Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq

Robust Estimates for One-Parameter Exponential Regression Model

Ahmed Joudah Irshayyid
Ministry of Agriculture, Bureau of the Minister,
Baghdad, Iraq.
Ahmed.jawda1201a@coadec.uobaghdad.edu.iq

Rabab Abdulrida Saleh
Department of Statistics, College of Administration and Economics,
University of Baghdad,
Baghdad, Iraq.
rabab.saleh@coadec.uobaghdad.edu.iq

Received:4/12/2022 Accepted: 26/12/2022 Published: December / 2022

Abstract

One-parameter exponential regression is one of the most common and widely used models in several fields, to estimate the parameters of the one-parameter exponential regression model use the ordinary least square method but this method is not effective in the presence of outlier values, so robust methods were used to treat outlier values in one-parameter exponential regression model are to estimate the parameters using robust method (Median-of-Means, Forward search, M-Estimation), and the simulation was used to compare between the estimation methods with different sample sizes and assuming four ratios from the outliers of the data (10%, 20%, 30%, 40%). And the mean square error (MSE) was made to reach the best estimation method for the parameters, where the results obtained using the simulation showed that the forward search is the best because it gives the lowest mean of error. On the practical side, expenditure and revenue data were used to estimate the parameters of the one-parameter exponential regression ,where the data was tested, it appeared to have an exponential distribution, and the boxplot and (COOK) test were used to detect the outliers present in the real data. The Goodness of fit test was used for the one-parameter exponential model, and it was found that the data did not follow the normal distribution, and it was found that it suffers from the problem of heterogeneity of variance. The one-parameter exponential regression model for the expenditure and revenue data was estimated using the advanced search method because it was the best estimate.

Paper type Research paper

Keywords: Exponential regression model, robust estimates, Median-of-Means, Forward search, robust M-estimator.

1. Introduction

The one-parameter exponential regression model is one of the nonlinear regression models when the relationship between the variables is exponential, that is, the model parameters appear exponential or the variables appear exponential.(Lim and Peddada,2013),(Liu et al,2021)

Classical methods are used to estimate one-parameter exponential regression parameter, however, in the presence of outliers affecting the data, these methods are not useful for data analysis. Classical estimation methods can lead to misleading values for the one-parameter exponential regression coefficient, and the estimates may not be reliable, and the relationship between response and explanatory variables is oblique, then estimates are biased. To address this problem and to reduce errors resulting from estimating parameters, robust methods have been developed that are not easily affected by outliers, and we need these robust methods to estimate the one-parameter exponential regression model more accurately. (Tabatabai et al. 2014)

For instance, the principle of median-of-means (MOM) the researcher (Liu et al,2021) used the median-of-means method to estimate the coefficients of the one-parameter exponential regression model in presence of outliers, and it was compared with the least squares method, and the estimator (MOM) was more efficient. The researchers (Alsalem and Altaher,2019) used the Gauss-Newton method to estimate the parameters of the nonlinear least squares method, which is the most used method for estimating the exponential model coefficients, and in the presence of outliers, it may have a significant impact on the parameter estimates. Therefore, some robust nonlinear techniques were used as an alternative method for the classical least squares. This is the (M) method, and the two methods were compared, and the simulation results with all sample sizes and pollution rates indicated that the robust method M is the best. The researcher (Atkinson et al,2010) used the forward search method as a general and robust method, and this method includes dividing the data into small subtotals to discover the outliers values affecting the estimation of the parameters, then close observations are added to the model to monitor the estimation of the parameters with increasing the sub-group by measuring the errors and choosing the sub-group that has the least error as it was applied to the daily ozone concentration data in the extreme hot spot. The researcher (Atkinson et al, 2017) used the iterative forward search method in the regression using observation weights, which produces a flexible and informative form in the robust regression in the presence of outliers, and versions of forward plots are used to exhibit the presence of multiple outliers in a data set.

