Journal of Economics and Administrative Sciences
2024; 30(143), pp. 523-535

P-ISSN 2518-5764
E-ISSN 2227-703X

Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq
DOI:https://doi.org/10.33095/p4nwgd31

Comparison between estimates of beta regression model using MSE

Melad Dheyaa Ismaiel*
Department of Statistics,
College of Administration and Economics
University of Baghdad
Baghdad, Iraq
mielad.diaa2101m@coadec.uobaghdad.edu.iq
https://orcid.org/0009-0005-0055-5414/print
***Corresponding author**

Ahmed Dheyab Ahmed
Department of Statistics,
College of Administration and Economics
University of Baghdad
Baghdad, Iraq
ahmedstatistic@ coadec.uobaghdad.edu.iq
https://orcid.org/0000-0002-1689-1376

Received:20/11/2023 Accepted:7/5/2024 Published Online First: 1 /10/ 2024

Abstract:

The multiple linear regression analysis shows the effects of explanatory variables on the dependent variable. Its function is to represent data to understand the shape and nature of the relationship between explanatory variables and the response variable. One of the main problems faced by linear regression is the nature of the data, as sometimes it does not follow a normal distribution. This has led to the emergence of different types of linear regression models such as the beta regression model, where the data of the dependent variable is less than one and confined within the period (0,1), and the data is relative and decimal. Non-traditional methods will be used to estimate beta regression parameters because some traditional methods do not provide suitable and accurate results when dealing with one of the main problems encountered in regression, which is multicollinearity. To address this issue, we used some methods include The Maximum Likelihood Estimation (MLE), ridge regression, modified Liu, and Ozkale and Kaciranlar method. Shrinking parameters are proposed for each method, and a comparison between the methods is conducted using the Mean Squared Error comparison criterion through a simulation experiment with two approaches, the simulation results showed that the ridge regression method is the best for the first and second methods of simulation.

Paper type Research Paper
Keywords: Beta Regression Model, Multicollinearity, Maximum Likelihood Estimation, Ridge Regression, Modified Liu, Ozkale and Kaciranlar Method, Simulation.

1.Introduction:

Linear regression is considered one of the topics used in statistics, in addition to its importance in the medical, economic, social and many fields of life (Rao et al, 2008; Thompson, 2003; Mohamed and Ahmed, 2022). Among the problems that linear regression suffers from is that the data do not follow a normal distribution and the emergence of different regression models such as the Poisson regression model, gamma regression model and exponential regression model (Irshayyid and Saleh, 2022; Radam and Hameed, 2023). James and Stein (1961) studied regression models in which random errors are distributed in spherical distributions when there is a problem of multicollinearity. They found a modern method to estimate model parameters and to address the problem of multicollinearity. They also explained the characteristics of a good estimator. Extracted from models suffering from multicollinearity. Hoerl and Kennard (1970) presented a proposal for a method that addresses the problem of multicollinearity in linear regression models by adding a small positive quantity (k) to the diagonal elements of the (x'x) matrix, and they called it the Ridge Regression Estimator Method (Abonazel and Taha,2023; Kadhim and Suslim, 2002; Mohammad and Ahmed, 2022). They showed theoretically that the estimator resulting from this addition is a biased estimator and is more efficient than the estimator produced by the least squares method. Therefore, the beta regression model in which the random errors follow a beta distribution will be studied, the beta distribution is considered one of the most important continuous probability distributions whose data are within the period (0, 1), and it is worth noting that the beta distribution is derived from the beta function (Hirmez, 1990; Yang and Chang, 2010). The beta distribution is one of the important and basic statistical distributions in its ability to adapt and model continuous random variables (Abonazel et al,2022; Abonazel et al,2023)., because its values are decimal numbers confined between the period(1, 0) (Hirmez, 1990), one of the most important problems that regression models suffer from is the problem of multicollinearity between the explanatory variables when a correlation is formed between the variables(Swindel, 1976; Sadullah and Selahattin, 2008), such as the connection of two or more variables (Abdulah et al, 2019; Ahmed et al, 2020), to reduce its effect on linear regression models, many methods have been developed to estimate model parameters and detect the problem(Ahmed et al, 2020; Kareem and Hashim, 2021).

1.1 Literature review :

Ferrari and Cribari-Neto(2004) presented a new study that includes a beta regression model when random errors follow a beta distribution and the distribution parameters are composed of the mean and Precision parameters. They also demonstrated the importance and usefulness of the beta regression model when The data is continuous and within the period (0, 1).

Ospina et al(2006) calculated the second-order deviations of the maximum likelihood estimators of the beta regression model and used them by defining (bias-adjusted) estimators as an alternative to the analytically corrected maximum likelihood estimators, using the (Parametric bootstrap) method, in addition to suggesting different strategies for estimating intervals.

Simas et al (2010) expanded the formula of the beta regression model and studied it when the regression and the Precision parameter is non-linear, they compared three estimators free of deviations, the simulation results indicate a better result by estimating the model directly in its non-linear form, in addition to choosing the appropriate dispersion variances for the regression model.

