Journal of Mathematics Education and Application (JMEA)
Vol. 2, No 3, Oktober 2023, pp. 117–129 DOI: https://doi.org/10.30596/jmea.v2i3.13327

Comparison of Newton Raphson Method And Ridge Method In Probit Regression Parameter Estimation

Yastri[1*], Rahmawati Pane [2]
[1]Bachelor Faculty of Mathematics and Natural Sciences, Universitas Sumatera Utara, Indonesia
[2]Lecturer atFaculty of Mathematics and Natural Sciences, Universitas Sumatera Utara, Indonesia
*Corresponding Author. E-mail: yastriasri123@gmail.com

Article Info

Article History
Received : 07 Juli 2023
Accepted: 31 Oktober 2023
Published: 31 Oktober 2023

Keywords:
Newton Raphson Method, Parameter Estimation, Probit Regression, Ridge Method

ABSTRACT

Probit regression model is a non-linear model used in the process of analyzing the relationship between a response variable that has categorical properties. The problem that is very often experienced in probit regression when the predictor variable consists of one or more is that there is a very high correlation between predictor variables called multicollinearity. To overcome this, the Newton Raphson method and the Rigde method are used. So this research was conducted to compare the Newton Raphson method and the Ridge method in the estimation of the Probit Regression parameter. The data used in this research is 1000 data generation that contains multicollinearity. Based on this research, the estimated mean square error of the Probit Regression model using the Newton Raphson method is 0.488. The estimation result of the mean square error of the Probit Regression model using the Ridge method is 0.488. The results of this study indicate that the estimation of the Probit Regression parameter using the Newton Raphson method is as good as the Ridge method. This can be seen from the estimated value of MSE using the Newton Raphson method and the Ridge method.

To cite this article:

INTRODUCTION

The Newton Raphson method is a good form of numerical method in determining the roots of an equation. This method will always converge if the initial point selection is within the settlement area. The drawback of this method is that The calculation process requires the derivative function $f'(x)$ of a function $f(x)$whose roots are to be determined. Regression analysis is a statistical method that is most often used in our daily lives. Regression analysis is a technique used in model development that connects response variables (Y) with predictor variables (X) and is used to predict future values. Estimation is made to see the impact that will occur if a change in one variable affects another variable, so that anticipatory handling can be carried out in overcoming the impact of this change.

According to Imam Ghozali (2011), the process of testing the classical assumptions on the form of the linear regression model used is done in order to know whether the regression model is a very good model or not. The purpose of testing the classical assumptions is to provide certainty that the regression equation which can be based on the model has a value of accuracy in estimation, is not biased, and is

Journal of Mathematics Education and Application (JMEA)
Vol. 2, No 3, Oktober 2023, pp. 117–129 DOI: https://doi.org/10.30596/jmea.v2i3.13327

consistent. In carrying out the regression analysis process, it is necessary to first test the assumptions. Some of the assumptions that must be fulfilled in the regression analysis process include: Assumption of normality, Assumption of homoscedasticity, Assumption of non-autocorrelation, Assumption of non-multicollinearity, and Assumption of linearity.

The Probit Regression Model is a non-linear model that is used in the process of analyzing the relationship between a response variable that has categorical properties, that is, a value of 1 for the presence of a characteristic and 0 for the absence of a characteristic, with one or several predictor variables that have numerical, categorical properties. or a combination of both.

The problem that is very often experienced in probit regression when the predictor variable consists of one or more is that there is a very high correlation between the predictor variables which is called multicollinearity. If multicollinearity occurs, then the classical assumption is not fulfilled. This has an impact on the estimator obtained is not efficient, causing the variance of a regression coefficient to be large and not minimum (Gujarati, 2006). Multicollinearity can produce regression coefficient values, intervals and standard *error* which is large and the variable relationship is not significant. The use of ridge regression is better when used to solve multicollinearity assumption problems because the resulting coefficient estimators tend to be more stable than the least squares method (Chatterjee and Hadi, 2006).

