CrossMark

Regression Models According to Birnbaum-Saunders Distribution

Mohammad Nasim Wafa [1,*], Ziauddin Zia[1],
and Sayed Abdul Moqim Hussaini[2]

ABSTRACT

The Birnbaum–Saunders Distribution (BSD) is a positive Chile distribution that is often used to analyze life span data. Today, Birnbaum Sanders Distribution (BSD) has gained increasing popularity in various fields such as air pollution, business, earth sciences, industry and medicine. BSD is a special model due to its specificity. We have presented some results of reconstructed version of the BS distribution and a method of generating random numbers from this distribution. This article examines three regression models according on Birnbaum-Sanders distribution. The first model is derived from the distribution. The second model is derived from parameter re-averaging. The third model is derived through the logarithmic transformation of the response variable. The main purpose of this article is to compare the performance of three Birnbaum-Sanders regression models. Finally, an example of genuine data is presented to compare the regression models.

Keywords: Birnbaum-Sanders distributions, generate random numbers, life expectancy, regression model.

1. Introduction

The linear regression is one of the most widely used and effective statistical tools. This model is based on the method of the least squared error. It has been around for many years due to the relationship that the standard linear regression model has with various fields of knowledge, especially in statistics. It is noted that the model suggested by Gauss is not applicable in researches where the data do not provide indication of normality. This model includes the change of response to the display family related to the subject. Such a model is known as a generalized linear model (GLM).

There are models used to model data that has a positive range and is skewed, such as the gamma or inverse Gaussian distribution. In general, working with these distributions is not easy. Biparametric distribution is a distribution created from the univariate transformation of the standard normal distribution. It was introduced to analyze the data of Chula with positive amplitude. The parameters of this distribution are called scale and location. This distribution is more flexible than the known distributions for analyzing this type of data. Birnbaum-Sanders distribution is used to analyse the wear life of aluminum coupons, insurance data, wear time of bearings and data related to minerals in bones. We tried to study regressions based on BS distribution. There are three approaches to Bernbaum-Sanders distribution which have been presented by [1] and used to build flexible regression models. The first approach is based on the logarithmic transformation of the response variable.

Here we examine three approaches. The first approach is according on the parameterization of the BS distribution, which relates the mean of the distribution to the explanatory variables. The second approach is according on the logarithmic transformation of the response variable. The third approach directly relates the scale parameter of the BS distribution to the explanatory variables. It should be noted that the first and third approaches are equivalent, the difference mainly lies in applying a logarithmic transformation to the response variable. To compare the three regression methods, an

example with material wear data is provided. The continuation of this article is organized as follows. In section 2, the distribution as well as the relevant regression models and, in particular, an algorithm that condenses the necessary steps for computational implementation are presented. Birnbaum and Saunders is a very important distribution that can be obtained for life span, which arose from a physical problem. Birnbaum-Saunders (BS) distribution which actually describes and defines the entire period of time studied that the type of cumulative damage caused by the initiation and growth of a dominant crack dominates the threshold and causes failure. This distribution is more flexible than known distributions for analyzing this type of data. Outside of the context of the reliability and usability of BSD distributions implemented in various other disciplines, more details about these programs and an extensive bibliographic review can be found in [2]. BSD is defined in terms of the standard normal distribution through the random variable.

We mentioned earlier that the Birnbaum-Saunders (BS) distribution has a positive skew distribution that is often used to analyze life span data. Which is widely and generally used for regression analysis in this context that more than auxiliary variables are involved in the life test. The Birnbaum-Saunders distribution (BSD) is considered to be positively skewed in the majority. The Generalized Birnbaum-Saunders (GBS) distribution are new class of positively skewed models with heavier and lighter tails than the standard Birnbaum-Saunders (BS) distribution which is mainly used to study data analysis and longevity. However, all parts of the theoretical reasoning and interesting features of the GBS model, arguably, enable its application beyond the analysis of lifetime data. This distribution is one-sided and we mentioned earlier that it has positive deviation, positive provision and two parameters related to its scale and shape [3]–[5]. Interest in the BS distribution is due to physical theoretical arguments, its attractive features, and its relationship to the normal model. Although the origin of BS distribution is material fatigue, it should be remembered that BS distribution has many applications, but most of its applications have been used in the following cases: agriculture, commerce, pollution, finance engineering, forest, food, and textile industries, and informatics, medicine, insurance, mortality, microbiology, pharmacology, psychology, nutrition, queuing theory, quality control, water quality, toxicology and wind energy.