Liu (2013) proposed a new method to reduce and address the problem of multicollinearity in multiple linear regression and they relied on the shrinkage parameters of the ridge regression method and the Liu method, to estimate the parameters of the multiple linear regression model and call it the Modified Liu Method.

Algamal and Abonazel (2021) presented the Modified Liu Method for the beta regression model to address the problem of multicollinearity. The estimator of the Modified Liu method was compared with the maximum likelihood method, Liu method and the ridge regression method, which depends on the mean squares error.
Qasim et al (2021) presented a new Ridge Regression estimator method for the beta regression model in order to reduce the instability of the maximum likelihood method for the beta regression model. They also proposed new formulas for calculating the shrinkage parameter for the Ridge Regression method, as well as proposing Median Squared Error(SE) This criterion can provide compelling evidence supporting the proposed method in the Monte Carlo simulation study.

Abonazel et al (2022) developed a new biased estimator and it was called (Two-Parameter Estimator) to estimate the parameters of the beta regression model. This estimator depends on several methods, namely the Ridge Regression method, the Liu method, the Modified Liu method, and the Maximum Likelihood Estimators method, because the new estimator depends on the parameters of these methods.

Geissinger et al (2022) presented a new case of beta regression in the natural sciences, meaning it is continuous data. They used two types of data related to the natural sciences, which is biochemistry. And environmental components, using the Maximum Likelihood method.

The multiple linear regression model has many types and hypotheses. Among those types, when the random error is distributed in a normal distribution, the model parameters can be estimated by using the ordinary least squares method and the maximum likelihood method, as its capabilities are accurate and good.

However, when the random error is distributed with other probability distributions such as the beta distribution and in the presence the problem of multicollinearity among its explanatory variables will be resorted to the Ridge Regression method (Yousif and Hussein, 2015), the modified Liu method, and the Ozkale and Kaciranlar method to estimate the model parameters and to obtain accurate and good results.

The objective of this research is to make a comparison between estimation methods using the MSE comparison standard, by estimating the Beta Regression model when there is a problem of multicollinearity using the Ridge Regression method, the modified Liu method, and the Ozkale and Kaciranlar method, using two simulation methods and using the proposed shrinkage parameters.

2.Material and Methods:

2.1 Beta Regression Model:

The probability density function of the beta distribution (Ospina et al, 2006; Simas et al, 2010) can be written after performing transformations in the equation (1):

$$f(y;\mu,\varphi)=\frac{\Gamma(\varphi)}{\Gamma(\mu\varphi)\Gamma((1-\mu)\varphi)}y^{\mu\varphi-1}(1-y)^{(1-\mu)\varphi-1}\ ;0<y<1 \qquad (1)$$

where : $0<\mu<1$, $\varphi > 0$

φ : Precision Parameter (Abonazel et al, 2022; Yasin, 2023).

μ : Mean of beta probability distribution.

2.2 Maximum Likelihood Method:

To estimate the parameters of the beta regression model using the maximum likelihood estimator (MLE) method (Cribari-Net and Vasconcellos, 2002; Smithson and Verkuilen, 2006), and by maximizing the observations of the beta regression equation and taking the natural logarithm, the equation is as follows:

$$L = \sum_{t=1}^{n} \{log\Gamma(\varphi) - log\Gamma(\mu_t(\varphi)) - log\Gamma((1-\mu_t)(\varphi)) + (\mu_t(\varphi) - 1)\log(y_t) + ((1-\mu_t)(\varphi) - 1)\log(1-y_t)\} \quad (2)$$

The derivation of the maximum likelihood method equation goes through two stages:
The first stage is according to the equation (3):

$$\beta^{(r+1)}{}_{ML1} = \beta^{(r)} + \left(I_{\beta\beta}^{(r)}\right)^{-1} S_{\beta}^{(r)}(\beta) \quad ; r = 1,2,\dots,R \quad (3)$$

$$\beta^{(1)} = \hat{\beta}_{ols} = (X'X)^{-1}X'Y \quad (4)$$

Where:
$\hat{\beta}_{ols}$: Ordinary Least Square Method (OLS).
R : Number of iterations.

Estimation equation (3) is considered the first stage in estimating the parameters of the beta regression model according to the maximum likelihood method.

Also, $I_{\beta\beta}^{(r)}$ represents Fisher's equation

$$I_{\beta\beta}^{(r)} = -E\left(\frac{\partial^2 \ell(\beta,\varphi)}{\partial\beta\partial\beta'}\right) = \varphi X'W\,X$$

$$S(\beta) = \varphi X'A(y^* - \mu^*) \quad (5)$$

The second stage of estimating the parameters of the beta regression model is according to the equation as following:

$$\beta^{r}{}_{ML2} = (x'\widehat{w}x)^{-1}x'\,\widehat{w}\,\hat{z} \quad (6)$$

$$\hat{z} = \hat{\eta} + \widehat{w}^{-1}\hat{A}(y^* - \mu^*) \quad (7)$$

$$\widehat{w}_t = \varphi\{\psi'(\mu_t\varphi) + \psi'((1-\mu_t)\varphi)\}\frac{1}{[g'(\mu_t)]^2}$$

X : is an (n$\times$ k) matrix of regressors.

