze the censored response variable with some covariates. In the last decade, researchers have introduced flexible location scale regression models to analyze the different characteristics of the data sets. The important paper on location-scale can be cited as follows: log-generalized Transmuted-Weibull, suppose x is a random variable following the **T-W** density function in equation (5). and y is defined $y = logx$, $x = e^{y}$, $\beta = e^{\mu}$, $\alpha = \frac{1}{\sigma}$, It is easy to verify that the density function of y reduces to:

$$f(y) = (|J|).f^{-1}(x) \quad \textbf{(7)}$$

To obtain the value of the Jacobean transformation, the differential is with respect to x and we get $\frac{dy}{dx} = \frac{1}{x}$ and then substituting in equation (7).

$$f(y,\mu,\sigma) = \frac{1}{\sigma} exp\left(\frac{y-\mu}{\sigma}\right) * e^{-exp\left(\frac{y-\mu}{\sigma}\right)} * \left[(1+k) - 2k\left(1 - e^{-exp\left(\frac{y-\mu}{\sigma}\right)}\right)\right] \quad \textbf{(8)}$$

By substituting in equation (8), let $Z = \frac{y-\mu}{\sigma}$ & $\sigma Z = y - \mu$ & $y = \mu + \sigma Z$

$$f(z) = \frac{1}{\sigma} exp(z) * e^{-exp(z)} * \left[(1+k) - 2k\left(1 - e^{-exp(z)}\right)\right] \quad \textbf{(9)}$$

where $-\infty < \mu < \infty$ and σ > 0. The parameters μ and σ are the location and scale parameters of the **T-W** distribution. Hereafter, the pdf in (9) is called as log-transmuted weibull regression model. The corresponding survival function is given by:

$$S\,(y,\mu,\sigma) \; = 1-(1+k)\,(1-e^{-exp\left(\frac{y-\mu}{\sigma}\right)})+k\left[1-e^{-exp\left(\frac{y-\mu}{\sigma}\right)}\right]^2 \quad \textbf{(10)}$$

Consider the following regression model:

$$y_i = x_i^T\beta + \sigma_i z_i \quad \textbf{(11)}$$

where the response variable y_i has the density function is given in (8). The covariates are linked to location of y_i with identity link function $\mu = x^T\beta$.

Where $X = (x_1, x_2, \dots \dots \dots \dots . . x_p)^T$ is the model matrix consists of the observations of independent variables, and $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots \dots \dots . \beta_p)$ is the unknown regression coefficients.

Estimation of The Model Parameters

Let the random sample $(y_1, y_2, \dots, y_n)$ follow a T-W distribution and the response variable is defined as $y_n = min(x_i, c_i)$. Where c_i is the censoring time and x_i is the observed lifetime. Assume that the censoring times and lifetimes are independent. Let F and C are the sets representing the observed lifetimes and censoring times. The general formulation of the log-likelihood function for the model given in (8) is given by:

$$L(\theta) = \sum_{i=1}^{n} \delta_i \log\big(f(y)\big) + \sum_{i=1}^{n} (1-\delta_i)\log\,(s(y)) \quad \textbf{(12)}$$

$$\log\big(f(y)\big) = \delta_i log\left[\frac{1}{\sigma}\left(\frac{y-\mu}{\sigma}\right) *\; e^{-exp\left(\frac{y-\mu}{\sigma}\right)} \;*\left[(1+k)\;-2k(1-e^{-exp\left(\frac{y-\mu}{\sigma}\right)})\right]\right]$$

$$\sum_{i=1}^{n} log\big(f(y)\big) = \delta_i \left[\sum_{i=1}^{n} log\left(\frac{1}{\sigma}\right) + \sum_{i=1}^{n} log\left(\frac{y-\mu}{\sigma}\right) - \sum_{i=1}^{n} exp\left(\frac{y-\mu}{\sigma}\right) + \sum_{i=1}^{n} log\left[(1+k)\;-2k(1-e^{-exp\left(\frac{y-\mu}{\sigma}\right)})\right]\right]$$

$$\sum_{i=1}^{n} log\big(f(y)\big) = \delta_i \left[\sum_{i=1}^{n} log\left(\frac{1}{\sigma}\right) + \sum_{i=1}^{n} log\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right) - \sum_{i=1}^{n} exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right) + \sum_{i=1}^{n} log\left[(1+k)\;-2k(1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)})\right]\right]$$

$$\sum_{i=1}^{n} \log\,(s(y)) = (1-\delta_i)\sum_{i=1}^{n} log\left[1-(1+k)\,(1-e^{-exp\left(\frac{y-\mu}{\sigma}\right)})+k\left[1-e^{-exp\left(\frac{y-\mu}{\sigma}\right)}\right]^2\right]$$

$$\sum_{i=1}^{n} \log\,(s(y)) = (1-\delta_i)\sum_{i=1}^{n} log\left[1-(1+k)\,(1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)})+k\left[1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2\right]$$

To estimate the coefficients of the proposed regression model, we deferential equation No. (12) with respect to the regression coefficients.