Linear regression is one of the most widely used statistical tools that started on the least square method. Due to the relationship that the normal linear regression model has had with various fields of knowledge for many years. However, the model proposed by Gauss is not applicable in research where the data do not provide evidence of normality. Motivated to solve this problem in [5], they defined a wide class of models that have a linear structure. These models include the fact that the distribution of the response variable belongs to the exponential family, which gives useful properties to the class of models. The model is known as a generalized linear model (GLM). Among GLMs there are members that are used to model data that have a positive range and are skewed, such as the gamma or inverse Gaussian distribution. In general, working with these distributions is not an easy task. Biparametric distribution is a distribution that is created from the univariate transformation of the standard normal distribution. It was introduced to analyze the data of Chula with registration range. The parameters of this distribution are called location and scale [7]. This distribution has more flexibility than known distributions for analyzing this type of data. Among the examples in which Birnbaum-Sanders distribution is used, we can mention the wear life of aluminum coupons, analysis of insurance data, wear time of bearings and data related to minerals in bones. In this study, we are looking for regression based.

Let's study and analyze the distribution of BS. In this regard, it is possible to use the three Birnbaum-Sanders distribution approaches that are presented according to [3] to build flexible regression models. The first approach is centered on the logarithmic revolution of the response variable [2]. The second approach is based on the parametrization of the BS distribution, which links the distribution mean to the explanatory variables [4]. The third approach, according to [2], directly defines the scale parameter of the BS distribution as Link explanatory variables. It should be noted that the first and second approaches are equivalent, the difference lies mainly in the application of logarithmic transformation for the response variable. To compare the three regression methods, an example with material wear data is presented.

2. Background of Birnbaum-Sanders Distributions

Data whose response falls in the interval (0, 1) such as indices, proportions, or rates appear very frequently in different fields of knowledge, mainly the areas of social sciences, engineering, economic sciences, and medicine. Some practical examples of these types of data are the proportion of patients who die from a certain disease or virus (SARS-CoV-2, Diabetes, HIV or Cancer) in a country or city; the Human Development Index or the illiteracy rate in a region or country; the proportion of deaths due to exposure to smoking or other exposure factors; the mortality rate from traffic accidents in a city;

the percentage of items that do not meet the minimum requirements in an assembly line; and the portion of income that a family spends on entertainment. For the analysis of data such as those described above, statistical methodologies developed from distributions with support in the interval (0, 1) are required. In this sense, several probability distributions and regression models have been proposed; see Ferrari and Cribari-Neto [1], Kumaraswamy [2], Martínez-Flórez *et al*. [3], [4], Kieschnick and Mccullough [5], Mazucheli *et al*. [6], [7].

The BS distribution depends on the standard normal model. Here we use the following expression:

$$\phi(z) = (2\pi)^{-\frac{1}{2}} = \exp\left(-\frac{1}{2}z^2\right) \ and \ \Phi(z) = \int_{-\theta}^{z} \phi(u)\,du; \quad z\epsilon R$$

We already mentioned that the BS distribution is continuous, unimodal and right-skewed, properties common to most lifetime distributions. The randomness of T, which follows the BS distribution, can be expressed as BSD (α, β). α is called the scale parameter and the β symptom is the place parameter. Using the $Z = 1/\alpha\left[(T/\beta)^{1/2} - (\beta/T)^{-1/2}\right]$ notation, Z follows the standard normal functions.