A=diag$\left\{\frac{1}{g`(\mu_1)},\dots,\frac{1}{g`(\mu_t)}\right\}$

$y^* = Log\frac{y_t}{1-y_t}$; $\left\{Log\frac{y_1}{1-y_1},\dots,Log\frac{y_n}{1-y_n}\right\}$

(8)

$$\mu^* = \{\psi(\mu_t\varphi) - \psi((1-\mu_t)\varphi)\} = \{\{\psi(\mu_1\varphi) - \psi((1-\mu_1)\varphi)\},\dots,\{\psi(\mu_n\varphi) - \psi((1-\mu_n)\varphi)\}\} \quad (9)$$

η_t : Linear Predictor (Asar, 2018).

A=diag$\left\{\frac{1}{g`(\mu_1)},\dots,\frac{1}{g`(\mu_t)}\right\}$

$y^* = Log\frac{y_t}{1-y_t}$; $\left\{Log\frac{y_1}{1-y_1},\dots,Log\frac{y_n}{1-y_n}\right\}$

(8)

$$\mu^* = \{\psi(\mu_t\varphi) - \psi((1-\mu_t)\varphi)\} = \{\{\psi(\mu_1\varphi) - \psi((1-\mu_1)\varphi)\},\dots,\{\psi(\mu_n\varphi) - \psi((1-\mu_n)\varphi)\}\} \quad (9)$$

η_t : Linear Predictor (Asar, 2018).

2.3 Ridge Regression (RR):

There are two types of multicollinearity, which are complete and semi-complete (Algamal et al,2022). When the multicollinearity is semi-complete (semi multicollinearity) and the determinant value of the information matrix is very small, this leads to the variance of the estimated information being large in very large proportions(Yalian and Yang, 2012; Hussein, 2019), we conclude from this that some explanatory variables are sometimes insignificant in the linear regression model, and one of the most important methods for dealing with this problem is

the ridge regression method (Kadhim and suslim, 2002; Khalaf and Shukur, 2005). The ridge regression method will be used for estimation the beta regression model in presence the problem of multicollinearity (Muniz and Kibria, 2009; Hussein,2016), in addition to proposing the shrinkage parameter (k) and the mean square error criterion based on Monte Carlo simulation.
Let $\hat{\beta}$ be an estimate of the vector β and the Weighted Sum of Squared Error (WSSE) is defined as follows (Kibria, 2003; Qasim et al, 2021):

$$\phi = \left(y - \hat{\beta}_{RR}\right)'\left(y - \hat{\beta}_{RR}\right)$$
$$= \left(y - X\hat{\beta}_{ML}\right)'\left(y - X\hat{\beta}_{ML}\right) + \left(\hat{\beta}_{RR} - \hat{\beta}_{ML}\right)' X'WX\left(\hat{\beta}_{RR} - \hat{\beta}_{ML}\right)$$

Where:
ϕ : Minimum value.
$\Theta(\hat{\beta}) > 0$ and B_{RR} It is obtained by minimizing the range of $\widehat{\beta_{RR}}'\widehat{\beta_{RR}}$ while setting the constraint(Kibria and Banik, 2016; Qasim et al, 2021):

$$Q = \hat{\beta}'_{RR}\hat{\beta}_{RR} + \left(\frac{1}{k}\right)\left\{\left(\hat{\beta}_{RR} - \hat{\beta}_{ML}\right)' X'WX\left(\hat{\beta}_{RR} - \hat{\beta}_{ML}\right) - \phi_0\right\} \quad (10)$$

Where :
$\left(\frac{1}{k}\right)$: It represents the Lagrange Multiplier
After simplifying, we obtain the $\hat{\beta}_{BRR}$ estimator (Awwad et al,2022)

$$\hat{\beta}_{BRR} = (X'WX + k_{PR}I)^{-1}X'WX\hat{\beta}_{ML} \quad (11)$$

k>0
K : It is the shrinkage parameter (Khalaf and Shukur, 2005).
I: Identity matrix (P × P).
The proposed shrinkage k is:

$$k_{PR} = \frac{1}{\varphi\, Min(\alpha_j)^2} \quad (12)$$

2.4 Modified Liu Method (MLR):

Algamal and Abonazel (2021) developed another method to reduce the problem of multicollinearity by developing the Liu method to the Modified Liu method by estimating the two shrinkage parameters and integrating the shrinkage parameter for Liu method and the shrinkage parameter for the Ridge Regression method (Liu, 1993; Asar and Genc, 2016; Algamal and Abonazel, 2021), the Modified Liu Regression (MLR) method is as follows (Liu, 2003; Omara, 2019):

$$\hat{\beta}_{MLR} = \left(X'\widehat{W}X + k_{PR}I\right)^{-1}\left(X'\widehat{W}X - d_{PL}I\right)\hat{\beta}_{ML} \quad (13)$$

$k > 0\,, -\infty < d < \infty$
The bias vector for modified liu method for beta regression is (Liu, 2004; Abonazel et al, 2022):

$$Bias\left(\hat{\beta}_{MLR}\right) = -(d_{PL} + k_{PR})\left(X'\widehat{W}X + k_{PR}I\right)^{-1}\beta_{ML} \quad (14)$$