2. Materials and Methods

2.1 One-Parameter Exponential Regression Model

This model has been applied to Chinese GDP data by the researcher (Liu et al,2021), and the formula is as follows:

$$y_i = e^{\beta x_i} + \varepsilon_i \quad \dots (1)$$

where $y = [y_1 . y_2 . \dots . y_n]^T$ is (n*1) response vector. x_i : it is the explanatory variable in the model whose values are assumed to be known. β : is parameter unknown. ε_i: the error is the random part and it is a small value. This equation can

be solved by transforming by taking the logarithm of each side of the above equation:

$$\ln y_i = \beta x_i$$

Where: $y_i^* = \ln y_i$

$$y_i^* = \beta x_i + \varepsilon_i \quad \dots (3)$$

The model parameters can be estimated using the least squares method:

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i^* - \hat{y}_i^*)^2 \qquad \dots (4)$$

Substitute the estimated equation for $\hat{y}_i^*$ of equation (3) in equation (4)

$$h(\alpha.\beta) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i^* - \beta x_i)^2 \quad \dots (5)$$

We derive equation (5) according to (β) and then we equal the derivatives to zero in order to obtain an estimator of β.

$$\frac{\partial h(\alpha.\beta)}{\partial \beta} = 2\sum_{i=1}^{n} (y_i^* - \beta x_i) - x_i = 0$$

$$-2\sum_{i=1}^{n} x_i y_i^* + 2\sum_{i=1}^{n} x_i^2 \beta = 0$$

$$\beta = \frac{\sum_{i=1}^{n} x_i y_i^*}{\sum_{i=1}^{n} x_i^2} \quad \dots (6)$$

2.2 Robust Estimation Methods

Classical methods for parameter estimation in a one-parameter exponential regression model in the presence of outliers do not give efficient estimates; therefore, we use robust methods because they are not affected by the presence of outliers; it can be used even if the basic assumptions are violated or the assessment and test conditions are not met. There are several robust methods for estimating the parameters of the statistical model. The objective of estimation by robust methods is to find new statistical methods that are less sensitive and susceptible to abnormal values and to obtain the best estimation results.

2.2.1 The Median-of-Means Method (MOM)

This estimator is used systematically to build robust estimates. The MOM methods estimate parameters for β are using the following steps:

1- We separate $(y_i . x_i)_{.i=1.2\dots.n}$ into g of groups, the number of observations in each group is equal to T as in the equation below

$$T = \frac{n}{g} \quad \dots (7)$$

For ease of calculation, let's say n is always divisible by g, then we extract the number of groups g from the following equation: (Zhang and Liu, 2020)

$$g = 8 * \log\left(\frac{1}{\zeta}\right) \quad \dots (8)$$

Where $\{\zeta \in (0,1)\}$ and g are rounded to the nearest number in the positive direction. The structure of observations is always undefined, and diagnosing outliers is complicated, so ζ is determined from the following equation: (Liu et al,2021)

$$\zeta = \frac{C}{\sqrt{n}} \quad \dots (9)$$

C : A positive integer regardless of the outliers present and equal to.

$$c = tr(\Sigma) \dots (10)$$

$tr(\Sigma)$: The rank of the covariance and covariance matrix. (Lerasle, 2019)

2- We estimate the parameters β in each group j by a least squares method so that $j = 1.2. \dots .g$

$$\hat{\beta}^{(j)} = \frac{\sum_{j=1}^{T} x_j {y_j}^*}{\sum_{j=1}^{T} {x_j}^2} \quad \dots (11)$$

3- We find the estimator MOM

$$\hat{\beta}^{MOM} = \left(\hat{\beta}^{j}\right)^{T} \quad \dots (12)$$

So that the median of the parameter is found in the following way:

$$\hat{\beta}^{MOM} = median\left(\hat{\beta}^{1}.\hat{\beta}^{2}. \dots \quad .\hat{\beta}^{g}\right)^{T} \quad \dots (13)$$

2.2.2 Forward Search Method (FS)

It is a robust data analysis method used for parameter estimation, which starts from a small selected subset of data and monitors the effect of adding observations to the subset until all data is added at the end. Thus, we avoid the influence of outliers, and estimating parameters is a very reliable method, suppose we have $Z_i = (y_i . x_i)$, in which the one-parameter exponential regression model contains p=1 parameter; the forward search algorithm starts by choosing a subset of size m where (m=p+1), the subset of size m to m+1 is increased at each step to form a new subset of observations, where the subgroups are $p + 1 \leq m \leq n$ and Z_m is a vector of the observations and the sizes of the groups are as follows (Atkinson et al. 2017)

$$S_{i1\dots \dots im}^{(m)} = \{Z_m . Z_{m+1} . Z_{m+2} \dots \dots . Z_n\} \quad \dots (14)$$