First, Estimation of parameter $\boldsymbol{\beta_0}$:

By using partials' derivatives of log likelihood function of Transmuted Weibull at (12) with respect to $\boldsymbol{\beta_0}$ as the following:

$\frac{\partial L(\theta)}{\partial\beta_0} = \sum_{i=1}\frac{\partial \log(f(y))}{\partial\beta_0} + \sum_{i=1}^{n}\frac{\partial \log(s(y))}{\partial\beta_0}) = 0$

$$\sum_{i=1}^{n} log(f(y)) = (\delta_i)\left[\sum_{i=1}^{n} log\left(\frac{1}{\sigma}\right) + \sum_{i=1}^{n} log\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right) - \sum_{i=1}^{n} exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right) + \sum_{i=1}^{n} log\left[(1+k)\ -2k(1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)})\right]\right]$$

$$\frac{\partial\sum_{i=1}^{n} log(f(y))}{\partial\beta_0} = (\delta_i)\sum_{i=1}^{n}\left(\frac{\sigma}{y-\beta_0-x\beta_1}\right)*\left(\frac{-1}{\sigma}\right) + (\delta_i)\sum_{i=1}^{n} exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)*\left(\frac{-1}{\sigma}\right) + (\delta_i)\sum_{i=1}^{n}\left[\frac{+\frac{2}{\sigma}k(e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\ *-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}{(1+k)\ -2k(1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)})})\right]$$

$$\frac{\partial\sum_{i=1}^{n}\log(s(y))}{\partial\beta_0} = (1-\delta_i)\ \frac{(1+k)e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\left(-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)\right)\frac{-1}{\sigma} - \frac{1}{\sigma}e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)} * exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)*}{\sum_{i=1}^{n} log\left[1-(1+k)\ (1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}) + 1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]}$$

second; Estimation of parameter $\boldsymbol{\beta_1}$:

By using partials' derivatives of log likelihood function of Transmuted Weibull at (12) with respect to $\boldsymbol{\beta_1}$ as the following:

$\frac{\partial L(\theta)}{\partial\beta_1} = (\delta_i)\sum_{i=1}\frac{\partial \log(f(y))}{\partial\beta_1} + (1-\delta_i)\sum_{i=1}^{n}\frac{\partial \log(s(y)}{\partial\beta_1})=0$

$$\frac{\partial \log(f(y))}{\partial\beta_1} = \sum_{i=1}^{n}\left(\frac{(\delta_i)}{y-\beta_0-x\beta_1}\right)*\left(\frac{-x}{\sigma}\right) + \sum_{i=1}^{n} exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)*\left(\frac{-x}{\sigma}\right) + \sum_{i=1}^{n}\left[\frac{+\frac{2x}{\sigma}k(e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\ *-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}{(1+k)\ -2k(1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)})})\right]$$

$$\frac{\partial\sum_{i=1}^{n}\log(s(y)}{\partial\beta_1} = (1-\delta_i)\frac{(1+k)e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\left(-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)\right)\frac{-x}{\sigma} - \frac{x}{\sigma}e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)} * exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)*}{\sum_{i=1}^{n} log\left[1-(1+k)\ (1-e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}) + 1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]}$$

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

third; Estimation of parameter σ:

By using partials' derivatives of log likelihood function of Transmuted Weibull at (12) with respect to $\boldsymbol{\sigma}$ as the following:

$\frac{\partial L(\theta)}{\partial \sigma} = \sum_{i=1} \frac{\partial \text{og}(f(y))}{\partial \sigma} + \sum_{i=1}^{n} \frac{\partial \log(s(y)}{\partial \sigma})=0$

$$\frac{\partial \text{og}(f(y))}{\partial \sigma} = \sum_{i=1}^{n} \left(\frac{(\delta_i)}{y - \beta_0 - x\beta_1}\right) * \left(\frac{-1}{\sigma}\right) + \sum_{i=1}^{n} exp\left(\frac{y - \beta_0 - x\beta_1}{\sigma}\right) * \left(\frac{-(\delta_i)}{\sigma^2}\right) + \sum_{i=1}^{n} \left[\frac{+\frac{2}{\sigma^2} k(e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)} * -exp\left(\frac{y - \beta_0 - x\beta_1}{\sigma}\right))}{(1+k) \ - 2k(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)})}\right] (-\frac{(\delta_i)}{\sigma})$$

$$\frac{\boldsymbol{\partial \log(s(y))}}{\boldsymbol{\partial \sigma}} = (\mathbf{1} - \boldsymbol{\delta_i}) \sum_{i=1}^{n} \frac{\frac{\mathbf{1}}{\boldsymbol{\sigma^2}} * \left((\mathbf{1}+\boldsymbol{k})\ (\mathbf{1} - \boldsymbol{e}^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}) * \boldsymbol{exp}\left(\frac{\boldsymbol{y - \beta_0 - x\beta_1}}{\boldsymbol{\sigma}}\right) + \mathbf{2}\boldsymbol{k}\left[\mathbf{1} - \boldsymbol{e}^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right] * -\boldsymbol{exp}\left(\frac{\boldsymbol{y - \beta_0 - x\beta_1}}{\boldsymbol{\sigma}}\right)\right)}{\left[\mathbf{1} - (\mathbf{1}+\boldsymbol{k})\ (\mathbf{1} - \boldsymbol{e}^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}) + \boldsymbol{k}\left[\mathbf{1} - \boldsymbol{e}^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2\right]}$$

Therefore, the model of Transmuted weibull in (12) can be written as linear log location-scale regression model as following:

$$log\mu(y_i) = \beta_0 + \beta_1 x \qquad \textbf{(13)}$$

where y_i is the dependent variable. x is the independent variable, β_0 is the intercept, β_1 is the slope in which the relation between dependent and the independent variables are determined, and $\sigma_i z_i$ is the error term.

