Journal of Economics and Administrative Sciences
2024; 30(143), pp. 510-522

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/10.33095/1xrve980

A Comparison between Methods for Estimating the Restricted Gamma Ridge Regression Model Using the Simulation

Mevan Shakur Abdullah *
Department of Statistics,
College of Administration and Economics,
University of Baghdad,Iraq
Mivan.shakur2101m@coadec.uobaghdad.edu.iq
https://orcid.org/0009-0002-9567-4624
***Corresponding author**

Suhail Najim Abood
Department of Statistics,
College of Administration and Economics,
University of Baghdad,Iraq
suhnaj2005@coadec.uobaghdad.edu.iq

Received: 11/2/2024 Accepted: 17/3/2024 Published Online First: 1 /10/ 2024

Abstract :

In this paper, we discuss estimating the parameters of the restricted gamma ridge regression model by combining gamma ridge regression with restricted maximum likelihood. The characteristics of the new estimator and its superiority over the restricted gamma ridge regression estimator and restricted maximum likelihood will be identified, using several formulas for the shrinkage factor k, and it will also be Using the Monte Carlo simulation method to generate data that suffers from the problem of multicollinearity with different sizes (n=25,50,100,250) in light of other influential factors (degree of correlation, number of explanatory variables), and subjecting the parameters to linear restrictions, to get rid of the problem of multicollinearity in light of the subjection of the parameters to the model has linear constraints and the model parameters will be estimated using four estimation methods that rely on the mean square error (MSE) as a standard for comparison between the estimation methods, Through the results of the simulation experiment it was shown that the compound estimator method is the best way to estimate the parameters of the finite gamma regression model .

Paper type : Research paper

Keywords: Restricted Gamma Ridge Regression (RGRR), Gamma Ridge Regression (GRR), Restricted Maximum Likelihood Estimator (RMLE), Shrinkage factor k, Mean Square Error (MSE).

1. Introduction :

The Gamma Regression Model (GRM) is considered one of the commonly used models in the field of economics and medicine, in addition to several other fields, with the presence of multicollinearity between the explanatory variables an ordinary ridge regression ORR used, and because the explanatory variable is positively skewed and follows the Gamma distribution this method has proven ineffective because it gives high variances. Therefore, the Gamma ridge Regression method (GRRE) is used to estimate the parameters of the restricted gamma ridge regression model, but when the parameters fall under the influence of a linear constraint Rβ=r, since R is m × p, which is a known matrix. And r is a vector of known elements m × 1, so in this case we resort to the restricted maximum likelihood (RMLE) to get rid of the effect of restrictions imposed on the model parameters, and with the problem multicollinearity In this case, we resort to using a method that combines constrained maximum likelihood and gamma ridge regression to overcome the problem of multicollinearity in the presence of restrictions imposed on the parameters.

1.1 Literature review:

Many research papers have been published on the study of choosing the best estimator for ridge regression with regard to linear regression, multiple regression, and other types of regression models, as well as with regard to the gamma ridge regression model. The following are the most prominent published researches on this topic:

Francis et al (2016) suggested the restricted Liu estimator to find a regression model parameter estimate in the presence of the multicollinearity problem, and they assumed the constraint Rb=r. The properties of this estimator were compared with the properties of the restricted maximum likelihood estimator RMLE, and the effectiveness of the restricted Liu estimator was demonstrated.

El-Gammal (2018) proposed the gamma ridge model by proposing a modification of the estimator with the gamma ridge regression model. The gamma regression model is considered common in practical application when the data are positively skewed in order to overcome the problem of multicollinearity, which has a negative impact on the variance of the model's estimators.

Qasim,et al (2018) used the maximum likelihood method to estimate the unknown gamma regression parameters in the presence of the multicollinearity problem. It was noted that the variance of the estimator using the maximum likelihood method MLE was exaggerated, so the Liu estimator was used, as this estimator is considered an important estimate to address the problem of multicollinearity in gamma regression.

Amin et al (2020) proposed some ridge estimators for the gamma regression model GRM, which is considered a special form of the generalized linear model (GLM) in which the response variable is positively skewed and suitable for the gamma distribution, and the maximum likelihood method ML is considered it is the most widely used method for estimating GRM coefficients if the explanatory variables are not related. However, if the explanatory variables are related, ML is unable to estimate GRM coefficients. Researchers have proposed ridge estimates as a method to address the problem of multicollinearity or correlation between variables.