The parameters are estimated by the least squares method for each sub-group so that we have a set of forward search parameters $\hat{\beta}_{Fc}$ where (Atkinson et al. 2010)

$$\hat{\beta}_{Fs} = \left(\hat{\beta}_m . \hat{\beta}_{m+1} . \quad \dots . \hat{\beta}_n\right) \quad \dots (15)$$

We find the random errors by replacing all the parameters of the subset with the main observations

$$\hat{y}_{im} = e^{\hat{\beta}_{mi*} x_i} \quad \dots (16)$$

$$e_{i_m} = y_i - \hat{y}_{im} \quad \dots i = 1.2. \quad \dots n \quad \dots (17)$$

We choose the parameters that have the least random errors. (RIAN and ATKINSO 2000)

$$e_i = \min\left(e_{i_m}\right) \quad \dots (18)$$

2.2.3 Robust M-method

It is one of the robust regression methods, which is considered one of the most well-known robust methods and is symbolized by the symbol (M), which deals with the problem of outliers by replacing the square of the residuals with the loss function ρ while keeping the main objective of the estimation method is to make the estimator as low as possible as in the following equation

$$\hat{\beta}_M = arg \min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{r_i}{\sigma}\right) \quad \dots (19)$$

The measurement parameter can be estimated using the median of the absolute deviations as in the following equation (Alsalem and Altaher,2019)

$$\sigma = \frac{medain|r_i - medain(r_i)|}{0 \cdot 6745} \quad \dots (20)$$

The residuals are standardized by the scale of scattering σ to ensure that the scale is equal; the loss function ρ is even and non-decreasing symmetric for positive values. It is easier to differentiate ρ with respect to β and find the root of the derivative, where ψ is the derivative of ρ, since the parameters depend on the weights, the weights are defined from the following equation (Ismail and Rasheed 2021)

$$w_i = \frac{\psi\left(\frac{r_i}{\sigma}\right)}{r_i} \quad \dots (21)$$

$$z_i = \frac{r_i}{\sigma} \quad \dots (22)$$

Where

$$\frac{\partial \sum_{i=1}^{n} \rho(z_i)}{\partial \beta} = \sum_{i=1}^{n} \psi(z_i) \quad \dots (23)$$

The estimated $\hat{\beta}_M$ is obtained by applying the nonlinear least squares method by multiplying both sides by the root of the weights (Marasovic et al 2016)

$$y_i = \eta(\beta; x_i) + \varepsilon_i \qquad , i = 1, 2, 3 \dots, n \quad \dots (24)$$

$\eta(\beta; x_i)$ is a nonlinear regression model function of known shape.

$$\rho^{-1} y_i = \rho^{-1} \eta(\hat{\beta}) + \rho^{-1}\, r_i \quad \dots (25)$$

where

$$\rho^{-1} = \sqrt{w_i} \quad \dots (26)$$

$$\rho^{-1} {r_i}^2 \rho^{-1} = \left(\rho^{-1} y_i - \rho^{-1} \eta(\hat{\beta})\right)^2 \quad \dots (27)$$

The weights equation can be written in matrix form as follows:

$$P\acute{P} = \begin{pmatrix} \frac{1}{w_1} & 0 & 0 \\ 0 & \frac{1}{w_2} & 0 \\ 0 & 0 & \frac{1}{w_n} \end{pmatrix} \quad \dots (28)$$

$$\rho^{-1} {r_i}^2 \rho^{-1} = \left(\rho^{-1} Y - \rho^{-1} \eta(\hat{\beta})\right)' \left(\rho^{-1} Y - \rho^{-1} \eta(\hat{\beta})\right) = 0 \quad \dots (29)$$

$$Y'WY - Y'W\eta(\widehat{\beta}) - YW\eta(\widehat{\beta}) + \eta(\widehat{\beta})'W\eta(\widehat{\beta}) \quad \dots (30)$$

where $\rho^{-1} * \rho^{-1} = W$

We derive the equation (29) with respect to β

$$\frac{\partial}{\partial \beta} = -2Y' * W * \dot{\eta}(\widehat{\beta}) + 2\dot{\eta}(\widehat{\beta})' * W * \dot{\eta}(\widehat{\beta})\widehat{\beta} = 0$$