Residual analysis

Residual analysis is an important step of any regression analysis to check the sufficiency of the fitted model. If the fitted model is accurate for the data used, the residuals have to meet the distributional assumptions. Here, we used four kinds of residuals modified under T-W regression model including the Martingale Residual, Deviance Component Residual, Pearson Residuals and Cox and Snell Residual.

The Martingale Residual:

Is much used in the counting process These residuals are a symmetric and take maximum values (+1) and minimum values (-∞) **(Alamoudi, et al. 2017)**: we defined the martingale residual as:

$$\boldsymbol{r_{Mi}} = \boldsymbol{\delta_i} + \left(\int_0^y h(u)\, du\right)$$

where $\boldsymbol{\delta_i} = \begin{cases} \mathbf{0} & \boldsymbol{the\ i^{th} observation\ is\ censored} \\ \mathbf{1} & \boldsymbol{the\ i^{th} observation\ is\ uncensored} \end{cases}$

As we known $\int_{\mathbf{0}}^{y} \boldsymbol{h(u)}\ \boldsymbol{du} = \boldsymbol{log}\ [\boldsymbol{s(y_i)}]$

Then, the martingale residual can be reduced to:

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

Here, the martingale residual for the log T-W regression model takes the following form:

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

$$r_{Mi} = \begin{cases} 1 + \left[log \left[1 - (1+k)\left(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right) + k\left[1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2 \right] \right] & ,\delta_i = 1 \\ log \left[1 - (1+k)\left(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right) + k\left[1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2 \right] & ,\delta_i = 0 \end{cases}$$

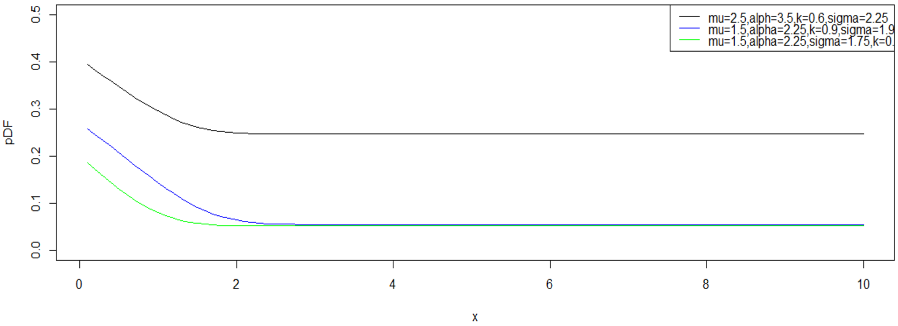

Figure (2). The martingale residual for transmuted weibull regression model

Deviance component Residual:

This residue was suggested to make the martingale residual more symmetric around zero **(Alamoudi, et al. 2017)**. The deviance component for the parametric regression model is given:

$$\hat{r}_{Di} \; = sign\;(\hat{r}_{Mi})\;\{-2\;[\hat{r}_{Mi} + log\;(1 - \hat{r}_{Mi})]\}^{1/2}$$

Where r_{Mi} is the martingale residual. * sign () function is a function that drives the (***+1***) values if the argument is positive and (***-1***) is negative. The deviance component residual for the Transmuted Weibull model is given by:

$$\hat{r}_{Di} = \begin{cases} \mathbf{sign}\;(\hat{\mathbf{r}}_{\mathbf{Mi}}) \left\{ -2 \left[\begin{array}{l} 1 + \left[log \left[1 - (1+k)\left(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right) + k\left[1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2 \right] \right] \\ + log \left[1 - \left(1 - \left[log \left[1 - (1+k)\left(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right) + k\left[1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2 \right] \right] \right) \right] \end{array} \right] \right\}^{1/2} & ,\delta_i = 1 \\ \mathbf{sign}\;(\hat{\mathbf{r}}_{\mathbf{Mi}}) \left\{ -2 \left[\begin{array}{l} 1 + \left[log \left[1 - (1+k)\left(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right) + k\left[1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2 \right] \right] \\ + \mathbf{log} \left(1 - 1 - \left[log \left[1 - (1+k)\left(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right) + k\left[1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2 \right] \right] \right) \end{array} \right] \right\}^{1/2} & ,\delta_i = 0 \end{cases}$$

Pearson Residuals:

are used to detect outliers. It depends on the idea of subtracting the mean and dividing by the standard deviation. The Pearson Residual knows the bounded regression of the inverse of the exponential distribution as follows:

$$r_i = \frac{y_i - \hat{\mu}}{\sqrt{\widehat{var(y)}}} \qquad \textbf{(14)}$$

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

Cox and Snell Residual:

Cox and Snell (1968) residual defined as follows:

$$e_i = -lin[1 - F(y_i, \beta)] \qquad \textbf{(15)}$$

Substituting in equation No. (15) for the value of the Cumulative Function, we get the Cox and Cox and Snell residual for Transmuted Weibull regression model as the following:

$$e_i = -lin\left[1 - \left[(1+k)\ \left(1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right) - k\left[1 - e^{-exp\left(\frac{y-\beta_0-x\beta_1}{\sigma}\right)}\right]^2\right]\ \right]$$

Numerical Experiment

simulation study

In this section, a simulation study is given to evaluate the performance of coefficients for the proposed regression model according to Transmuted Weibull regression model. All results were obtained from 1,000 Monte Carlo replications. The following steps is used in Monte Carlo simulation.

1. Generate the data from this model with $\beta_0 = 0.8, \beta_1 = 1.4, K = 1\ and\ \sigma = 1$
2. Generate $y \sim T - W(\mu, \sigma)$ **where** $\boldsymbol{\mu} = \beta_0 + \beta_1 x$.
3. Generate $x \sim U(0,1)$.
4. Generate $Z \sim U(0,1)$.
5. The sample sizes are taken as *n* = 20, 100, 200.
6. Each sample size is replicated 1000 times.
7. For each generated sample sizes, the biases, (AS) and (MSE) are evaluates at level the three censoring rates (20%, 30%, 40%).

The simulation results are reported in Table (1), Table (2) and Table (3). As seen from the results, the estimated biases, average of estimates (AS) and mean square error (MSEs) are near the desired value, zero.

Table (1) Simulation Study Results of (T-W) Regression model at Censoring Rate=20%

Censoring rate= 0.20	**n=20**			**n=50**			**n=100**		
Parameter	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**
$\widehat{K}$	.324	0.351	0.626	.320	0.210	0.524	.189	0.023	0.241
$\hat{\beta}_0$	0.976	0.176	0.334	0.675	0.125	0.210	0.741	0.112	0.110
$\hat{\beta}_1$	1.923	0.477	0.830	1.624	0.211	0.622	1.522	0.110	0.322
$\hat{\sigma}$	1.256	0.256	0.326	1.130	0.115	0.245	1.012	0.012	0006

Table (1); explains that **at censoring rate =0.20**, Average Estimation (AE), Bias, and Mean square error (MSE) are estimated at different sample size n=20, n=50, n=100 for different parameters ($\widehat{K}$, $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\sigma}$). The previous table indicates that the estimated AE, Bias, MSE are tends to desire value (zero) as be shown in figure

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

(1). When the sample size increases, the estimated AE, Bias, MSE decreases, this mean that the difference between true and estimated values for parameters declines up to desired value (zero).

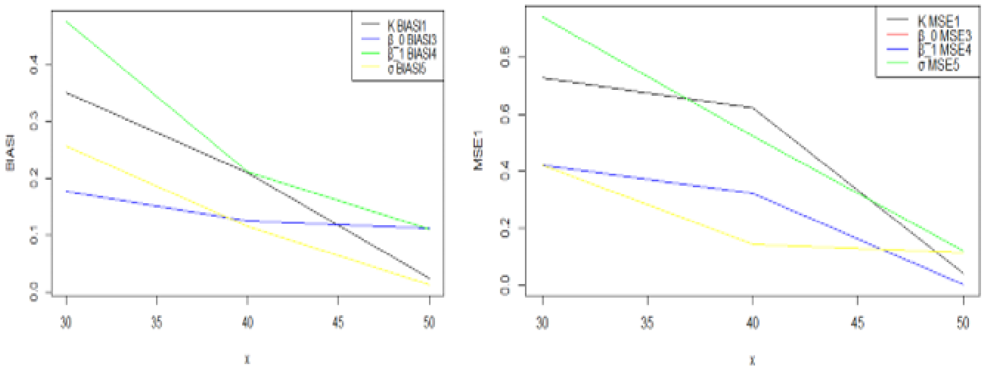

Figure (3). The Bias and MSE for different sample sizes at censoring rate (20%)

Table (2), Simulation Study Results of (T-W) Regression model at Censoring rate 30%

Censoring rate= 0.30	n=20			n=50			n=100		
Parameter	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**
$\widehat{K}$	.834	0.634	0.816	.623	0.423	0.629	.062	0.042	0.016
$\hat{\beta}_0$	0.843	0.312	0.147	0.712	0.213	0.131	0.613	0.113	0.013
$\hat{\beta}_1$	1.842	0.441	0.921	1.552	0.141	0.625	1.343	0.131	0.016
$\hat{\sigma}$	1.239	0.239	0.412	1.220	0.020	0.312	1.125	0.021	0.022

Table (2), Average Estimation (AE), Bias, and Mean square error (MSE) are estimated but at **censoring rate =0.30** with different sample size n=20, n=50, n=100 for different parameters ($\widehat{K}$, $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\sigma}$). Also, as the sample size increases, the estimated AE, Bias, MSE decreases, this mean that the difference declines up to desired value (zero) and this can be explained by the figure (2) that shown the patterns of relationships between the different sample sizes and the estimated AE, Bias, MSE.