$$T = \frac{\beta}{4}\left[\alpha Z + \sqrt{(\alpha Z)^2 + 4}\right]^2$$

Here T~BS(α, β). Then, the cumulative distribution function (CDF) for the random variable T that follows the two-parameter Birnbaum–Saunders distribution is as:

$$F_T(t; \alpha, \beta) = \Phi\left[\left(\frac{1}{\alpha}\right)\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{2}{3}}\right\}\right], 0 < t < \infty \quad \alpha, \beta > 0$$

The density function of this distribution is expressed in the form:

$$F_T(t; \alpha, \beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta}\left[\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{2}{3}}\right\}\right]\exp\left[-\frac{1}{2a^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right], \quad t, \alpha, \beta > 0$$

In addition, the following characteristics are: $E(T) = \beta/2\left(\alpha^2 + 2\right)$, $var(T) = \beta^2/2\left(5\alpha^4 + 4\alpha^2\right)$, $\alpha T\sim$ BS $(\alpha, \alpha\beta)$ *for* $\alpha > 0$ and $T^{-1}\sim$BS (α, β^{-1}). These characteristics are useful for various statistical purposes such as generating random numbers, estimating parameters and modeling based on regression.

One of the main motivations to study the RBS distribution is easier way to obtain statistical properties which are not possible through the classical parametrization. Such a reparameterization is useful because, first, moment estimates for the original parameterization of the BS distribution do not have a closed-form, but this is possible [9] and, second, it allows a response variable to be modeled in its original scale [8], which is not possible with the parameterizations proposed until now. The objectives of this chapter are: (i) to provide some results on moments of a reparameterized version of the BS distribution and a generator of random numbers; (ii) to propose estimators for this reparameterization; (iii) to study the performance of these estimators; and (iv) to apply the results to real data. The studied estimators are based on ML, moment, MM methods and (GMM).

2.1. Log-BS-Distribution

True random variable of Y value with density is function in the form:

$$f_Y(y; \alpha, \beta, \sigma) = \frac{2}{\alpha\sigma\sqrt{2\pi}}\cosh\left(\frac{y - \gamma}{\alpha}\right)\exp\left[-\frac{2}{\alpha^2}\sinh^2\left(\frac{y - \gamma}{\alpha}\right)\right]$$

It is defined that $\cosh(x)$ *is displayed in the form of* $\cosh(x) = (\exp^x + \exp^{-x})/2$. *Here* α *is the shape parameter*, σ *is the scale parameter and* γ *is the place parameter*.

2.2. Regression Model

In this section, we explain regression model. In this model, the methods of determining the validity of regression models include comparison of predictions, analyzing lifetime data, model coefficients with theory, and collecting new data to check model predictions. Attempts have been made to compare the results with theoretical model calculations and data division or cross-validation in which the majority of data and sometimes a part of the data are used to estimate the model coefficients and the rest of the data are used to measure the accuracy of the model prediction. Below is a descriptive overview of these methods.

Multivariate regression models (MRMs) are widely used in engineering and finance research, health agricultural and biological sciences, mainly for two purposes: prediction and effect estimation. When building a regression model, various strategies and ways have been recommended:

A) An appropriate statistical method that matches the data structure will be used.
B) Prevent or overcorrect the model.
C) By limiting the number of variables based on the number of events, an adequate sample size should be ensured.
D) Always evaluate the performance of the final model against calibration and detection criteria. If resources allow, validate the prediction model on external data.
E) Be aware of the pitfalls associated with automated variable selection methods (such as stepwise).

2.3. BS Regression Model

If $t_1, t_2, t_3, \ldots .t_n$ are independent random variables, then $T_i \sim BS(\alpha, \theta_i)$.

In this case, the observations corresponding to them are in the form of $t = (t_1, t_2, t_3, \ldots .t_n)^T$.

In addition, p is a deterministic variable with values $x = (1, x_1, x_2, \ldots, x_p)^T$, and the vector of regression coefficients $\beta = (\beta_0, \beta_1, \beta_3, \ldots, \beta_p)^T$ is considered, so that $\theta_i = \exp(x_i^T\beta)$.