$\dot{\eta}(\widehat{\beta})$: Derivative of a nonlinear regression model function

$$-2\dot{\eta}(\widehat{\beta})' * W * \dot{\eta}(\widehat{\beta})\widehat{\beta} = -2Y' * W * \dot{\eta}(\widehat{\beta})$$

$$\widehat{\beta}_M = \left(\dot{\eta}(\beta)' * W * \dot{\eta}(\beta)\right)^{-1} \dot{\eta}'(\beta) * W * Y \quad \dots (31)$$

where

$$\dot{\eta}(\beta) = x_i e^{\beta x_i} \quad \dots (32)$$

$$\widehat{\beta}_M = \left(x_i e^{\beta x_i}{}' * W * x_i e^{\beta x_i}\right)^{-1} x_i e^{\beta x_i}{}' * W * Y \quad \dots (33)$$

where W is (n*n) a square diagonal matrix whose elements are positive and w_i represents the weights. The model parameter estimated by M-robust method depends on the loss function by the weighted least squares method repeatedly re-weighted. The one-parameter exponential regression parameter cannot be obtained until the errors are obtained by the least squares method; therefore, the weighted least squares method is used frequently and according to the following steps: (Khan et al 2021)

i. **The one-parameter exponential regression line is estimated on the data by the least squares method by setting the counter I=0 to find the initial values of the estimates of the one-parameter exponential regression parameter $\beta^{(0)}$.**
ii. **The initial estimates of the weights are found from the residual values ε_{i0} that are obtained from the estimates of the one-parameter exponential regression from the least squares method of the data.**
iii. **The weight function is applied to the prime values of the residuals to create the prime weights $w(\varepsilon_{i0})$.**
iv. **The weighted least squares method is applied according to the equation (33), and thus we get the one-parameter exponential regression parameters $\beta^{(1)}$ from the initial iteration.**
v. **We extract the new weights $w(\varepsilon_{i1})$, which we get from extracting the residuals from the parameters of the initial weighted least squares method $\beta^{(1)}$.**
vi. **We continue with the solution and extract the rest of the parameters and stop if the condition is met:**

$$\beta^{(I)} - \beta^{(I-1)} \cong 0$$

The solution is considered close if the deviation in the estimates is not more than 0.01% from the previous iteration. The efficiency of the M estimators is about 95% compared to the ordinary least squares method according to the assumptions of the regression model.
Since the estimator by M-method depends on the loss function, Andrew's function has been used:

- **Andrew's Sin. function**

$$\psi(z_i) = \begin{cases} c \sin\left(\frac{z_i}{c}\right) & if\ |z_i| \le c\pi \\ 0 & if\ |z_i| > c\pi \end{cases} \quad \dots (34)$$

where c=1.339 or 1.5 or 2.1 (Khan et al 2021)

3. Discussion of Results

3.1 Simulation

The simulation experiments for this study included writing a number of programs in the language (MATLAB) to generate simulated data to compare methods with different sample sizes. (n1 = 25, n2 = 50, n3 = 100, n4 = 150) In generating data for random variables and different outliers ratio, each experiment was repeated 1000 times, taking into account the one-parameter exponential regression model that has been relied upon. The stages of the simulation experiment are described through the following steps:

i. **Set default values for parameters, and it is one of the important and basic steps on which exponential regression models depend, where B=2 was chosen.**
ii. **Generating the values of the independent variable that is assumed to be uniformly distributed over a period u(0,1)**
iii. **In the case of the data does not contain outliers the error is generated according to the following equation $\varepsilon = \tau \sim N(0,1)$ and in the case of the presence of outliers value the error is generated according to the following equation**
$$\varepsilon = (1-\tau) \sim N(0,1) + (\tau) \sim N(0,10) \quad \dots (38)$$
where the ratio of outliers are $(\tau = 10\%, 20\%, 30\%, 40\%)$
iv. **Calculating the dependent variable (y), the following one-parameter exponential regression model is applied**
$$y_i = e^{\beta x_i} + \varepsilon_i$$
v. **The estimation methods are compared on the statistical error scale of error squares. (MSE).**