Mahmoudi et al (2020) proposed improved estimators based on the initial test and Stein-type strategies for estimating parameters in the gamma regression model, two penalty estimators were introduced, such as lasso and ridge regression.

Qasim et al (2021) proposed a new estimator for beta series regression (BRR) as a treatment for the instability of MLE due to the presence of the multicollinearity problem, and Monte Carlo simulation was used as a tool to evaluate the performance of BRR and MLE.

Yassin,et al (2022) studied the estimation of the parameters of the gamma regression model using the hillslope regression estimates used in estimating the parameters of the linear regression model in the presence of the multicollinearity problem and generalizing them to the gamma regression model. These estimators were compared with Least squares estimator and prove the effectiveness of these estimators.

The research problem is summed up in estimating the parameters of the restricted gamma ridge regression model when the explanatory variable are related to each other, as it is difficult to reach sound estimates of the model parameters using ordinary estimation methods such as the maximum likelihood method MLE, Especially in light of imposing restrictions on the model's parameters, as it will give estimates with high variances and thus it will not be possible to the researcher is able to know which explanatory variables have an impact on the regression model.

This research aims to find the best estimator for the restricted gamma ridge regression model by combining the gamma ridge regression GRR method with the restricted maximum liklihood RMLE to obtain a new estimator for the model. The mean square error MSE will be used as a comparison standard to test the effectiveness of the new estimator.

2. Materials and Methods:

2.1 Gamma Regression Model:

Gamma Regression Model GRM is considered an extension of the topic of generalized linear models GLM, as generalized linear models differ from the well-known linear regression in that the distribution of the dependent variable is required to belong to the exponential family and that the expected values μ_i are for the random variable Y, It is replaced by a link function η_i=g(μ_i) and η is a linear combination of independent variables. The goal of the link function is to make the error variance more stable. In addition, the error distribution of the model can be chosen in a way that is independent unlike linear regression, which should be the error distribution normal distribution (Hardin and Hilbe, 2007; Qasim et al. 2021).

Because the gamma distribution is a specific form of the family of exponential distributions, the scientists Hardin and Hilbe in 2007 formulated the equation of the probability density function for the gamma distribution to become as in equation (1)(Algamal 2018 ; Amin et al. 2020) :

$$f(Y,\mu,\phi) = \frac{1}{\Gamma(\phi)^{-1}}\left(\frac{1}{\mu\phi}\right)^{\phi^{-1}} Y^{\phi^{-1}-1} e^{-Y/\mu\phi} \qquad Y \geq 0 \qquad (1)$$

Assuming that $\alpha = \phi^{-1}$; $\beta = \mu\phi$

ϕ: dispersion parameter

μ: arithmetic mean

The probability mass function for the exponential family is given as in equ(2):

$$f(Y,\theta,\phi) = \exp\left[\frac{Y\theta - b(\theta)}{a(\phi)} + c(Y,\phi)\right] \qquad (2)$$

θ: location parameter

b(θ): cumulative distribution function (c.d.f) of the variable Y if

b(θ)= - ln (μ) θ = 1/μ

a(ϕ)= - ϕ: constant value required to estimate the standard error

$$C(Y,\phi) = \frac{1-\phi}{\phi}\ln(Y) - \frac{\ln(\phi)}{\phi} - ln\Gamma(\frac{1}{\phi})$$

Equation (2) can be written as in equ(3):

$$f(Y,\theta,\phi) = \left[\frac{y/\mu-(-\ln(\theta))}{-\phi} + \frac{1-\phi}{\phi}\ln(Y) - \frac{\ln(\phi)}{\phi} - \ln\Gamma(\frac{1}{\phi})\right] \qquad (3)$$

Therefore, the expectation and variance take the following form:

$$E(Y) = b'(\theta) = \frac{\partial b}{\partial \mu}\frac{\partial \mu}{\partial \theta} = \frac{-1}{\mu} \times (-\mu^2) = \mu \qquad (4)$$

$$b''(\theta) = \frac{\partial^2 b}{\partial \mu^2}\left(\frac{\partial \mu}{\partial \theta}\right) + \frac{\partial b}{\partial \mu}\left(\frac{\partial^2 \mu}{\partial \theta^2}\right) = 0(1)^2 + (1)(-\mu^2) = -\mu^2$$

$$\text{Var (Y)} = a(\phi)V(\mu) = -\phi\, b''(\theta) = (-\phi)(-\mu^2) = \phi\,\mu^2 \quad (5)$$