Ridge regression has an advantage over other methods because it can reduce the effect of multicollinearity by finding a biased estimator that has a smaller variance than the variance of the multiple linear regression estimator (Pratiwi, 2016). The process of the Ridge regression method is carried out in a way similar to the least squares method. Ridge regression uses the bias constant in the least squares equation so that the coefficient results are reduced and tend to be close to zero (Hestie, 2008).

It is difficult to estimate the parameters of the probit regression model using the maximum likelihood method to obtain the maximum point. This is because the form of the likelihood function is complex so it cannot be differentiated. To overcome these problems, the Newton Raphson method is used. The Newton Raphson method can solve complex function problems with its approach. Therefore, forecasting with loglikelihood is a very good approximation and close to its maximum point (Storvik, 2011). Parameter estimation is the estimation of population characteristic values (parameters) based on the sample characteristic values. The parameter estimation is classified into two, namely the first is point estimation and the second is interval estimation.

Regression Analysis

Linear regression analysis has two types, namely simple linear regression or multiple linear regression. In simple linear regression analysis, the predictor variable used is one. Whereas in multiple linear regression analysis, the predictor variable used is more than one (Wasilane, 2014).

The model of simple linear regression is:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \,, \quad i = 1,2,\dots,n \tag{1}$$

The model of multiple linear regression is:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i \tag{2}$$

Probit regression

Probit regression is a form of nonlinear regression that is used to analyze the relationship between one response variable and its predictor variable, where the response variable is qualitatively dichotomous,

Vol. 2, No 3, Oktober 2023, pp. 117–129 DOI: https://doi.org/10.30596/jmea.v2i3.13327

which means it has a value of 1 which indicates the presence and value 0 denotes absence (Candra, 2009:3). Probit regression namely the development of a logistic regression model by making the logistic regression equation have a normal distribution (Widhiarso, 2012).

Probit regression is expressed in the form Φ (Z). The Φ symbol describes the function of the standard deviation of the normal distribution.

$$P(Y=1)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{Z^2}{2}}dZ=\Phi(Z) \tag{3}$$

The probit regression model is

$$y_i^*=x_i^{'}\beta+\varepsilon_i \tag{4}$$

Multicollinearity

Multicollinearity shows that there is a linear relationship with a perfect level between the predictor variables in the regression model. Multicollinearity is a condition where there is a very strong correlation between the predictor variables (X) involved in the development of the linear regression model. Collinearity means a single linear relationship, while multicollinearity means there is more than one almost perfect linear relationship. In its application, it is often not distinguished by either one relationship or more, the name multicollinearity is used (Supranto, 2005).

Multicollinearity or poor conditioning can lead to inaccurate estimates of the regression coefficients, inflate the standard errors of the regression coefficients, deflate the partial *t* test for the regression coefficients, give incorrect and insignificant *p*-values, and decrease the predictability of the model. It also causes a change in the direction of the estimated coefficient signs (Saleh, et al 2019).

One measure to test the effect of multicollinearity (Setiawan and Kusrini, 2010) is*Variance Inflation Factors* (VIF). The VIF value is calculated by the equation, namely:

$$VIF=diag(X^{'}X)=\frac{1}{1-R^2} \tag{5}$$

Another way to determine multicollinearity with value*Tolerance* (TOL) with the equation

$$TOL\ =\frac{1}{VIF} \tag{6}$$

Parameter Estimation

Estimation is a value for estimating the relationship regarding the value of a population parameter whose value is unknown using a sample (statistics), thus a random variable is obtained from a population in question.$Y=X\beta+\varepsilon$estimate parametersβ is$\hat{\beta}=(X^TX)^{-1}(X^TY)$.