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \quad \dots (39)$$

Table (1): It shows the (MSE) values of the one-parameter exponential regression model

size	Outliers ratio	MSF			
		ols	MOM	FS	M
N=25	0%	5.415e-5	2.465e-5	3.721e-5	1.783e-5
	10%	0.006832	0.007432	0.004249	0.005255
	20%	0.006868	0.007432	0.003772	0.005169
	30%	0.006961	0.007432	0.003772	0.005163
	40%	0.006713	0.006976	0.003772	0.005039
N-50	0%	5.701e-5	5.660 e-5	5.572e-5	7.761e-5
	10%	0.005611	0.005621	0.005568	0.005866
	20%	0.005589	0.005621	0.005483	0.005878
	30%	0.005577	0.005621	0.005484	0.005876
	40%	0.005597	0.005621	0.005484	0.005831
N=100	0%	7.825e-7	8.592e-7	6.244e-7	6.822e-6
	10%	7.069e-5	7.268e-5	6.264e-5	1.078e-4
	20%	6.964e-5	7.268e-5	5.667e-5	1.090e-4
	30%	6.971e-5	7.312e-5	5.667e-5	1.092e-4
	40%	6.895e-5	7.037e-5	5.667e-5	1.085e-4
N=150	0%	5.246e-8	5.060e-8	4.196e-8	6.180e-7
	10%	4.963e-6	4.977e-6	4.664e-6	8.304e-6
	20%	4.862e-6	4.977e-6	4.028e-6	8.437e-6
	30%	4.759e-6	4.926e-6	4.527e-6	8.583e-6
	40%	4.648e-6	4.926e-6	4.039e-6	8.499e-6

It was noted that values (MSE) were close to all estimation methods and for all the ratio of outliers, the methods can be considered equivalent, and all methods are characterized by the accuracy of the assessment because they gave reliable results.

It is clear from Table (1) that the results of the sample size (n = 25) and through the criterion (MSE) of the one-parameter exponential regression model shows that the forward search was the best estimation method for outliers ratio (10%, 20%, 30%, 40%) because it outperformed all methods and achieved the lowest level (MSE), while the M method was the best estimation method for the outliers ratio (0%) because it achieved the least mean square error (MSE); the simulation results clearly indicate the efficiency and robust of the forward search method and the M method with respect to methods other in the sample size (n = 25) and the presence of influential outliers.

It is clear from Table (1) the results of the sample sizes (n = 50), (n = 100) and (n = 150) ,and through the (MSE) criterion of the one-parameter exponential regression model shows that the forward search method (FS) was the best estimation method for all the ratio of outliers because it outperformed all methods and achieved less (MSE), and the simulation results indicate the efficiency and robustness of the forward search (FS) relative to other methods at the sample sizes (n = 50), (n = 100) and (n = 150) and the presence of influential outliers.

3.2 Real Data Application

In order to achieve the objective of the research and verify the performance of robust methods for treating outliers in real data, this part is devoted to the study of real data, especially for revenues (x) revenues are defined as all amounts collected by the government during a certain period of time, whether they are revenues from taxes, fees, grants or loans. And expenditures (y) are defined as all expenditures incurred by the government during a certain period of time, obtained from the Ministry of Finance and then apply the estimation methods in the one-parameter exponential regression model to the real data to reach the best method through the mean squares error (MSE). A sample size (155 views) was selected from January 2009 to November 2021, where the expenditures variable (y) was tested by (easyfit) program, where the distribution of the dependent variable has an exponential distribution with a parameter equal to $\lambda = 0.03$ as in Figure (1).