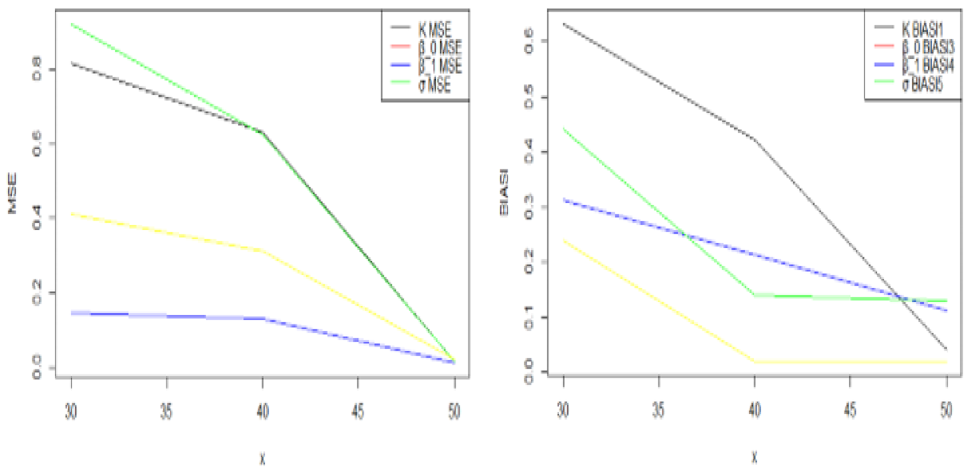

Figure (4). The MSE and Bias for different sample sizes at censoring rate (30%)

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

It is clear from the previous figure (2) that the mean square error and bias estimators goes to zero when the sample size increases.

Table (3) Simulation Study Results of (T-W) Regression model at censoring rate 40%

Censoring rate 0.40	n=20			n=50			n=100		
Parameter	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**
$\hat{K}$	2.456	0.456	0.562	2.342	0.342	0.413	2.023	0.023	0.004
$\hat{\beta}_0$	0.842	0.242	0.423	0.624	0.024	0.462	0.321	0.321	0.023
$\hat{\beta}_1$	0.795	0.195	0.423	0.675	0.075	0.049	0.657	0.057	0.054
$\hat{\sigma}$	1.423	0.423	0.162	1.125	0.125	0.625	1.025	0.025	0.004

Through Table (3), the Average Estimation (AE), Bias, and Mean square error (MSE) are estimated but at **censoring rate =0.40** with different sample sizes. Also, the results of estimators ensure that there is negative relationship between the sample size and estimated AE, Bias, MSE.

Table (4) coefficient of regression at different sizes of sample with different rates

	Rate20%			Rate30%			Rate50%		
	n=20	n=50	n=100	n=20	n=50	n=100	n=20	n=50	n=100
$\hat{\beta}_0$	0.976	0.675	0.741	0.843	0.712	0.613	0.242	0.624	0.321
$\hat{\beta}_1$	1.923	1.624	1.522	1.842	1.552	1.343	0.195	0.675	0.657
AIC	125.24	123.85	109.25	108.56	99.96	80.54	78.39	66.97	52.65
BIC	124.37	122.56	115.86	114.67	97.35	83.96	82.42	78.54	53.54
R^2	0.623	0.692	0.763	0.793	0.832	0.893	0.932	0.972	0.983

-Table (4); illustrates the estimated values for coefficient of regression at different sample sizes 20, 50 and 100 under different censoring rates = 20%, 30% and 50%.

- Also noticed that coefficient of determination R^2 increases when the censoring rates and sample size increases. At rate 50% with sample size 100, it is observed that R^2 =.983 at higher level than other rates and sample size as be shown in figure (3).

- also, it is observed from the previous table that goodness of fit criteria AIC, BIC declining with increasing censoring rates and sample size as be shown in figure (3).

-it is concluded that the higher the censoring rates and sample size, the higher the quality of regression model for the real data sets.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