The proposed shrinkage parameters for the Modified Liu method for beta regression are according to the following equations:

$$k_{PR} = \frac{1}{\varphi\, Min(\alpha_j)^2} \quad (15)$$

$$d_{PL} = Median\left(\sqrt{\frac{\alpha^2{}_j}{\lambda_j}}\right) \quad (16)$$

2.5 Ozkale and Kaciranlar Method (O.K.):

This is the method proposed by (Ozkale and Kaciranlar, 2007) for the beta regression model to address the problem of multicollinearity and to obtain more stable and accurate results, and the estimation equation is as follows (Kurnaz and Akay, 2015; Lukman et al, 2019; Omara, 2019; Abonazel et al, 2022):

$$\hat{\beta}_{TPBR} = \left(X'\widehat{W}X + k_R I\right)^{-1}\left(X'\widehat{W}X + k_R d_M I\right)\hat{\beta}_{ML} \tag{17}$$

Where:

$k > 0$; $0 < d < 1$

The bias vector for two beta regression parameters is as follows (Abonazel et al, 2022):

$$Bias\left(\hat{\beta}_{TPBR}\right) = k_R(d_M - 1)\left(X'\widehat{W}X + k_R I\right)^{-1}\beta_{ML} \tag{18}$$

The shrinkage functions are as follows:

The shrinkage parameter (k) is according to the equation (19):

$$k_R = \frac{1}{\varphi \sum_{j=1}^{p} \hat{\alpha}_j^2} \tag{19}$$

The shrinkage parameter (d) is according to the equation (20):

$$d_M = \frac{\sum_{j=1}^{p}\left[\frac{\left(\frac{1}{\varphi} - k_M \hat{\alpha}_j^2\right)}{\left(\lambda_j + k_M\right)^2}\right]}{\sum_{j=1}^{p}\left[\frac{\left(\frac{1}{\varphi} + \lambda_j \hat{\alpha}_j^2\right)}{\lambda_j\left(\lambda_j + k_M\right)^2}\right]} \tag{20}$$

3.Discussion of Results:

3.1 Simulation:

The comparison process between the estimation methods was carried out based on the mean square error in generating random data on the MATLAB program.

1. the random error variable (Y_i) was generated in a beta regression model according to a beta distribution with two parameters (μ_i) and (φ).

Where:

$y_i \sim Beta\ (\mu_i, \varphi)$

$\varphi = \{1, 2\}$

As for (μ_t)its value is according to the equation (21):

$$\mu_t = \frac{e^{x_t'\beta}}{1 + e^{x_t'\beta}} \tag{21}$$

2. The explanatory variables $x_i = (x_{i1}, \dots, x_{ip})'$are subjected to correlational relationships to form the problem of multicollinearity between the explanatory variables of the model using the following equation(Arashi et al, 2021):

$$x_{ij} = (1 - \rho^2)^{0.5} Z_{ij} + \rho Z_{ij} \tag{22}$$

i= 1 , 2 , … , n , j= 1 , 2 , ... , p

ρ : It represents the simple correlation coefficient between each pair of explanatory variables.

Z_{ij} : They represent random numbers that are independent and generated in two different ways, as follows:

1. The first method of simulation :Z_{ij} is generated when it has a pseudo-standard normal distribution

2.The second method of simulation: it is done by generating Z_{ij} values according to real data, as follows:

$$Z_{ij} = L(X) + (U(X) - L(X))rand\ (n, P) \tag{23}$$

Since:

L(X): represents the lowest value of the explanatory variable in the real applied experiment data.

U(X): represents the upper value of the explanatory variable in the real applied experiment data.

n : represents the sample size.

P-ISSN 2518-5764
E-ISSN 2227-703X

P: represents the number of explanatory variables.

3. Six sizes were chosen for the selected samples (20, 30, 50, 60, 100, 150) and three values for the correlation coefficients (0.90, 0.95, 0.99). The number of explanatory variables in the experiment was assumed to be (2, 5) and the experiment was repeated 1000 times.

4. The mean square error estimates for the model are calculated according to the following equation:

$$MSE(\hat{Y}) = \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n-k} / L \qquad (24)$$

3.2 Results of Simulation :

The simulation results for the first method of simulation include the following results, which are shown below in Table 1:

Table 1: Represents a comparison between the estimation methods in the presence of the proposed shrinkage parameters and with respect to the number of variables (2, 5) by adopting the mean square error criterion for the first method of simulation.