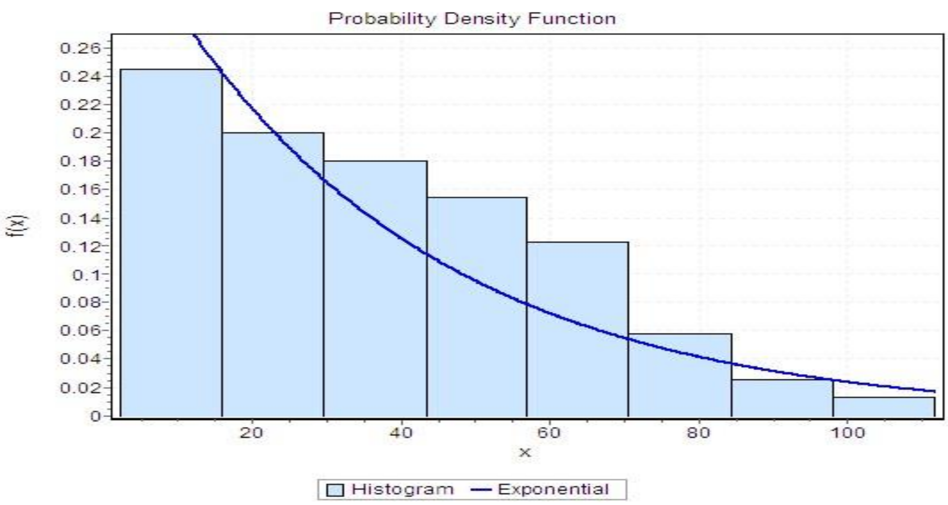

Figure (1): shows the exponential distribution of the dependent variable for (n = 155)

By plotting the boxplot in Figure (2), we notice that there are outliers in the real data for variable y and variable x.

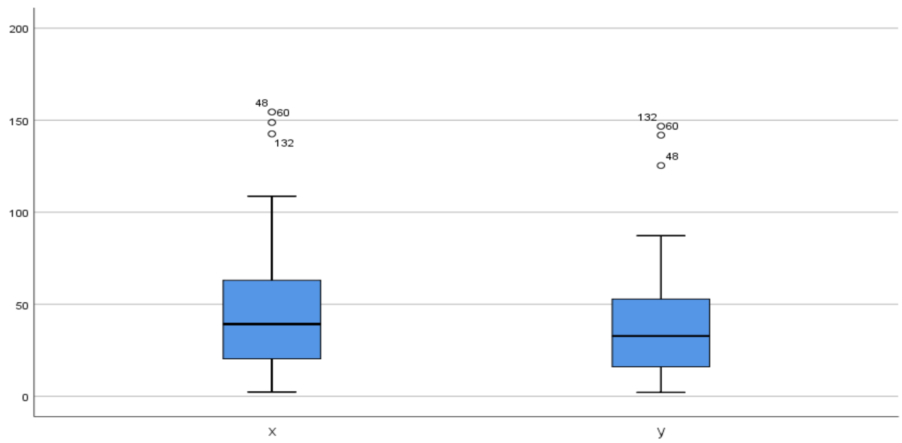

Figure (2): shows the outliers of the dependent variable and the independent variable

Some tests and measures have been conducted for the one-parameter exponential regression model to ensure the integrity of the data before estimating the parameter of the model using robust methods and to ascertain whether the data follow a normal distribution or not, and to detect the problem of heterogeneity of error variance and to detect the presence of outliers in the data. By the least squares method, the model was as follows

$$\hat{y}_i = e^{0.0579 * x_i} \quad , i = 1, \dots \quad , 155 \qquad \dots (40)$$

We note that an increase in revenues by one unit affects an increase in expenditures by $e^{0.0579}$=(1.0596).

The Jarque-Bera test (h) was used to test the normal distribution of random error for the one-parameter exponential regression model, and it was found that the random error does not follow the normal distribution, where the test value was (h =1), which means rejecting the null hypothesis with a significance level (0.05). This indicates that the data either suffer from the problem of heterogeneity of variance or the presence of outliers in the data that were detected by the boxplot , so we resort to the robust methods for estimating the parameters of the one-parameter exponential regression model.

Gold field Quandt test was used to detect the problem of heterogeneity of error variance of the one-parameter exponential regression model. It turns out that the error suffers from the problem of heterogeneity of variance, as in Table (2):

Table (2): shows the heterogeneity test for the one-parameter exponential regression model

test	F	$F_{(78,78,0.05)}$
Gold field-Quant	3846.1	1.47

Arithmetic methods were used to detect and check the data the presence of outliers in the one-parameter exponential regression model, where Cook's method was used, as shown in Table (3).