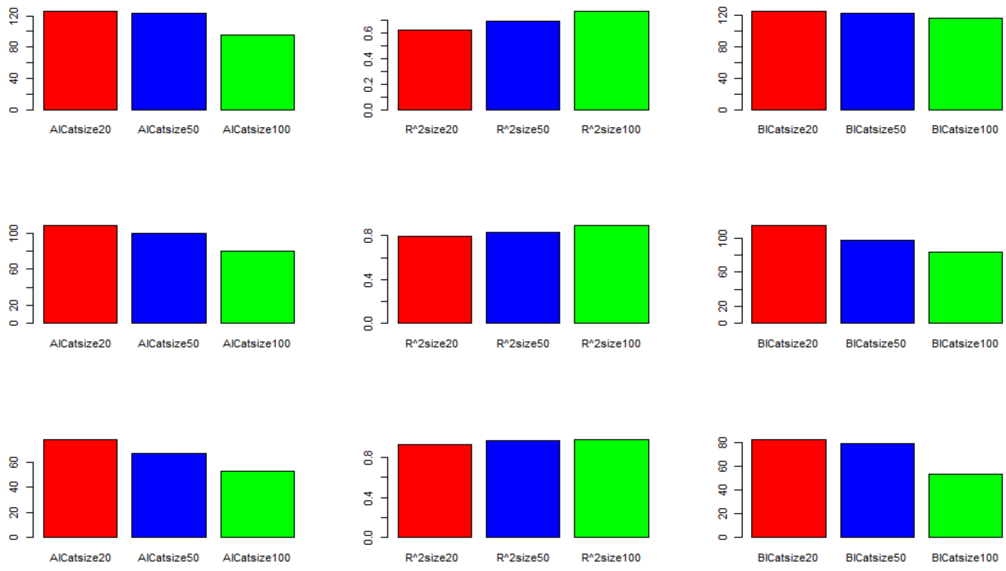

Figure (5). Some Statistical Criteria to Judge the Quality of the Model

From the previous figure (3), we notice that the coefficient of determination reaches its maximum value when the sample size increases, reaching .983 when the sample size was large, and that the statistical criteria through which the model was evaluated decrease with the increase in sample size. Therefore, it is necessary to use the proposed regression model with large samples.

Applications to Real data

This section provides the applications of Transmuted Weibull regression model on real data sets of firm's accounting items, real data represents some accounting items of seven companies listed on the stock exchange and each company we took data for 3 years. The accounting data involving one dependent variable refer to bad debt rate, and three independent variables involving the cash ratio, the quick liquidity ratio, and the Rate of loan maturities granted to the company.

Descriptive statistics:

Through the following table, we notice that the bad debt rate had a mean value of **.291429**, that the median had a value of **.22**, that the first quartile had a value of **.09**, and the third quartile had a value of **.46**.,and the cash ratio had a mean value **1.735831**, that the median had a value of **1.065453**, that the first quartile had a value of **.958781**, and the third quartile had a value of **1.582895**. and the quick liquidity had a mean value **.397745**, that the median had a value of **.104810**, that the first quartile had a value of **.054675**, and the third quartile had a value of **.216721**. and, the Rate of loan maturities had a mean value **.844688**, that the median had a value of **.657652**, that the first quartile had a value of **.068553**, and the third quartile had a value of **.899976.**

Table (5). Summary of Real Data Set

statistics	**y**	**X_1**	**X_2**	**X_3**
mean	.291429	1.735831	.397745	.844688
median	.220000	1.065453	.104810	.657652
min	.0100	.0239	.0117	.0000
max	0.780000	9.454016	5.388563	5.227127

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

Q1	.090000	.958781	.054675	.068553
Q3	.460000	1.582895	.216721	.899976

-The Relationship between bad debt rate, the liquidity ratio, the quick liquidity ratio, and the Rate of loan maturities:

The following represent the Correlation matrix between bad debt rate (y), the liquidity ratio (x_1), the quick liquidity ratio(x_2), the Rate of loan maturities (x_3) in order to determine the explanatory variable that is more related to the bad debt rate.

$$\begin{bmatrix} 1.0000000 & -0.7841519 & -0.6049875 & -0.6484035 \\ -0.7841519 & 1.0000000 & 0.5753677 & 0.5796666 \\ -0.6049875 & 0.5753677 & 1.0000000 & 0.6705442 \\ -0.6484035 & 0.5796666 & 0.6705442 & 1.0000000 \end{bmatrix}$$

Through the correlation matrix, we notice that the maximum value of the correlation between the bad debt rate (y), the liquidity ratio (x_1) reached **-0.78**, which is an inverse relationship as be shown in figure (4).

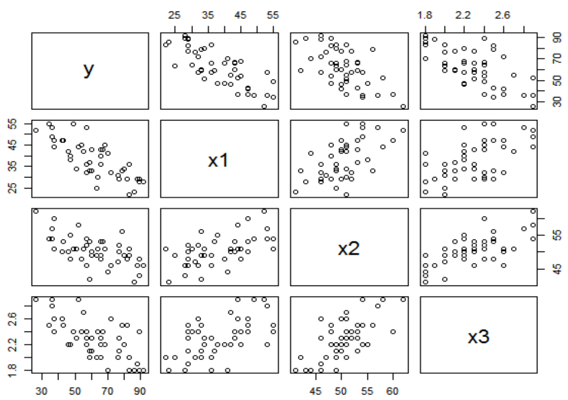

Figure (4). Correlation matrix between dependent and independent variables

Goodness- of – Fit Test of Real Data.

some statistical criteria, such as the Kolmogorov- Smirnov test, the Cramer- von statistic test, and Anderson-darling will be presented to test that the data follows Transmuted weibull distribution.