***Mean Square Error* (MSE)**

Vol. 2, No 3, Oktober 2023, pp. 117–129 DOI: https://doi.org/10.30596/jmea.v2i3.13327

Mean Square Error (MSE) is another method for evaluating estimation methods (Suryaningrum & W, 2015). MSE is a way to measure the overall estimation error.This method can be used to see the error rate of a model estimated from existing data.

$$MSE = \frac{(Y-\hat{Y})^T(Y-\hat{Y})}{n}$$

Metode Ridge

Ridge regression is a method that can be used in dealing with multicollinearity cases. Hoerl and Kennard (1970) formulated the Ridge estimator for the first time to deal with multicollinearity problems in a linear regression by displaying the results of multiple linear regression values which produce matrix values X^'X having a determinant close to zero resulting in an unstable parameter estimator. The Estimator of the Ridge regression parameter ($\hat{k}$) utilizes the Least Squares (OLS) method, namely by adding a small positive number (ε) to the diagonal value of the matrix X^' X, so that bias can be overcome. The ε value is between 0 and 1, so that the Ridge regression estimator will still be biased towards the β parameter, but tends to be stable (Sunyoto, 2009).

The Ridge estimator value is obtained by minimizing the sum of squares*error* using the OLS method:

$$\varepsilon_i = Y_i - X\hat{\beta}$$
$$\sum \varepsilon_i{}^2 = \sum (Y_i - X\hat{\beta})^2 \quad (8)$$

Estimator $\hat{\beta}$obtained by using the OLS method, namely by minimizing the sum of squares*error* and equate it to zero as follows:

$$\sum_{i=1}^{n} \varepsilon_i{}^2 = \varepsilon_1{}^2 + \varepsilon_2{}^2 + \cdots + \varepsilon_n{}^2$$
$$= \varepsilon^T \varepsilon$$

So that residue is:

$$\varepsilon_i = y - y_i$$
$$\sum \varepsilon_i{}^2 = \sum_{i=1}^{n} (y - y_i)^2$$

Temporary:

$$\varepsilon^T \varepsilon = (Y - X\beta)^T (Y - X\beta)$$

For:

$$S = Y^T Y - Y^T X\beta - \beta^T X^T Y + \beta^T X^T X\beta$$

To determine the linear regression estimator, it can be calculated from the first derivative of equation (8) partially to and equated with zero results.

$$\frac{\partial(S)}{\partial \beta} = \frac{\partial(Y^T Y - Y^T X\hat{\beta} - \hat{\beta}^T X^T Y + \hat{\beta}^T X^T X\hat{\beta})}{\partial \hat{\beta}^T}$$
$$\frac{\partial(S)}{\partial \beta} = -X^T Y + X^T X\hat{\beta}$$
$$\frac{\partial(S)}{\partial \beta} = 0$$
$$-X^T Y + X^T X\beta = 0$$

Vol. 2, No 3, Oktober 2023, pp. 117–129 DOI: https://doi.org/10.30596/jmea.v2i3.13327

$$X^T X\beta = X^T Y$$
$$\hat{\beta} = (X^T X)^{-1}(X^T Y) \tag{9}$$

By using the Lagrange multiplier denoted by the Lagrange function G is defined as:

$$G = \sum {\varepsilon_i}^2 + kg \tag{10}$$

with limited conditions

$$g = \hat{\beta}^T\hat{\beta} - c^2$$

for

$$G = (Y^T Y - Y^T X\hat{\beta} - \hat{\beta}^T X^T Y + \hat{\beta}^T X^T X\hat{\beta}) + k(\hat{\beta}^T\hat{\beta} - c^2)$$

that meet the requirements

$$\frac{\partial G}{\partial \beta} = 0$$

Where *k* the Lagrange multiplier is not affected by $\hat{\beta}$ and *k* which is a finite positive constant.

Looking for the value of β by breaking up $\frac{\partial G}{\partial \beta}$=0

$$\frac{\partial(\sum {\varepsilon_i}^2 + kg)}{\partial \hat{\beta}^T} = 0$$

$$\frac{\partial((Y^T Y - Y^T X\hat{\beta} - \hat{\beta}^T X^T Y + \hat{\beta}^T X^T X\hat{\beta}) + k(\hat{\beta}^T\hat{\beta} - c^2))}{\partial \hat{\beta}^T} = 0$$

$$\hat{\beta} = (X'X + kI)^{-1}X'Y \tag{11}$$

Vol. 2, No 3, Oktober 2023, pp. 117–129 DOI: https://doi.org/10.30596/jmea.v2i3.13327

Where $\hat{\beta} = (X^{'}X + kI)^{-1}X^{'}Y$ and $0 \le k \le \infty$, the form of the equation is called the Ridge regression estimator.