The link function for the mean of the explanatory variable Y, which follows the gamma model, can be written in the equation(6) form (Qasim et al,2021 ; Yasin et al, 2022;Kamary et al,2023):

$$g(\mu_i) = \left(\frac{1}{\mu_i}\right) = x_i^T\beta \qquad i=1, 2, \ldots, n \quad (6)$$

x_i: rows of the variable matrix X=$(x_{i1}, x_{i2}, \ldots, x_{ip})'$
X: explanatory variable matrix of degree n × (p+1)
β: the normal vector of regression coefficients of degree (p+1)×1
The maximum likelihood function MLE:

$$L=I(y_i,\phi)=\sum_{i=1}^{n}\left\{\left(\frac{y_i/\mu_i-(-\ln(\mu_i))}{-\phi}\right)+\frac{1-\phi}{\phi}\ln(y_i)-\frac{\ln(\phi)}{\phi}-\ln\Gamma\left(\frac{1}{\phi}\right)\right\} \quad (7)$$

The MLE is calculated by means of the reweighted least squares algorithm IRLS

$$\hat{\beta}_{ML} = (X^T\widehat{W}X)^{-1}\; X^T\,\widehat{W}y^* \quad (8)$$

Since:

$$\widehat{W} = \text{diag}\{(1/\mu_i^2)\mu_i^2\} \quad \text{and} \quad y^* = \hat{\eta}_i + (y_i - \hat{\mu}_i)/\hat{\mu}_i^2$$

They are modified variables that use an inverse link function(Abdeljabbar 2020):

$$\hat{\mu}_{i=}\,[x_i^T\,\hat{\beta}_{ML}\,]^{-1} \quad \eta_i = x_i'\beta \quad , \quad i = 1,2,\ldots,n$$

The covariance matrix for $\hat{\beta}_{ML}$is:

$$\text{Cov}\,(\hat{\beta}_{ML}) = \hat{\phi}\,(X^T\widehat{W}X)^{-1} \quad (9)$$

The estimated value of the dispersion coefficient $\hat{\phi}$ is calculated by the equation(10):

$$\hat{\phi} = \frac{1}{n-q}\sum_{i=1}^{n}\frac{(y_i-\hat{\mu}_i)^2}{(\hat{\mu}_i)^2} \quad (10)$$

Where q=p+1
P : Number of independent variables
The mean square error MSE is extracted:

$$E(L_{ML}^2)=E(\hat{\beta}_{ML}-\beta)^T(\hat{\beta}_{ML}-\beta) = tr(X\widehat{W}X)^{-1} = \hat{\emptyset}\sum_{j=1}^{p}\frac{1}{\lambda_j} \quad (11)$$

λ_j: Eigenvalue of the matrix $(x'\widehat{w}x)$
The observation recorded when using the MLE method in the presence of the problem of multicollinearity is that the MSE value is large as a result of the increased correlation between the explanatory variables, which leads to the value of λ_j being small.

2.2 Ridge Regression:

The least squares method (OLS) is considered the best unbiased linear estimation method for regression model parameters because it gives parameters with the lowest variance value, but in the presence of the multicollinearity problem, using this method will give us inaccurate estimates for the regression model, because the parameters will have large variances, so the researchers proposed Horel - Bennared in 1970 developed a method to address the problem of multicollinearity by adding a small positive quantity to the diagonal elements of the information matrix $(X'X)$ and Its calculated according to the equation(12)(Sampreet 1989; Kazem 2002):

$$B_{RR} = \begin{bmatrix} b_{1RR} \\ \vdots \\ b_{pRR} \end{bmatrix} = (X'X + kI_k)^{-1}X'Y \quad (12)$$

β_{RR} represents the Ordinary Ridge Regression estimator and is symbolized by the symbol (ORR).