The BS regression model studied by Wafa *et al.* (2019) [9] can be written in the form:

$$T_i = \theta_i \phi_i = exp\left(x_i^T \beta\right) \phi_i$$

where $\phi_i \sim BS(\alpha, 1)$ and $T_i \sim BS(\alpha, \theta_i)$.

2.4. Log-BS Regression Model

If $t_1, t_2, t_3, \ldots .t_n$ are independent random variables, then $T_i \sim BS(\alpha, \theta_i)$.

In this case, $t = (t_1, t_2, t_3, \ldots .t_n)^T$ are the observations corresponding to them.

Furthermore, $Y_1 \sim \log BS(t_1), Y_2 \sim \log BS(t_2), \ldots .Y_n \sim \log BS(t_n)$ that is $Y_i \sim \log BS(\alpha, \mu_i)$, where $\mu_i = \log(\theta_i)$ and $y = (y_1, t_2, y_3, \ldots y_n)^T$. By applying the logarithm to the model, we have regression:

$$Y_1 \sim \log BS(T_i) = \mu_i + \epsilon_i = x_i^T\beta + \epsilon_i, i = 1, 2, 3, 4, 5, \ldots .n.$$

where $\mu_i = x_i^t\beta, \epsilon_i = \log(\phi_i) \sim \log{-}BS(\alpha, 0)$, β and x_i are defined as in model (BS Regression model).

2.5. A New Parameterization for the BS Distribution

The RBS distribution is reparameterized with respect to the original parameterization by $\alpha = \sqrt{2/\delta}$ and $\beta = \delta\mu/[\delta+1]$, such that $\delta = 2/\alpha^2$ and $\mu = \beta[1+\alpha^2/2]$, where $\delta > 0$ and $\mu > 0$ are shape and mean parameters, respectively. Motivations and justifications for this new parameterization are detailed in the next section. Thus, if $Y \sim RBS(\mu, \delta)$, then its PDF is given by:

$$f(y; \mu, \delta) = \frac{\exp\left(\frac{\delta}{2}\right)\sqrt{\delta+1}}{4\sqrt{\pi\mu}y^{\frac{3}{2}}}\left[y + \frac{\delta\mu}{\delta+1}\right]\exp\left(-\frac{\delta}{4}\left[\frac{\{\delta+1\}y}{\delta\mu} + \frac{\delta\mu}{\{\delta+1\}y}\right]\right), y > 0$$

From one can note that RBS and standard normal RVs are related by:

$$Y = \frac{\delta\mu}{\delta+1}\left[\frac{Z}{\sqrt{2\delta}} + \sqrt{\left(\frac{Z}{\sqrt{2\delta}}\right)^2 + 1}\right] \ and$$

$$Z = \sqrt{\frac{\delta}{2}}\left[\sqrt{\frac{\{\delta+1\}Y}{\mu\delta}} - \sqrt{\frac{\mu\delta}{\{\delta+1\}Y}}\right]$$

Thus, from the above equation, the cumulative distribution function (CDF) and the quantile functin (QF) of $Y \sim RBS(\mu, \delta)$ are given as:

$$F(y; \mu, \delta) = \Phi\left(\sqrt{\frac{\delta}{2}}\left[\sqrt{\frac{\{\delta+1\}Y}{\mu\delta}} - \sqrt{\frac{\mu\delta}{\{\delta+1\}Y}}\right]\right), y > 0, and$$

$$y(q; \mu, \delta) = F^{-1}(q) = \frac{\delta\mu}{\delta+1}\left[\frac{z(q)}{\sqrt{2\delta}} + \sqrt{\left\{\frac{z(q)}{\sqrt{2\delta}}\right\}^2 + 1}\right] 0 < q < 1,$$

where z(q) is the q_{th} quantile of the standard normal distribution and $F - 1$ is the inverse CDF of Y. The hazard rate function of Y is defined by:

$$h(y; \mu, \delta) = \frac{\exp\left(\frac{\delta}{2}\right)\sqrt{\delta+1}}{4\sqrt{\pi\mu}y^3}\left[y + \frac{\delta\pi}{\delta+1}\right]\frac{\exp\left(-\frac{\delta}{4}\left[\frac{\{\delta+1\}y}{\mu\delta} + \frac{\mu\delta}{\{\delta+1\}y}\right]\right)}{\Phi\left(-\sqrt{\frac{\delta}{2}}\left[\sqrt{\frac{\{\delta+1\}t}{\mu\delta}} - \sqrt{\frac{\mu\delta}{\{\delta+1\}y}}\right]\right)}, y > 0$$

2.6. Regression Models Based on BS Distributions

Rieck and Nedelman defined that if $Y \sim BS(\alpha, \beta)$, then $Z = \log(Y)$ follows a log-BS distribution with shape parameter and location parameter $\gamma = log(\beta) \in R$, denoted by $Z \sim \log - BS(\alpha, \gamma)$. In this regression model, the original response must be transformed to a logarithmic scale. Thus, although in this scale one is modeling the mean, say $\gamma = log(\beta)$, in the natural scale one is modeling $\beta = exp(\gamma)$, which in the BS case corresponds to the median.

Reference [8] proposed a new approach to the BS regression model. Formally, $Y_1, \ldots, Y_n$ are independent random variables, where $Y_i \sim RBS(\mu_i, \delta)$, for $i = 1, \ldots, n$, and $y = [y_1, \ldots, y_n]^\top$. Then, a statistical model based on by the systematic can be defined as:

$$g(\mu_i) = \eta_i = x_i^\top \beta, i = 1, 2, \ldots, n,$$

where $\beta = [\beta_1, \ldots, \beta_p]^\top$, for $p < n$, is a vector of unknown parameters to be estimated, and $x_i\top = [1, x_{i2}, \ldots, x_{ip}]$ represents the values of **p** regressors, such that $\mu_i = g^{-1}(x_i^\top \beta)$, with $g^{-1}(\cdot)$ being the inverse function of $g(\cdot)$. In the model given in the link function $g : R \to R+$ is strictly monotone, positive, and at least twice differentiable; for example, $g(\mu) = log(\mu)$ or $g(\mu) = \sqrt{\mu}$.

Formally, $Var[Y_i]$ is a function of μ_i and, consequently, of the regressors x_i. Then, because we are modeling the mean based on a particular structure, we are also modeling the variance. Therefore, situations in which the variance is not constant can be analyzed by using this model. The log-likelihood function of the model given in $(g(\mu_i) = \eta_i = x_i^\top \beta)$ for $\theta = [\beta^\top, \delta]^\top$ is $\ell(\theta; y) = \sum n_i = \ell_i(\mu_i, \delta; y_i)$, where

$$\ell_i(\mu_i, \delta; y_i) = \frac{\delta}{2} - \frac{1}{2} log \left(\frac{\{\delta + 1\} \mu_i y_i^3}{\{\delta y_i + y_i + \delta \mu_i\}^2} - \frac{\{\delta + 1\} y_i}{4\mu_i} - \frac{\mu_i \delta^2}{4\{\delta + 1\} y_i} \right)$$

The score functions for βj, with $j = 1, \ldots, p$, and δ are, respectively, given by:

$$U\beta_j(\theta) = \frac{\partial \ell(\theta)}{\partial \beta_j} = \sum_{1=1}^n \left[-\frac{1}{2}\mu_i + \frac{\delta}{\{\delta y_i + \delta y_i + \mu_i\}} + \frac{\{\delta + 1\} yi}{4\mu_i} - \frac{\delta^2}{4\{\delta + 1\} y_i} \right] \frac{1}{g'(\mu_i)} x_{ij}$$