P	n	ρ	ϕ	RR	MLR	O.K.
2	20	0.90	1	0.242853161169977	1.73984764404361	1.47433295517998
			2	0.0794658304821197	0.710342813069322	0.519179254523815
		0.95	1	0.236351878744249	1.92021106316904	1.59966488919547
			2	0.0744184927728133	0.747731921356797	0.530102112052892
		0.99	1	0.223681915780450	2.35970654028362	1.89185510390365
			2	0.0653917986791959	0.839121132617783	0.556622444075752
	30	0.90	1	0.236740611004212	1.20775086704980	1.07784270557839
			2	0.0815625052548073	0.508168210071924	0.410661322543856
		0.95	1	0.231305984208223	1.31000026851344	1.15480469020285
			2	0.00922671547699422	0.136909707663703	0.102024136224316
		0.99	1	0.220788160851393	1.56479459568447	1.34175369017031
			2	0.0685251495267468	0.582622160158033	0.439191410869473
	50	0.90	1	0.242288935073918	0.910472046823098	0.851643549555290
			2	0.0872390887535289	0.376166643953343	0.331477754163610
		0.95	1	0.237696612513651	0.977355716939770	0.907562126944415
			2	0.0109638885470376	0.0922055443616729	0.0775552318512533
		0.99	1	0.228945692580354	1.14032140912100	1.04157277101508
			2	0.0750903478062654	0.420654774393536	0.355194667963601
	60	0.90	1	0.239058864943827	0.822251258602106	0.777266287438765
			2	0.0872327889184683	0.339295119939626	0.304932772395718
		0.95	1	0.234740236769663	0.879598347883308	0.826394583452940
			2	0.0830244895817542	0.349564571187446	0.310674930471660
		0.99	1	0.226565503965725	1.01790869740000	0.943139084814303
			2	0.0755061536985577	0.375428714133321	0.325415962496382
	100	0.90	1	0.243705843745015	0.666307907690692	0.644265630493102
			2	0.0905544119307057	0.270174586608284	0.253314975232123
		0.95	1	0.239869752403356	0.705257836203813	0.679430799189001
			2	0.0865793654515260	0.275968935583050	0.256999349417598
		0.99	1	0.232740096498724	0.799577334421131	0.763915563050519
			2	0.0141161267129018	0.0827748785355776	0.0743965642973238
	150	0.90	1	0.242920816645181	0.564816713587232	0.552192688079442
			2	0.0916056230084848	0.228878325520578	0.219148577557707
		0.95	1	0.239356549288510	0.592589578058194	0.577917448728386
			2	0.0877738113879950	0.232017075283025	0.221135645679100
		0.99	1	0.232818839044347	0.659528329064194	0.639606021066771

P-ISSN 2518-5764
E-ISSN 2227-703X

			2	0.0810177351155632	0.241443195447440	0.227752651832593
5	20	0.90	1	0.224262646318021	2.39324739293209	1.46162789431889
			2	0.0698904596873413	1.06938118586365	0.473271153448016
		0.95	1	0.216209944965279	2.65536007168053	1.54455323257877
			2	0.0649776317726388	1.12963323535035	0.466764619556106
		0.99	1	0.201106547970498	3.29982702121196	1.72651326852207
			2	0.0566604908457195	1.27184231186771	0.455021899590286
	30	0.90	1	0.223732132375628	1.67996272333220	1.24595933725202
			2	0.0717823376721720	0.735518239983831	0.434479263266498
		0.95	1	0.216547729153008	1.85241837304808	1.32809855004189
			2	0.0668073761265380	0.774453957437023	0.434917184797799
		0.99	1	0.202764067127237	2.26676499820914	1.51288848041864
			2	0.0581417754961798	0.867446002263011	0.436672151935356
	50	0.90	1	0.219900636255829	1.21016718391146	1.01721004087943
			2	0.0750895795204167	0.523702140596940	0.383493271651176
		0.95	1	0.214035411041681	1.31829267464633	1.08713189964915
			2	0.0703781078601558	0.546927688974045	0.387938039969871
		0.99	1	0.202703705526162	1.58128042314348	1.24920502089825
			2	0.0620052905954948	0.603622731361240	0.399035165572294
	60	0.90	1	0.219688024481672	1.09383903872220	0.947252485800825
			2	0.0767193341825638	0.470103496075259	0.362452365169379
		0.95	1	0.214220146828383	1.18754116816863	1.01214765481335
			2	0.0721197346404005	0.489548931155511	0.367396060724617
		0.99	1	0.203673203760534	1.41513274599361	1.16357783267493
			2	0.0639169198171012	0.537174811468390	0.379697097343320
	100	0.90	1	0.221822353912346	0.825711627839514	0.759284684786979
			2	0.0810872506876259	0.352140477437405	0.301710417450121
		0.95	1	0.217265939946426	0.886203630908605	0.807306157592234
			2	0.0768166276148607	0.363478491986448	0.306356944302587
		0.99	1	0.208585679104381	1.03276182780285	0.921026020924472
			2	0.0691902121797065	0.392026157552303	0.318496952740095
	150	0.90	1	0.223003837070396	0.679743483889331	0.643200516589705
			2	0.0833049546002092	0.286554156074498	0.258504386964927
		0.95	1	0.218962124672076	0.723001732998766	0.679908052915333
			2	0.0792781451042453	0.293628325861748	0.261976508331279
		0.99	1	0.211365980442014	0.827617927107038	0.767402946162064
			2	0.0721155341243539	0.312156421825117	0.271674775362802

As for the simulation results for the second simulation method, the following results are included in Table 2:

P-ISSN 2518-5764
E-ISSN 2227-703X

Table 2 : Represents a comparison between the two estimation methods in the presence of the proposed shrinkage parameters and with respect to five explanatory variables, adopting the mean squared error criterion for the second method of simulation.