Table (3): shows the outliers and influential observations in the one-parameter exponential model

observation	47	48	60	120	132
cook	✔	✔	✔	✔	✔

It is clear from Table (3) that there are five values affect the estimation of parameters of the one-parameter exponential according to the Cook test, so we use the robust methods.

The parameters of the one-parameter exponential regression model are estimated using the best methods that were produced by the experimental side of the research through the mean squares error squares for the model, which was the method of forward search in the case of the distribution of the error a normal distribution with different the ratios of outliers, because it gave less (MSE) for the model than other methods when (n=155). As shown in Table (4)

Table (4): shows the value of the parameters and mean of the error squares for the model the one-parameter exponential regression of the expenditures and revenues data

mothed	B	MSE	Comparison
OLS	0.0579	133.37	
MOM	0.0617	318.94	
FS	0.0466	44.539	best estimator
M	0.0208	46.58	

The robust forward research method is characterized by its ability of strength and robustness and achieved the highest resistance to outliers than the rest of the methods and is characterized by the accuracy of the estimate because it gave reliable results. We note that an increase in revenues by one unit affects an increase in expenditures by 1.0477= $e^{0.0466}$, as shown in the equation below:

$$\hat{y}_i = e^{0.0466 * x_i} \quad \ldots (41)$$

4. Conclusion

i. When the size is (25), (50), (100) and (150), the simulation results showed that the estimator using the forward search method proved its efficiency in the estimation method in the one-parameter exponential regression model, where it achieved the least mean squares of error compared to the rest of the estimation methods with an increase in the percentage of outliers.

ii. The effect of the sample size was clear on the accuracy of the results for one-parameter exponential regression model and all estimation methods, as the simulation results showed that the value of the mean squares error decreases gradually with increasing sample size and for different outliers ratios.

iii. Through the applied side, it was found that the one-parameter exponential regression model is achieved the least criterion for mean square error in the forward search method, and it appeared to us through the analysis of expenditure and revenue data, and using the exponential regression model, we notice that the data is appropriate for the model. In this study, the effect of the revenue variable on the expenditure variable was shown.

References

1. Alsalem, A. K. and Altaher, A.M., (2019), "A comparative study of some robust nonlinear regression methods", Journal of Pure & Applied Sciences, ISSN 2521-9200.

2. Atkinson, A.C. , Corbellini, A. and Riani, M., (2017), "Robust Bayesian regression with the forward search", Funding for this research was provided by: London School of Economics and Political Science, UK

3. Atkinson, A.C., Riani, M. and Cerioli,A., (2010), 'The forward search: Theory and data analysis', Journal of the Korean Statistical Society 39, 117–134 .

4. Ismail, M.I. and Rasheed, H. A., (2021), "Robust Regression Methods/ a Comparison Study", Turkish Journal of Computer and Mathematics Education Vol.12 No.14 (2021), 2939- 2949 .

5. Khan, D. M., Ali, M., Ahmad, Z. and Manzoor, S., (2021), "A New Efficient Redescending M-Estimator for Robust Fitting of Linear Regression Models in the Presence of Outliers", Hindawi, Mathematical Problems in Engineering Volume 2021, Article ID 3090537, 11 pages.

6. Koenker, R. and Park, B. J., (1994), "An interior point algorithm for nonlinear quantile regression", Journal of Econometrics 71 (1996) 265-283.

7. Lerasle, M., (2019), "Selected topics on robust statistical learning theory Lecture Notes", Paris-Saclay university.

8. Lim, C., Sen, P. K. and Peddada, S. D., (2013), "Robust nonlinear regression in applications", Journal Indian Soc Agric Stat. 2013; 67(2): 215–234.

9. Liu, P., Zhang, M., Zhang, R. and Zhou, Q., (2021), "Robust Estimation and Tests for Parameters of Some Nonlinear Regression Models", Journal Mathematics Volume 9.

10.Marasović, M., Marasović, T. and Milos, M., (2016), "Robust Nonlinear Regression in Enzyme Kinetic Parameters Estimation", Hindawi Journal of Chemistry Volume 2017, Article ID 6560983, 12 pages.