2.3 Gamma Ridge Regression:

The best method used to estimate the regression coefficient in the presence of the multicollinearity problem is the ridge method, and the most important advantages and disadvantages of this method are reducing the value of the mean square error MSE and increasing the value of the shrinkage factor k, In the presence of the problem of multicollinearity, the MSE value of the maximum likelihood estimates MLE is inflated and misleading. To solve this problem, the ridge regression method for generalized regression was proposed. In the same way, the scientists adopted the estimation of the regression coefficient for the gamma ridge regression model (Asar et al. 2017):

$$\hat{\beta}_{RR}=(X^T\hat{W}X + kI_p)^{-1}(X^T\hat{W}X)\ \hat{\beta}_{ML} \quad (13)$$

$$\hat{\beta}_{ML}=(X^T\hat{W}X)^{-1}X^T\hat{W}y^*$$

2.4 Restricted Maximum Likelihood Estimator:

One of the methods used to prove the efficiency of the estimator is by using previous information, such as information related to the regression coefficients. The primary goal is to estimate the coefficient β when β is subject to a linear constraint $R'_m\beta = r_m$ Since

r_m: is a constant number

R_m: is a known vector $p\times 1$

m is an independent restriction imposed on the feature vector β

With this case, a restricted estimator used for β, where the restricted maximum likelihood method RMLE gives us the largest value of the link function for the GRM gamma regression model on β under the restrictions $R'_m\beta = r_m$, and it is calculated RMLE according to the equation (14) (Sarkar 1992 ; Kurtoglu 2017 ; Qasim et al. 2021):

$$\hat{\beta}_{RMLE} = \hat{\beta}_{MLE} + A^{-1}\acute{R}\left(RA^{-1}\acute{R}\right)^{-1}\left(r - R\hat{\beta}_{MLE}\right) \quad (14)$$

2.5 Built-in estimator:

To obtain the RMLE estimator, It is done by maximizing the maximum potential function of the gamma regression model, taking into account the constraint $R_m{}'\beta = r_m$, where m=1,2,..t, and with the presence of the multicollinearity problem, the RMLE method can give us weak estimators and thus give false information as This is the case for the MLE estimator with existence of the multicollinearity problem, so itis necessary to Add modifications to the RMLE estimator in order to obtain an effective estimator under the framework of a set of linear constraints. The restricted letter regression estimator for the general regression model GLM was presented by Ozkale and Kurtoglu (2017) and itis the same estimator used for gamma regression model GRM (Qasim et al. 2018 ; Amine et al. 2020 Qasim et al. 2021):

$$\hat{\beta}_R(k)=A_k\ \hat{\beta}_{RMLE} \quad (15)$$

2.6 Estimation of the shrinkage coefficient:

It is better in practical application for the shrinkage coefficient (k) estimation to be in a way that reduces the value of the mean square error MSE of the maximum possible estimator RMLE and the combined estimator RGRRE. For this purpose, equation (16) was used (Tibshirani 1996 ; Qasim et al. 2021 ; Yasin et al. 2022):

$$EMSE=\hat{\emptyset}\sum_{j=1}^{p+1}\frac{\lambda_j}{(\lambda_j+k)^2}+k^2\sum_{j=1}^{p}\frac{\alpha_j{}^2}{(\lambda_j+k)^2} \quad (16)$$

Equation (16) is derived with respect to k and equalized to zero to find the best value for k... Researchers Qasim , Akram , Amin and Manson in (2021) proposed a number of methods to estimate the shrinkage coefficient for the purpose of choosing the best method that gives us the lowest MSE, which is :

$$k_{(1)} = \hat{k}_{mean} = \frac{\sum_{j=1}^{p}\sqrt{\frac{\widehat{\emptyset}\lambda_j m_{jj}}{\hat{\alpha}_j^{2}}}}{p} \quad (17)$$

$$k_{(2)} = \hat{k}_{median} = median\left(\frac{\widehat{\emptyset}\lambda_j m_{jj}}{\hat{\alpha}_j^{2}}\right) \quad (18)$$

$$k_{(3)} = \hat{k}_{HM} = \frac{p}{\sum_{j=1}^{p}(1/(\frac{\widehat{\emptyset}\lambda_j m_{jjj}}{\hat{\alpha}_j^{2}}))} \quad (19)$$

$$k_{(4)} = \hat{k}_{max} = max\left(\sqrt{\frac{\widehat{\emptyset}\lambda_j m_{jj}}{\hat{\alpha}_j^{2}}}\right) \quad (20)$$

$$k_{(5)} = \hat{k}_{GM} = \left(\prod_{j=1}^{p}(\frac{\widehat{\emptyset}\lambda_j m_{jj}}{\hat{\alpha}_j^{2}})\right)^{\frac{1}{p}} \quad (21)$$