P	*n*	ρ	ϕ	RR	MLR	O.K.
5	100	0.90	1	0.00386699039904131	0.00386690803511351	0.00386704158299637
			2	0.00333038413818084	0.00333042724363634	0.00333040360031634
		0.95	1	0.00378747227886261	0.00378756811272587	0.00378752088672872
			2	0.00324208068553154	0.00324213251340113	0.00324210207719897
		0.99	1	0.00364908274521512	0.00364914156203378	0.00364912606087515
			2	0.00309346710701350	0.00309352396647516	0.00309349691050538

The results of the simulation experiment showed that the ridge regression method for the beta regression model is better than the rest of the estimation methods by comparing the results of the mean square errors for all the estimation methods used, where the ridge regression method was the smallest among all the results of the estimation methods in the presence of the proposed shrinkage parameters and when there are explanatory variables equal to (2, 5) and at two levels of Precision Parameter (1, 2) and for the first and second simulation methods..

4 . Conclusion:

1. The simulation results showed that the Ridge Regression method is better than the Modified Liu method and Ozkale and Kaciranlar method in the presence of the proposed shrinkage parameters, for the first and second simulation methods.
2. When using the second method of simulation, the Ridge Regression method when there are five explanatory variables is the best in the presence of the proposed shrinkage parameters.
3. The results of the mean square errors showed that the second best method after the Ridge Regression method is the Ozkale and Kaciranlar method when there are two explanatory variables and also in the presence of 5 explanatory variables as for the first simulation method.
4. Based on the simulation results, it is recommended to use the ridge regression method in the presence of the proposed shrinkage parameter to estimate the parameters of the beta regression model as an alternative to the modified Liu method and the Ozkale and Kaciranlar method.
5. Other shrinkage parameters suitable for estimation methods can be suggested and compared.

Authors Declaration:

Conflicts of Interest: None

-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, Which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

- Ethical Clearance: The Research Was Approved By The Local Ethical Committee in The University.

References:

1.Abdulah, E., Ahmed, A. and Aboulwahhab, B. (2019). “Comparison of Estimate Methods of Multiple Linear Regression Model Auto –correlated Error when the Error Distribution with General Logistic” Journal of Engineering and Applied Sciences, Vol.14, No.19, pp.7072-7076.
2.Abonazel, M.R., Algamal, Z., Awwad, F. and Taha, I.M. (2022). “A New Two-Parameter Estimator for Beta Regression Model: Method, Simulation, and Application” Frontiers in Applied Mathematics and Statistics, Vol.7, pp.780322.
3.Abonazel, M.R. and Taha, I.M. (2023) “Beta ridge regression estimators: simulation and application” Communications in Statistics-Simulation and Computation, Vol.52, No.9, pp.4280-4292.

4. Abonazel, M.R., Dawoud, I., Awwad, F.A. and Lukman, A.F. (2022). "Dawoud–Kibria estimator for beta regression model: simulation and application" Frontiers in Applied Mathematics and Statistics, Vol.8, p.775068.
5. Abonazel, M.R., Said, H.A., Tag-Eldin, E., Abdel-Rahman, S. and Khattab, I.G. (2023) "Using beta regression modeling in medical sciences: a comparative study" Commun. Math. Biol. Neurosci., pp.Article-ID.
6.Ahmed, A., Abdulah, E., Abdulwahhab, B. and Abed, N. (2020). "Solving Multicollinearity Problem of gross domestic product using ridge regression method" Periodicals of Engineering and Natural Sciences, Vol.8, No.2, pp.668-672.
7.Ahmed, A., Abdulah, E. and Abdulwahhab, B. (2020). "Estimating the Coefficient of Polynomial Regression Model when the Error Distributed Generalized Logistic with three Parameters" International Journal of Psychosocial Rehabilitation, Vol.24, No.8, pp.4487-4496.
8.Algamal, Z.Y. and Abonazel, M.R. (2021). "Developing a Liu-type estimator in beta regression model" Concurrency and Computation: Practice and Experience, Vol.34, No.5, pp.1-11.
9. Algamal, Z.Y., Abonazel, M.R. and Lukman, A.F. (2022) "Modified Jackknife Ridge Estimator for Beta Regression Model with Application to Chemical Data" International Journal of Mathematics, Statistics, and Computer Science, Vol.1, pp.15-24.
10.Arashi, M., Roozbeh, M., Hamzah, N.A. and Gasparini, M. (2021). "Ridge regression and its applications in genetic studies" Plos One, Vol.16, No.4, pp. e0245376.
11.Asar, Y. and Genc, A. (2016). "New shrinkage parameters for the Liu-type logistic estimators" Communications in Statistics - Simulation and Computation, Vol.45, No.3, pp.1094-1103.
12.Asar, Y. (2018). "Liu-type negative binomial regression: a comparison of recent estimators and applications" Trends and Perspectives in linear statistical inference, springer, pp.23-39, doi.org/10.1007/978-3-319-73241-1_2
13. Awwad, F.A., Dawoud, I. and Abonazel, M.R. (2022) "Development of robust Özkale–Kaçiranlar and Yang–Chang estimators for regression models in the presence of multicollinearity and outliers" Concurrency and Computation: Practice and Experience, Vol.34, No.6, p. e6779.
14.Cribari-Neto, F. and Vasconcellos, K.L.P. (2002). "Nearly unbiased maximum-likelihood estimation for the beta distribution" Journal of Statistical Computation and Simulation, Vol.72, No.2, pp.107-118.
15.Ferrari, S.L.P. and Cribari-Neto, F. (2004). "Beta Regression for Modelling Rates and Proportions" Journal of Applied statistics, Vol.31, No.7, pp.799-815.
16. Geissinger, E.A., Khoo, C., Richmond, I.C., Faulkner, S.J. and Schneider, D.C. (2022). "A casefor beta regression in the natural sciences" Ecosphere, Vol.13, No.2, pp. e3940
17.Hirmez, A.H. (1990). "Mathematical Statistics" University Press, Iraq.
18.Hoerl, A.E. and Kennard, R.W. (1970). "Ridge regression: biased estimation for non-orthogonal problems" Technometrics, Vol.42, No.1, pp.80-86.
19. Hussein, S.M. (2016) "compared the proposed method (AUGJRR) with biased methods to estimate the Generalized ridge regression of the existence of multicollinearity" Al-Rafidain University College for Sciences, No.37.
20. Hussein, S.M. (2019) "Comparison of Some Suggested Estimators Based on Differencing Technique in the Partial Linear Model Using Simulation" Baghdad Science Journal, Vol.16, No.4, pp.0918-0918.
21. Irshayyid, A. and Saleh, R. (2022). "Robust Estimates for One – parameter Exponential Regression Model" Journal of Economics and Administrative Sciences (JEAS), Vol.28, No.134, pp.147-159.
22. Kadhim, A.H. and Suslim, B.S. (2002). "The Advance Econometrics Theory and Application" University of Baghdad, Iraq.