11.Motulsky, H. J. and Brown, R. E., (2006), "Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate", Journal BMC Bioinformatics 7, 123 (2006).

12.RIAN, M. and ATKINSO, A. C., (2000), 'Robust Diagnostic Data Analysis: Transformations in Regression' Technometrics, Vol. 42, No. 4 (Nov., 2000), pp. 384-394, Published by: American Statistical Association and American Society for Quality.

13.Riazoshams, H., Midi, H. and Ghilagaber, G., (2019), "Robust Nonlinear Regression: with Applications using R", John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA.

14.Tabatabai, MA., Kengwoung, K., Eby, W., Bae, S., Manne, U., Fouad, M. and Singh, K., (2014), "A New Robust Method for Nonlinear Regression", Journal of Biometric s & B iostatistics.vol.5.

15.Zhang, Y. and Liu, P., (2020), "Median-of-means approach for repeated measures data', Communications in Statistics - Theory and Methods, ISSN: 0361-0926 (Print) 1532-415X.

Journal of Economics and Administrative Sciences Vol.28 (NO. 134) 2022, pp. 147-159

التقديرات الحصينة لنموذج الانحدار الأسي أحادي المعلمة

الباحث/ احمد جودة ارشيد
قسم الاحصاء، كلية الادارة والاقتصاد، جامعة بغداد
Ahmed.jawda1201a@coadec.uobaghdad.edu.iq

ا.م د رباب عبد الرضا صالح
قسم الاحصاء، كلية الادارة والاقتصاد، جامعة بغداد
rabab.saleh@coadec.uobaghdad.edu.iq

Received:4/12/2022 Accepted: 26/12/2022 Published: December / 2022

مستخلص البحث

يعد الانحدار الأسي أحادي المعلمة أحد النماذج الأكثر شيوعًا والمستعملة على نطاق واسع في العديد من المجالات ، لتقدير معلمات أنموذج الانحدار الأسي أحادي المعلمة باستخدام طريقة المربعات الصغرى العادية ولكن هذه الطريقة ليست فعالة في وجود القيم الشاذة ، لذلك تم استعمال طرق الحصينة لمعالجة القيم الشاذة في أنموذج الانحدار الأسي أحادي المعلمة لتقدير المعلمات باستعمال طرائق الحصينة منها (وسيط المتوسطات MOM، البحث الأمامي FS، تقدير M) ، وتم استعمال المحاكاة للمقارنة بين طرق التقدير مع أحجام عينات مختلفة وافتراض أربع نسب من القيم الشاذة للبيانات (10٪ ، 20٪ ، 30٪ ، 40٪). ومن خلال متوسط الخطأ التربيعي (MSE) تم الوصول إلى أفضل طريقة تقدير للمعلمات ، حيث أظهرت النتائج التي تم الحصول عليها باستخدام المحاكاة أن البحث الأمامي FS هو الأفضل لأنه يعطي اقل متوسط لمربعات الخطأ . وفي الجانب التطبيقي ، تم استعمال بيانات النفقات والإيرادات لتقدير معلمات الانحدار الأسي ذي المعلمة الواحدة ، وحيث تم اختبار البيانات ، ظهر أن لها توزيعًا أسيًا ، وتم استخدام اختبار boxplot و (COOK) لاكتشاف القيم المتطرفة الموجودة في البيانات الحقيقية ، تم استعمال اختبار جودة المطابقة ومشكلة عدم تجانس تباين الأنموذج الأسي أحادي المعلمة ووجد أن البيانات لا تتبع التوزيع الطبيعي وأنها تعاني من مشكلة عدم تجانس التباين. تم تقدير أنموذج الانحدار الأسي ذي المعلمة الواحدة لبيانات النفقات والإيرادات باستخدام طريقة البحث الامامي لأنها كانت أفضل في التقدير.

نوع البحث: ورقة بحثية.

المصطلحات الرئيسة للبحث أنموذج الانحدار الاسي ، التقديرات الحصينة ، مقدر وسيط المتوسطات ، مقدر البحث الامام ، مقدر M الحصين.

*البحث مستل من رسالة ماجستير