Table (4.6) Goodness- of – Fit Test of Real Data

Goodness of fit statistics	**Transmuted Weibull**	**Weibull**	**MG-Weibull**
Kolmogorov- Smirnov	**0.08707688**	0.0989061	0.164695
Cramer-Von statistic	**0.01712300**	0.0234912	0.097667
Anderson- darling	**0.14797663**	0.1734884	0.624825

From the previous table, we notice the following that the value of some statistical criteria, such as the Kolmogorov-Smirnov test, the Cramer- von statistic test, and Anderson- darling, for the Transmuted Weibull distribution is smaller than the rest of the other distributions, and thus the data represents the Transmuted Weibull distribution.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

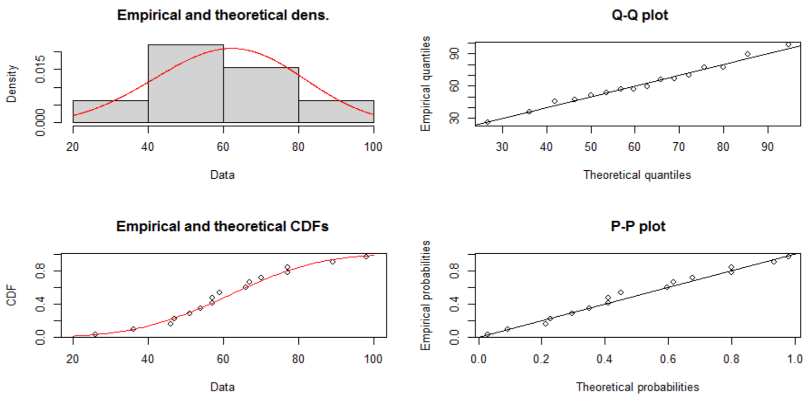

Figure (5). Goodness-of-fit

The previous figure show that the data follows the transmuted Weibull distribution, then, Fit of The Regression Model for Transmuted Weibull

Comparison between T-W With Kw-Lindley and MG Weibull regression model:

The comparison between T-W With Kw-Lindley and MG Weibull regression model by using R-program according to AIC, BIC and HQIC criteria as the following:

Table (7) The comparison between T-W With Kw-Lindley and MG Weibull regression

Regression model	$\hat{\beta}_0$	$\hat{\beta}_1$	HQIC	BIC	AIC	R^2
Transmuted Weibull	0.42352	0.6234	132.2815	125.3456	124.5423	0.8986
Kumaraswamy Lindley	0.03465	0.5234	134.8452	130.7256	128.4213	0.4239
LG Weibull	0.34214	0.4652	138.232	139.341	136.423	0.3492

Table (7) displays the results of previous criteria which show that the T-W regression model is more appropriate model compared to the KW-Lindley and MG Weibull regression model according to smaller values of HQIC AIC and BIC for T-W. Additionally, the higher coefficient of R-square for T-W compared to the KW-Lindley and LG Weibull regression model as be shown in Figure (6).

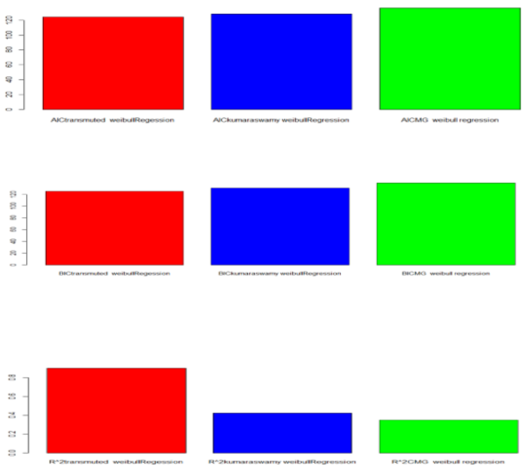

Figure (6) AIC, BIC and R^2 for transmuted weibull regression model, kumaraswamy weibull Regression and MG weibull regression Models

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume IX Issue VIII August 2024

From Figures (6) we notice that each of the statistical criteria AIC, BIC, HQIC for the transmuted-Weibull regression is smaller than the rest of the same criteria for the other regressions. Thus, the transmuted Weibull regression is the best regression model. The same is true for the coefficient of determination, we note that the coefficient of determination for the transmuted Weibull regression is larger than the rest of the other regressions. Thus, the transmuted Weibull regression is the best regression model. According to equation (13) the regression model of T-W is:

$$\boldsymbol{log\mu(y) = 0.42352 + 0.6234\, x\, (liquidity\, ratio)}$$

Global Influence

Is the diagnostic influence depending on case drop that represents one of the tools to perform sensitivity analysis introduced by Cook case drop is a famous method depending on the influence of deleting the i^{th}observation from the data on the parameter estimation. This method compares $\boldsymbol{\gamma_i}$ and $\boldsymbol{\gamma_{i-1}}$, where $\boldsymbol{\gamma_i}$ is MLE when the i^{th} observation is deleted from the original data. Then the ith case could be considered as influential observation if $\boldsymbol{\gamma_{i-1}}$ is far from $\boldsymbol{\gamma_i}$ (De Bastiani, F, et al 2015).

Table (8). Sensitive Analysis for regression

Case	$\widehat{\boldsymbol{\beta}}_0$	$\widehat{\boldsymbol{\beta}}_1$	AIC	BIC	HQIC	A Min -log (Likelihood)
Before	0.3211	0.6234	112. 842	115.432	114.9633	356.312
After	2.9364	0. 3544	137.574	138.542	135.213	324.624

Source: Developed by the researcher.