23.Kareem, U. and Hashim, F. (2021). “Comparison of some methods for Estimating maxture of linear Regression models with Application” Journal of Economics and Administrative Sciences (JEAS), vol.27, No. 129, pp.171-184.

24. Khalaf G. and Shukur G. (2005). “Choosing ridge parameter for regression problems” Communications in Statistics - Theory and Methods, Vol.34, No.5, pp.1177-1182.

25. Kibria, B.M.G. and Banik, S. (2016). “Some ridge regression estimators and their performances” Journal of Modern Applied Statistical Methods, Vol.15, No.1, pp.206-238.

26. Kibria, B.M.G. (2003). “Performance of some new ridge regression estimators” Communications in Statistics - Simulation and Computation, Vol.32, No.2, pp.419-435.

27.Kurnaz. F.S. and Akay, K.U. (2015). “A new Liu-type estimator” Statistical Papers, Vol.56, No.2, pp.495-517.

28.Liu, k. (2004). “More on Liu-type estimator in linear regression” Communications in Statistics - Theory and Methods, Vol.33, No.11, pp.2723-2733.

29.Liu, K.A. (1993). “New class of biased estimate in linear regression” Communications in Statistics - Theory and Methods, Vol.22, No.2, pp.393-402.

30. Liu, K. (2003). “Using Liu-Type Estimator to Combat Collinearity” Communication in Statistics –Theory and Methods, Vol.32, No.5, pp.1009-1020.

31. Lukman, A.F., Ayinde, K., Binuomote, S. and Clement, O.A. (2019). “Modified ridge-type estimator to combat multicollinearity: application to chemical data” Journal of Chemometrics, Vol.33, No.5, pp. e3125.

32. Mohamed, M. and Ahmed, A. (2022). “Comparison of Some Methods for Estimating a Polynomial Regression Model using Simulation” Mathematical Statistician and Engineering Applications, Vol.71, No.4, pp.693-706.

33. Mohammad, M. and Ahmed, A. (2022). “Estimating parameters via L-Linear method for Second-order Regression of polynomial Model” Journal of Economics and Administrative Sciences (JEAS), Vol.28, No.134, pp.160-167.

34. Muniz, G. and Kibria, B.M.G. (2009). “On some ridge regression estimators: an empirical comparison” Communications in Statistics - Simulation and Computation, Vol.38, No.3, pp.621-630.

35.Omara, T. (2019). “Modifying Two-Parameter Ridge Liu Estimator Based on Ridge Estimation” Pakistan Journal of Statistics and Operation Research, Vol.15, No.4, pp.881-890.

36.Ospina, R., Cribari-Neto, F., Vasconcellons, K. (2006). “Improved point and interval estimator for a beta regression model” Computational Statistics and Data Analysis, Vol.51, No.2, pp.960-981.

37.Ozkale, M.R. and Kaciranlar, S. (2007). “The restricted and unrestricted Two – Parameter estimators” Commun Stat Method, Vol.36, No.15, pp.2707-2725.

38.Qasim, M., Mansson, k. and Kibria, B.M. (2021). “On Same Beta Regression estimators: method, simulation and application” Journal of Statistical Computation and simulation, Vol.91, No.9, pp.1699-1712.