Calculate the value of k based on the following HKB formula

$$k_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\boldsymbol{\beta}}^{*\prime}\hat{\boldsymbol{\beta}}^{*}} \quad (12)$$

Maximum Likelihood Method

For example $x_1, x_2, \cdots, x_n$ is a random sample of population with density $f(x, \boldsymbol{\theta})$ where $\boldsymbol{\theta}(\theta_1, \theta_2, \cdots, \theta_k)$ is an unknown parameter function *probability* written:

$$L(\theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^{n} f(x_i, \boldsymbol{\theta})$$

Newton Raphson Method

The Newton Raphson approach is a very powerful and most widely applied numerical method for solving the problem of finding the roots of an equation $f(x) = 0$. This method have formulas, namely:

$$X_{n+1} = X_n - \frac{F(X_n)}{F'(X_n)}$$

RESEARCH METHOD

Comparing the Newton Raphson method and the Ridge method in estimating the Probit regression parameters. The first step is to study the literature by searching for literature such as theories from books, journals, previous research articles that are relevant to cases or parameter estimation problems. The next steps are to determine the appropriate problem, estimate the probit regression parameters using the Ridge method, estimate the probit regression parameters using the Newton Raphson method, generate data that contains multicollinearity, estimate parameters in the probit regression model using the ridge method on the data, perform parameter estimation in the probit regression model using the newton raphson method on the data and comparing the estimation results of the Ridge method and the Newton Raphson method.

RESULTS AND DISCUSSION

Research data

In this study, 1000 data were generated with one response variable (Y) and 2 predictor variables (X_1, X_2). Response variables (Y) are generated 0 and 1. Variables X_1 are generated from 1 to 1000 and to produce a very high correlation between variables X_1 and X_2, variables X_2 are generated by multiples of 2 of X_1 added value is ε . ε generated 1000 data distributed $N \sim (0,1)$.

Journal of Mathematics Education and Application (JMEA)
Vol. 2, No 3, Oktober 2023, pp. 117–129 DOI: https://doi.org/10.30596/jmea.v2i3.13327

Tabel 1 Awakening Data

No	y	x_1	x_2
1	0	1	3.043829
2	1	2	5.637644
3	0	3	9.586139
4	0	4	12.88854
5	0	5	13.81167
6	1	6	18.2067
7	0	7	21.26554
8	1	8	24.33165
…	…	…	…
999	0	999	2998.228
1000	0	1000	2999.617

Vol. 2, No 2, Juni 2023, pp. 62–72 DOI: https://doi.org/10.30596/jmea.v2i3.13327

Estimation of Probit Regression Model Parameters with the Newton Raphson Method

A probit regression model can be written

$$\Phi^{-1}(\pi_i) = \boldsymbol{X_i'\beta}$$

Estimation of the parameters of the probit regression model can be obtained using the method*Maximum Likelihood Estimation* (MLE). Parameter estimation using this method starts with defining the function*probability*and functions*log likelihood* as follows:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n}\left[\left(\Phi(\boldsymbol{X_i'\beta})\right)^{\boldsymbol{y_i}}\left(1-\Phi(\boldsymbol{X_i'\beta})\right)^{1-\boldsymbol{y_i}}\right]$$

$$LL = ln\ L(\boldsymbol{\beta}) = \sum_{i=1}^{n}\{\boldsymbol{y_i}[\ln\Phi(\boldsymbol{X_i'\beta})] + (1-\boldsymbol{y_i})\,\mathbf{ln}[1-\Phi(\boldsymbol{X_i'\beta})]\}$$

Then, by optimizing the function*log-likelihood*, find the first partial derivative of the function*log-likelihood* and equated to zero.

$$\frac{\partial LL}{\partial\boldsymbol{\beta}} = \frac{\partial\left(\sum_{i=1}^{n}\{y_i[\ln\Phi(\boldsymbol{X_i'\beta})] + (1-y_i)\ln[1-\Phi(\boldsymbol{X_i'\beta})]\}\right)}{\partial\boldsymbol{\beta}} = 0$$

$$= \sum_{i=1}^{n}\left[\frac{q_i\phi(\boldsymbol{X_i'\beta})}{\Phi(\boldsymbol{X_i'\beta})}\right]\boldsymbol{X_i} \ = 0$$