$$\sum_{i=1}^n [y_i^* - \mu_i^*] a_i x_{ij}$$

and

$$U\delta(\theta) = \frac{\partial \ell(\theta)}{\partial \delta} = \sum_{i=1}^n \frac{\delta}{2\{\delta + 1\}} + \frac{\{y_i + \mu_i\}}{\{\delta y + y_i + \delta \mu_i\}} - \frac{y_i}{4\mu} - \frac{\delta\{\delta + 2\}\mu_i}{4\{\delta + 1\}^2 y_i} = \sum_{i=1}^n [y_i^* - \mu_i^*]$$

where $g'(\cdot)$ is the derivative of $g(\cdot)$. In matrix form, we have:

$$U_\beta(\theta) = X^\top A [y^* - \mu^*] \text{ and } U_\delta(\theta) = 1^\top [y^\star - \mu^*]$$

where $y_i^* = \frac{\delta}{\{\delta y + y_i + \delta \mu_i\}} + \frac{\{\delta + 2\} y_i}{4\mu_i^2} - \frac{\delta^2}{4\mu_i}, \mu_i^** = \frac{1}{2\mu_i}, y_i^\star = \frac{\{y_i + \mu_i\}}{\{\delta y + y_i + \delta \mu_i\}} - \frac{y_i}{4\mu_i} - \frac{\delta\{\delta + 2\}\mu_i}{4\{\delta + 1\}^2 y_i},$
$\mu_i^\star = -\delta/2[\delta + 1] = I = [1, \ldots, 1]^\top$ is a vector of ones with **n** elements, $X = (x_1, \ldots, x_n)$, with x_i given in $(g(\mu_i) = \eta_i = x_i^\top \beta)$, for $i = 1, \ldots, n, [y^* - \mu^*] = [\{y_1^* - \mu_1^*\}, \ldots, \{y_n^* - \mu_n^*\}]^\top, [y^* - \mu^*] = [\{yy_1^* - \mu_1^*\}, \ldots, \{y_n^* - \mu_n^*\}]^\top$ and $A = a_i \delta_{ij}^n = \text{diag}(a)$ is $n \times n$ matrice, with $\mathbf{a} = [\mathbf{a}1,...,\mathbf{an}]\top$ with ™**i j n** being the Kronecker Delta that is a function of the 3 arguments i, j and **n**. If i and j are equals then the function takes the value 1, otherwise is equal to zero. The value n is the order of the square matrix.

Thus, the score vector is $U(\theta) = [U_\beta^\top(\theta), U_\delta(\theta)]^\top$.

2.7. RBS Regression Model

If $t_1, t_2, t_3, \ldots t_n$ are independent random variables then $T_i \sim RBS(\eta_i, \sigma)$, $i = 1, 2, 3, 4, 5, \ldots n$ and $t = (t_1, t_2, t_3, \ldots t_n)^T$ are the observations corresponding to them. RBS regression model density function liquor has a systematic component in the form:

$$h(\eta_i) = \eta_i = x_i\beta, i = 1, 2, 3, \ldots n$$

In the above expression, $\eta_i = h^{-1}(x_i^T \beta)$ and $h : R \longrightarrow R^+$.

A function is a relation that is strictly uniform, positive and at least twice differentiable [6].

In this paper, $h(\eta) = log(\eta)$ is used. It is noteworthy that the variance of Y_i is a function of η_i. Therefore, although the mean is being modeled, the variance is also modeled by the simple fact that $var(Y_i) = \eta_i^2/\Phi$ is modeled.