Cook =2 (A Min – log (Likelihood) before **-** A Min -log (Likelihood) After)

Cook = (356.312- 324.624) = 63.376

Noted that the cook is a positive value, meaning that the log-likelihood in small values (after deletion) is less than in large values (before deletion). This mean that the model is more sensitive to the data.

CONCLUSION

In this paper we introduce a new regression model for Transmuted Weibull distribution with application on real data set represent bad debit rate. The T-W regression model has wide contribution including flexibility, versatility and improved interpretability (reliability analysis, survival, environmental studies, finance, COVID 19). The method of maximum likelihood was utilized to estimate the model parameters. As well as the residuals analysis for T-W regression model are provided. Monte Carlo simulation is conducted to assess the performance of parameters of T-W regression model. We conclude that the higher the censoring rate and sample size, we noted that coefficient of determination increasing and AIC, BIC, HQIC decreasing as be shown in table (4.4). Based on goodness of fit criteria it is shown that T-W regression model is more fitted to real data set than other compared model the KW-Lindley and LG Weibull regression model. Finally, the Global influence is conducted to assess the T-W regression model before and after the i^{th} observation is deleted shown that the cook is a positive value, meaning that the log-likelihood in small values is less than in large values.

Highlights

i. This paper focuses on the process of creating regression models for probability distributions.
ii. The scope of this study is creating a regression model for the transmuted Weibull distribution.
iii. A new Regression model was applied to real data.

REFERENCE

1. Afify, A. Z., & Mead, M. E. (2017). On five-parameter Burr XII distribution: properties and applications. South African Statistical Journal, 51(1), 67-80. DOI:10.37920/sasj.2017.51.1.4
2. Afify, A. Z., Cordeiro, G. M., Yousof, H. M., Saboor, A., & Ortega, E. M. (2018). The Marshall-Olkin additive Weibull distribution with variable shapes for the hazard rate. Hacettepe Journal of Mathematics and Statistics, 47(2), 365-381. DOI:10.1007/s00009-017-0955-1
3. Afify, A. Z., Nofal, Z. M., & Butt, N. S. (2014). Transmuted complementary Weibull geometric distribution. Pakistan Journal of Statistics and Operation Research, 435-454.
4. Alamoudi, H. H., Mousa, S. A., and Baharith, L. A. (2017). Estimation and application in log-Fréchet regression model using censored data. Int. J. Adv. Stat. Probab, 5(1), 23-31.
5. AL-Kadim, K. A., & Mohammed, M. H. (2017). The cubic transmuted Weibull distribution. Journal of University of Babylon, 3, 862-876.
6. Aryal G.R, Tsokos C.P (2009) On the transmuted extreme value distribution with applications. Nonlinear Anal 71:1401–1407.
7. Aryal, G. R., & Tsokos, C. P. (2011). Transmuted Weibull distribution: A generalization of the Weibull probability distribution. European Journal of pure and applied mathematics, 4(2), 89-102.
8. Ashour, S. K., & Eltehiwy, M. A. (2013). Transmuted exponentiated modified Weibull distribution. International Journal of Basic and Applied Sciences, 2(3), 258.
9. Cox, D. R., & Snell, E. J. (1968). A general definition of residuals. Journal of the Royal Statistical Society: Series B (Methodological), 30(2), 248-265.
10. De Gusmao, F. R., Ortega, E. M., & Cordeiro, G. M. (2011). The generalized inverse Weibull distribution. Statistical Papers, 52(3), 591-619.
11. Dey, S., Kumar, D., Anis, M. Z., Nadarajah, S., & Okorie, I. (2021). A review of transmuted distributions. Journal of the Indian Society for Probability and Statistics, 22(1), 47-111.
12. Elbatal, I., & Aryal, G. (2013). On the transmuted additive weibull distribution. Austrian Journal of statistics, 42(2), 117-132.
13. Khan, M. S. (2019). Transmuted Modified Inverse Weibull distribution: Properties and application. Pakistan Journal of Statistics and Operation Research, 667-677.
14. Louzada, F., & Granzotto, D. C. (2016). The transmuted log-logistic regression model: a new model for time up to first calving of cows. Statistical Papers, 57, 623-640.
15. Merovci, F. (2013a). Transmuted lindley distribution. Int. J. Open Problems Compt. Math, 6(2), 63-72.
16. Merovci, F. (2013b). Transmuted rayleigh distribution. Austrian Journal of statistics, 42(1), 21-31.
17. Merovci, F., Elbatal, I., & Ahmed, A. (2013). Transmuted generalized inverse Weibull distribution. arXiv preprint arXiv:1309.3268.
18. Nofal, Z. M., Afify, A. Z., Yousof, H. M., Granzotto, D. C. T., & Louzada, F. (2018). The transmuted exponentiated additive weibull distribution: properties and applications. Journal of Modern Applied Statistical Methods, 17(1), 4.
19. Sarstedt, M., & Mooi, E. (2014). A Concise Guide to Market Research. The Process, Data, and 12. (p-p. 193-233). DOI:10.1007/978-3-642-53965-77 .
20. Shaw, W. T., & Buckley, I. R. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. arXiv preprint arXiv:0901.0434.