$= \sum_{i=1}^{n}\lambda_i\,\boldsymbol{X_i}$ =0

The numerical method that can be used is the Newton-Raphson method. This method requires a second partial derivative of the function*log-likelihood* with respect to the parametersestimated:

$$\frac{\partial^2 LL}{\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}'} = -\sum_{i=1}^{n}-\lambda_i(\lambda_i + \boldsymbol{X_i'\beta})\,\boldsymbol{X_i}\ \boldsymbol{X_i'}$$

The equation used to obtain the parameter estimator using the Newton-Raphson method is:

$$\widehat{\boldsymbol{\beta}}^{(n)} = \widehat{\boldsymbol{\beta}}^{(n-1)} + \boldsymbol{I}\ \ [\widehat{\boldsymbol{\beta}}^{(n)}]\boldsymbol{g}[\widehat{\boldsymbol{\beta}}^{(n)}]$$

Estimation of the Probit Regression Model Parameters with the Ridge Method

The equation used to obtain the parameter estimator with the Maximum method*probability*with the help of Newton Raphson

$$\boldsymbol{I}[\widehat{\boldsymbol{\beta}}] = E\left[\frac{\partial^2 LL}{\partial\beta_j\,\partial\beta_k}\right]$$

$$\boldsymbol{I}[\widehat{\boldsymbol{\beta}}] = E\left[\sum_{i=1}^{N}\frac{y_i-\pi_i}{\pi_i(1-\pi_i)}x_{ji}\left(\frac{\partial\pi_i}{\partial(x_i'\beta)}\right)\sum_{I=1}^{N}\frac{y_I-\pi_I}{\pi_I(1-\pi_I)}x_{kI}\left(\frac{\partial\pi_I}{\partial(x_I'\beta)}\right)\right]$$

Vol. 2, No 2, Juni 2023, pp. 62–72 DOI: https://doi.org/10.30596/jmea.v2i3.13327

$$= \sum_{i=1}^{N} \left[\frac{x_{ji}\, x_{ki}}{\pi_i(1-\pi_i)} \left(\frac{\partial \pi_i}{\partial(x_i'\beta)} \right)^2 \right]$$

In matrix form it can be written as follows:

$$\boldsymbol{I}[\widehat{\boldsymbol{\beta}}] = X'WX$$

maximum estimate is obtained*probability*with the Newton Raphson approach

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{NR}} = (\boldsymbol{X}'\widehat{\boldsymbol{W}}\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\widehat{\boldsymbol{W}}\hat{\boldsymbol{z}})$$

When the explanatory variable is*collinear*, some eigenvalues will be small, which inflates the MSE value. In this situation, the probit regression estimator with the Ridge method is a better alternative.The estimators using the probit method are as follows:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{RR}} = (\boldsymbol{X}'\widehat{\boldsymbol{W}}\boldsymbol{X} + \boldsymbol{kI})^{-\mathbf{1}}(\boldsymbol{X}'\widehat{\boldsymbol{W}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{NR}})$$

Data Multicollinearity Test

The results of the multicollinearity test between variables $X_1$and X_2

$$R^2 = \frac{(n(\sum X_1X_2) - (\sum X_1 \sum X_2))^2}{(n(\sum X_1^2) - (\sum X_1)^2)(n(\sum X_2^2) - (\sum X_2)^2)}$$

$$R^2 = \frac{(1000(1001490258) - (751484419422))^2}{(1000(333833500) - (500500)^2)(1000(300441067) - (1501467)^2)}$$

R^2 =0.9999993

The value of the coefficient ofdetermination(R^2)of 0.9999993

$$VIF = \frac{1}{1-R^2}$$

$$VIF = \frac{1}{1 - 0.9999993}$$

$$VIF = 1428571{,}4$$

$$TOl = \frac{1}{1428571{,}4}$$

$$TOL = 0.0000007$$

In the generated data, values are obtained VIF > 10 and TOL < 0.1 which indicates multicollinearity between predictor variables. From the results of the above calculations, the values VIF =1428571,4 and TOL =0.0000007 .