2.7 Comparison between $\hat{\beta}_R(k)$ and $\hat{\beta}_{RMLE}$ when the previous constraints are true τ=r-Rβ=0

$\hat{\beta}_{RMLE}$ is an unbiased estimate when τ=0, but $\hat{\beta}_R(k)$ is a biased estimate under the GRM gamma regression model, so (Qasim et al. 2021):

$Cov\ \hat{\beta}_R(k) \le Cov\ \hat{\beta}_{RMLE} \quad \forall \quad k \ge 0$

$$[Cov(\hat{\beta}_{RMLE}) - Cov(\hat{\beta}_R(k))]$$
$$= \widehat{\emptyset}[A^{-1} - A^{-1}R'(RA^{-1}R')^{-1}RA^{-1}] - \widehat{\emptyset}[A_k[A^{-1} - A^{-1}R'(RA^{-1}R')^{-1}RA^{-1}]A'_k]$$
$$= [\widehat{\emptyset}A_k[k^2A^{-1}GA^{-1} + kGA^{-1} + kA^{-1}G]A'_k]$$
$$= \hat{\phi}G - \hat{\phi}[A_k G A_k{}'] \quad (22)$$

From equation (22) we conclude that:

$G = A^{-1} - A^{-1}R'(RA^{-1}R')^{-1}RA^{-1}$ (G= positive semidefinite (psd))

$A = X^T\widehat{W}X$ (Positive definite)

$A_k = (I_p + kA^{-1})^{-1}$

$A^{-1}G$ is a matrix whose values are definite and non-negative

$A^{-1}GA^{-1}$ psd matrix

So$[Cov(\hat{\beta}_{RMLE}) - Cov(\hat{\beta}_R(k))]$ psd matrix $\forall$ k ≥ 0

It is clear from the above that the field of variation for $\hat{\beta}_R(k)$is less than the field of variation for $\hat{\beta}_{RMLE}$

To discuss the MSE characteristic of $\hat{\beta}_R(k)$, we can show that the $\hat{\beta}_R(k)$ estimator is better than the $\hat{\beta}_{RMLE}$ estimator, and the MSE of $\hat{\beta}_{RMLE}$ can be written for the GRM gamma regression model as in equation(23):

$$MSE\ (\hat{\beta}_{RMLE}) = \widehat{\emptyset}tr(G) = \widehat{\emptyset}\sum_{j=1}^{p} m_{jj} \quad (23)$$

So m_{jj} represents the elements of the main diagonal of the matrix $M = Q'GQ$, and Q is an orthogonal matrix, so $\Lambda = Q'GQ$ and Λ is a diagonal matrix whose main diagonal elements $(\lambda_1, \lambda_2, \dots, \lambda_p)$ represent the characteristic roots of the matrix G, and the MSE of RGRRE is calculated by the equation(24):

$$MSE\left(\hat{\beta}_R(k)\right) = \widehat{\emptyset}tr\left[Cov\left(\hat{\beta}_R(k)\right)\right] + \left[Bias\left(\hat{\beta}_R(k)\right)\right]'\left[Bias\left(\hat{\beta}_R(k)\right)\right] \quad (24)$$

If we assume that the constraint value τ = 0, then equation(23) can be written as follows:

$$MSE(\hat{\beta}_R(k)) = \widehat{\emptyset}tr(A_k G A'_k) + k^2\beta'(A + kI_p)^{-2}\beta$$

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

After simplifying we arrive at the equation(24):

$$MSE\ \hat{\beta}_R(k) = \hat{\emptyset}\sum_{j=1}^{p}\frac{\lambda_j^{\,2}}{(\lambda_j+k)^2}m_{jj} + k^2\sum_{j=1}^{p}\frac{\alpha_j^{\,2}}{(\lambda_j+k)^2}$$

$$= \Upsilon_1\,(F_k) + \Upsilon_2\,(F_k) \qquad (25)$$

The value of α_j represents the elements of $Q'\beta$ and λ_j is the eigenvalue of matrix A, and the value of $\Upsilon_1\,(F_k)$ and $\Upsilon_2\,(F_k)$ represent variance and bias square for $\hat{\beta}_R(k)$ respectively, from equation (25) we can find We conclude that the amount of bias for RGRRE and GRRE is the same when the constraints imposed on the equations are true $\tau=(f\text{-}F\beta)=0$.