TABLE I: Descriptive Statistics of the Response Variable (Losing Time)

Minimum	Mean	Median	SD	CS	CV	Maximum
0.013	3.76	0.85	7.43	3.2	198.2	37.5

TABLE II: Estimates, ML, Standard Errors, and P-values of Specified Models for Downtime Data

Model	Parameter	Estimate	Standard error	p-value	BIC	AIC
BS	α	1.279	0.1438	<0.001		
	β_0	0.0978	0.1707	0.566	130.18	125.24
	β_1	−14.1163	1.5714	<0.001		
Log-BS	α	1.279	0.1438	<0.001		
	β_0	0.0977	0.1707	0.566	134.30	129.11
	β_1	−14.1163	1.5714	<0.001		
RBS	α	1.22	0.2748	<0.001		
	β_0	0.6961	0.1935	<0.001	130.18	125.24
	β_1	−14.1170	1.5718	<0.001		

2.7.1. *Example of Material Exhaustion Data*

The presented regression are applied on a real data set related to the failure time of ten thousand hardened steel pieces placed at four different stress levels. These data were taken from laboratories.

Princeton Mobil Research and Development Company was acquired at Princeton University in the United States of America.

Table I provides descriptive statistics for the downtime data set, which includes items such as central statistics, Coefficient of variation (CV), standard deviation (SD) and coefficient of skewness (CS).

Table I provides descriptive statistics for downtime data sets, including Standard deviation (SD) Coefficients Variation (CV), and Coefficient of Skewness (CS).

Table II shows the results of estimates, corresponding standard errors (CSE) and $p-values$ of the $t-test$ for the parameters of the log-BS, BS and RBS regression model. In addition, AIC and BIC values associated with these models are reported. It can be seen from Table II that the estimates of β_1 are completely similar. However, the estimates of β_0 are different. The log-BS and BS models lower values than those observed for the $log-BS$ model. That is, to fit the model, RBS and BS models are more suitable. Log-BS had Equal values of 0.0977, while the RBS model has an estimate equal to 0.6961.

It is noteworthy that the width from the origin in the log-BS and BS models is not statistically significant at the 5% level.

This means that this parameter is not influential in the model. In addition, the BIC and AIC values for the BS and RBS models are equal, but these values are lower than the values observed for the $log-BS$ model. That is, RBS and BS models are more suitable for fitting the model.

3. Conclusion

In this research, an attempt has been made to compare three Birnbaum-Sanders regression models. These models are based on Birnbaum-Sanders, Birnbaum-Sanders and Birnbaum-Sanders logarithms. According to the real data related to the failure time, the values of AIC and BIC for RBS and BS models are equal to and less than the values observed for the log-BS model. That is, the fit of RBS and BS models are more suitable for these data. Therefore, three Birnbaum-Sanders regression models are compared and the numerical consequences of the performance of these regressions are reported and show that different models can be used when dealing with regression data with positive amplitude and following a skewed distribution to the right is appropriate.

References

[1] Ferrari S, Cribari-Neto F. Beta regression for modelling rates and proportions. *J Appl Stat*. 2004 Aug 1;31(7):799–815.
[2] Kumaraswamy P. A generalized probability density function for double-bounded random processes. *J Hydrol*. 1980 Mar 1;46(1–2):79–88.
[3] Martínez-Flórez G, Tovar-Falón R. Regression models based on the unit Sinh-normal distribution. *Math*. 2021;9:1231.
[4] Martínez-Flórez G, Azevedo-Farias RB, Tovar-Falón R. New class of unit-power-skew-normal distribution and its associated regression model for bounded responses. *Math*. 2022;10:3035.
[5] Kieschnick R, Mccullough BD. Regression analysis of variates observed on (0,1). *Stat Model*. 2003;3:193–213.
[6] Mazucheli J, Menezes AFB, Ghitany ME. The unit-Weibull distribution and associated inference. *J Appl Probab Stat*. 2018;13:1–22.

[7] Wafa MN, Hussani SA, Pazhman J. Evaluation of students' mathematical ability in Afghanistan's schools using cognitive diagnosis models. *Eurasia J Math Sci Technol Educ*. 2020;16(6):57–64.

[8] Wafa MN. Assessing school students' mathematic ability using DINA and DINO models. *Int J Math Trends Technol*. 2019;65(12):153–65.

[9] Wafa MN, Zia Z, Frozan F. Consistency and ability of students using DINA and DINO models. *Eur J Math Stat*. 2023 Jul 6;4(4):7–13.