Estimation of Probit Regression Model Parameters with the Newton Raphson Method on Generation

Vol. 2, No 2, Juni 2023, pp. 62–72 DOI: https://doi.org/10.30596/jmea.v2i3.13327

Data

The parameter estimation results obtained by the newton Raphson method with the help of the R program with the number of iterations of four iterations are as follows:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{NR}} = (\boldsymbol{X}'\widehat{\boldsymbol{W}}\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\widehat{\boldsymbol{W}}\hat{\boldsymbol{z}})$$

$$\widehat{\boldsymbol{\beta}}_{NR} = \begin{bmatrix} -0.07668 \\ 0,16221 \\ -0,05402 \end{bmatrix}$$

So that the estimation results of the probit regression model are as follows:

$$\boldsymbol{y}_{\boldsymbol{i}}^{*} =- 0.07668 + 0,16221x_1 - 0,05402x_2$$

Based on the identification results of the Probit Regression model, a comparison was made between the data and the estimation results. From these results, the MSE calculation was carried out. The results of the comparison between the data and the estimation results of the Newton Raphson method

Table 2 Comparison Data with Newton Raphson Method Estimation Results

No	Data (y)	Estimasi ($\hat{y}$)	$(y-\hat{y})'(y-\hat{y})$
1	0	0	0
2	1	0	1
3	0	0	0
4	0	0	0
5	0	0	0
6	1	0	1
7	0	0	0
8	0	1	1
9	0	0	1
10	1	0	1
…	…	…	…
1000	0	1	1
	Jumlah		488

Based on the estimated parameters and models obtained, the value is then calculated*Mean Square Error*-his. The results of estimating the MSE value are as follows:

$$MSE = n^{-1}(y-\hat{y})'(y-\hat{y})$$

$$MSE = 0.488$$

Estimation of Probit Regression Model Parameters with the Ridge Method on Generation Data

The first step is to define the ridge (k_{HKB}) method's lagrange multiplier. So that the value is obtained (k_{HKB}).

 DOI: https://doi.org/10.30596/jmea.v2i3.13327

$$k_{HKB} = \frac{p\hat{\sigma}^2}{\widehat{\boldsymbol{\beta}}^{*\prime}\widehat{\boldsymbol{\beta}}^{*}}$$

K_{HKB} = 0,001000667

the results obtained from the estimation of the parameters of the Ridge method and the value$\widehat{W}$ obtained with the help of the R program are as follows:

$$\hat{\beta}_{RR} = (X'\widehat{W}X + kI)^{-1}(X'\widehat{W}X\hat{\beta}_{NR})$$

$$\widehat{\boldsymbol{\beta}}_{RR} = \begin{bmatrix} -0.07667598 \\ 0,1622049 \\ -0,05401504 \end{bmatrix}$$

So that the estimation results of the probit regression model are as follows:

$$\boldsymbol{y_i^*} =- 0.07667598 + 0,1622049x_1 - 0,05401504x_2$$

Based on the identification results of the probit regression model, a comparison was made between the data and the estimation results. From these results, the MSE calculation was carried out. The results of the comparison between the data and the estimation results of the Ridge method are as follows:

Table 3 Comparison Data with Ridge Method Estimation Results

No	Data (y)	Estimasi ($\hat{y}$)	$(y-\hat{y})'(y-\hat{y})$
1	0	0	0
2	1	0	1
3	0	0	0
4	0	0	0
5	0	0	0
6	1	0	1
7	0	0	0
8	0	1	1
9	0	0	1
10	1	0	1
…	…	…	…
1000	0	1	1
	Jumlah		488