39. Radam, S. and Hameed, L. (2023). “comparison of poisson Regression and Conway Maxwell poisson models using simulation” Journal of Economics and Administrative Sciences (JEAS), Vol.29, No.136, pp.81-89.

40. Rao, C., Toutenburg, H., Shalabh and Heumann C. (2008). “Linear models and generalizations” Springer, Berlin.

41. Sabri, H.M. (2013). “A comparison of parameters estimation methods for the poisson regression model under existing of multicollinearty problem with application” PhD thesis, College Administration and Economics at the University of Baghdad, Iraq.

42. Sadullah, S. and Selahattin, K. (2008). “A new biased estimator based on ridge estimation” Statistical papers, Vol.49, No.4, pp.669-689.

43. Simas, A.B., Barreto-Souza, W. and Rocha, A.V. (2010). “Improved estimators for general class of beta regression models” Computational Statistics and Data Analysis, Vol.54, No.2, pp.348-366.
44. Smithson, M. and Verkuilen, J. (2006). “A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables” Psychol Methods, Vol.11, No.1, pp.54-71.
45. James, W. and Stein, C. (1961). “Estimation with Quadratic Loss” Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol.4, No.1, pp.361-379.
46. Swindel, B. (1976). “Good ridge estimators based on prior information” Communications in Statistics-Theory and Methods, Vol.5, No.11, pp.1065-1075.
47.Thompson, J.R. (2003). “Simulation: A Modeler’s Approach” IIE Transactions, Vol.35, No.3, pp.325-328.
48.Yalian, L. and Yang, H. (2012). “Anew Liu- type estimator in linear regression model” Statistical Papers, Vol.51, No.2, pp.1-11.
49. Yasin, A.S.A.R. (2023) “Performance of A New Bias Corrected Estimator in Beta Regression Model: A Monte Carlo Study” Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, Vol.5 No.2, pp.38-49.
50. Yang, H. and Chang, X. (2010). “A new two-parameter estimator in linear regression” Communications in Statistics-Theory and Methods, Vol.39, No.6, pp.923-934.
51. Yousif, H.M. and Hussein, S.M. (2015) “A comparison among some biased estimators in generalized linear regression model in present of multicollinearity” AL-Qadisiyah Journal for Administrative and Economic sciences, Vol.17, No.2.

Journal of Economics and Administrative Sciences
2024; 30(143), pp. 523-535
P-ISSN 2518-5764
E-ISSN 2227-703X

المقارنة بين مقدرات انموذج انحدار بيتا باستعمال MSE

ميلاد ضياء إسماعيل
جامعة بغداد / كلية الإدارة والاقتصاد / قسم الإحصاء
بغداد ، العراق
mielad.diaa2101m@coadec.uobaghdad.edu.iq
https://orcid.org/0009-0005-0055-5414/print

احمد ذياب احمد
جامعة بغداد / كلية الإدارة والاقتصاد / قسم الإحصاء
بغداد ، العراق
ahmedstatistic@ coadec.uobaghdad.edu.iq
https://orcid.org/0000-0002-1689-1376

Received:20/11/2023 Accepted:7/5/2024 Published Online First: 1 /10/ 2024

مستخلص البحث:

يوضح تحليل الانحدار الخطي المتعدد التأثيرات بين المتغيرات التفسيرية والمتغير التابع، حيث تكمن وظيفته في تمثيل البيانات لفهم شكل وطبيعة العلاقة بين المتغيرات التفسيرية ومتغير الاستجابة، ومن اهم المشاكل التي يعاني منها الانحدار الخطي هي طبيعة البيانات اذ انها في بعض الأحيان لا تتبع التوزيع الطبيعي وذلك أدى الى ظهور أنواع مختلفة من نماذج الانحدار الخطي كأنموذج انحدار بيتا، اذ تكون فيه بيانات المتغير المعتمد أصغر من الواحد وتكون محصورة ضمن الفترة (0,1) وتكون بياناته نسبية وعشرية. وسوف يتم استعمال طرائق غير تقليدية لتقدير معلمات انحدار بيتا وذلك لكون بعض الطرائق التقليدية لا تعطي نتائج مناسبة ودقيقة عند وجود احد اهم المشاكل التي يعاني منها الانحدار وهي مشكلة التعدد الخطي ولمعالجة هذه المشكلة يتم استعمال طريقة الإمكان الأعظم (MLE) وطريقة انحدار الحرف وطريقة ليو المعدلة وطريقة (Ozkale And Kaciranlar) مع اقترح معلمات تقلص لكل طريقة واجراء مقارنة بين الطرائق باستعمال معيار المقارنة متوسطات مربعات الخطأ وذلك بأجراء تجربة المحاكاة بأسلوبين، وقد بينت نتائج المحاكاة ان طريقة انحدار الحرف هي الأفضل وذلك بالنسبة للأسلوب الأول والأسلوب الثاني للمحاكاة.

نوع البحث: ورقة بحثية.

المصطلحات الرئيسة للبحث: انموذج انحدار بيتا، التعدد الخطي، طريقة الإمكان الأعظم، طريقة انحدار الحرف، طريقة ليو المعدلة، طريقة (Ozkale And Kaciranlar)، المحاكاة.