Based on the estimated parameters and models obtained, the value is then calculated*Mean Square Error*-his. The results of estimating the MSE value are as follows:

$$MSE = n^{-1}(y-\hat{y})'(y-\hat{y})$$

Vol. 2, No 2, Juni 2023, pp. 62–72 DOI: https://doi.org/10.30596/jmea.v2i3.13327

$$MSE = 0.488$$

Comparison of the Newton Raphson Estimation Method and the Ridge Method

Parameter estimation values and mean square error Probit regression using the Newton Raphson method and the Ridge method can be presented in tabular form as follows:

Table 4 Parameter Estimation and*Mean Square Error* Using Newton's Method Raphson and the Ridge Method Parameter Metode Newton Raphson Metode Ridge

Parameter	Metode Newton Raphson	Metode Ridge
$\hat{\boldsymbol{\beta}}_0$	-0.07668	-0.07667598
$\hat{\boldsymbol{\beta}}_1$	$0,16221$	$0,1622049$
$\hat{\boldsymbol{\beta}}_2$	$-0,05402$	$-0,05401504$
$\boldsymbol{MSE}$	0,488	0,488

From Table 4 after the simulation with the R program, the estimation results are obtained*MSE* using the Newton Rapson method is 0.488. While estimation*MSE* using the Ridge method is 0.488. It can be seen that the value*MSE* in both methods have the same value

CONCLUSION

In data simulation using the Probit Regression model, the estimation results are obtained*mean square error* using the Newton Raphson method is 0.488.In the data simulation using the Probit Regression model, the estimation results are obtained*mean square error* using the Ridge method is 0.488.The results of this study indicate that the estimation of the Probit Regression parameter using the Newton Raphson method is as good as the Ridge method. This can be seen from the estimated value*MSE* using the Newton Raphson method is the same as the Ridge Method which is equal to 0.488

REFERENCES

Candra Y. (2009). *Pembentukan Model Probit Bivariat*. Semarang: Universitas Diponegoro

Chatterjee S, Hadi AS. (2006) *Regression Analysis by Example*. 4 ed. USA: John Wiley & Sons

Ghozali, Imam. (2011). *Aplikasi AnalisisMultivariate dengan Program IBMSPSS 21*. Semarang: Badan PenerbitUniversitas Diponegoro.

Gujarati DN. (2006). *Dasar Dasar Ekonometrika* Jilid I. Jakarta: Erlangga.

Vol. 2, No 2, Juni 2023, pp. 62–72 DOI: https://doi.org/10.30596/jmea.v2i3.13327

Hastie Te. (2008). *The Elemen of StatisticalLearning. Data Mining, Inference, andprediction*. Edisi Kedua. New York:Spring.

Hoerl AE, Kennard RW. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55-67.

Pratiwi N. (2016). Perbandingan RegresiKomponen Utama dengan RegresiRidge untuk Mengatasi MasalahMultkikolinearitas. Skripsi S1.Semarang: Fakultas Matematika danIlmu Pengetahuan Alam UniversitasNegeri Semarang.

Saleh AME, Arashi M, Kibria BG. (2019). *Theory of ridge regression estimation with applications* (Vol. 285). John Wiley & Sons.

Setiawan, Kusrini DE. (2010).*Ekonometrika*. Yogyakarta: Andi

Sunyoto. (2009). Regresi Logistik Ridge : Pada Keberhasilan Siswa SMA Negeri 1 Kediri Diterima Di Perguruan Tinggi Negeri. Surabaya: Institut Teknologi Sepuluh Nopember

Supranto J. (2005). *Ekonometri Buku I*. Bogor: Ghalia Indonesia

Suryaningrum KM, W SP. (2015). Analisis dan Penerapan Metode Single Exponential SmoothinguntukPrediksi Penjualan pada Periode Tertentu (Studi Kasus: PT. Media Cemara Kreasi). Prosiding Senatif, (2), 259–266.

Storvik G. (2011). Numerical optimization of likelihoods: additiona literature for STK2120. University of Oslo

Wasilane Te. (2014). Model Regresi Ridge Untuk Mengatasi Model Regresi Linier Berganda Yang Mengandung Multikolinearitas. *Jurnal Barekeng*, 31- 37

Widhiarso W. (2012). *Berkenalan dengan Regresi Probit*. Yogyakarta: Universitas Gadjah Mada.