Also, the total variance of $\Upsilon_1\,(F_k)$ is continuous and decreases directly with respect to k, as in equation (26):

$$\frac{\partial\{\Upsilon_1(F_k)\}}{\partial k} = -2\emptyset\sum_{j=1}^{p}\frac{\lambda_j^{\,2}}{(\lambda_j+k)^3}m_{jj} \qquad (26)$$

From equation (26) we conclude that the variance is directly decreasing for k as long as

$$\frac{\partial\{\Upsilon_1(F_k)\}}{\partial k} \longrightarrow -\infty \ when\ k \to 0\ \ and\ \lambda_j \to 0$$

The square of the bias of $\Upsilon_2\,(F_k)$ is a directly increasing continuous function of k.

$$\frac{\partial\{\Upsilon_2(F_{(k)})\}}{\partial k} = 2k\sum_{j=1}^{p}\frac{\lambda_j\alpha_j^{\,2}}{(\lambda_j+k)^3} \quad k>0\,,\quad \lambda_j > 0 \qquad (27)$$

In the GRM gamma regression model, there will be a value of k > 0 with range

$$0 < k < \frac{\phi}{\left[max\left(\frac{\alpha_j^{\,2}}{\lambda_j m_{jj}}\right)\right]}$$

It will lead to the value of MSE $\beta_R(k)$ <MSE β_{RMLE} when τ=0

$$\frac{\partial\left(MSE(\hat{\beta}_R k)\right)}{\partial k} = -2\emptyset\sum_{j=1}^{p}\frac{\lambda_j^{\,2}}{(\lambda_j+k)^3}m_{jj} + 2k\sum_{j=1}^{p}\frac{\lambda_j\alpha_j^{\,2}}{(\lambda_j+k)^3}$$

$$= 2\sum_{j=1}^{p}\frac{k\lambda_j\alpha_j^{2} - \emptyset\lambda_j^{\,2}m_{jj}}{(\lambda_j+k)^3} \qquad (28)$$

It was previously pointed out that the value $of\ m_{jj} \geq 0$ and $\lambda_j > 0$ for each j=1,2,…,p, and that the value of the variance and the square of the bias are increasing and decreasing functions of k, so we can say that the value of ∂MSE(β_R(k))/∂k It will be negative if the condition is true

$$0 < k < \frac{\phi}{\left[max\left(\frac{\alpha_j^{\,2}}{\lambda_j m_{jj}}\right)\right]}$$

and that there is a certain value for k at which we proved that the built-in estimator RGRRE is better than gamma regression estimator GRRE Since it is a value

M=$\acute{Q}GQ = \Lambda^{-1} - B$

$B = \acute{Q}A^{-1}\acute{R}(RA^{-1}R)^{-1}RA^{-1}Q$

And M positive semidefinite matrix

Therefore $-b_{jj} \leq \frac{1}{\lambda_j}m_{jj} = \frac{1}{\lambda_j}$, also $b_{jj} = diag(\,B)$. It can be noted that:

$$\emptyset/\left(max\,\frac{\alpha_j^{\,2}}{\lambda_j m_{jj}}\right) \leq \emptyset/\alpha^2{}_{max}$$

So we conclude from the above that the availability of accurate prior information is of great importance because it will reduce the range value of k for the domination of the RGRRE estimator over the restricted maximum possible estimator RMLE compared to the dominance the GRRE gamma regression estimator over the MLE maximum possible estimator with the MSE comparison standard.

2.8 Comparison of $\hat{\beta}$(k) and $\hat{\beta}_R$(k) when the prior information is correct τ=(r-Rβ)=0

In this part, we will compare the estimator $\hat{\beta}_R(k)$ and $\hat{\beta}\ (k)$ when the imposed constraint *τ=r-Rβ=0* in the gamma regression model is

$Cov\ \hat{\beta}_R(k) < Cov\hat{\beta}(k) \quad \forall \quad k \geq 0 \qquad if\ \tau = 0$

if τ=0 Both $\beta\ (k)$ and $\beta_R(k)$ have the same bias magnitude and it could be written as follows:

$$Bias\left(\hat{\beta}_R(k)\right) = Bias\left(\hat{\beta}(k)\right) = -k\left(A + kI_p\right)^{-1}\beta$$

We can compare the value of the variance and covariance matrices for $\hat{\beta}_R$(k) and $\hat{\beta}$ (k)

$$[Cov(\hat{\beta}(k)) - Cov(\hat{\beta}_R(k))]$$
$$= \hat{\phi}(W(k)^{-1}AW(k)^{-1}) - \widehat{\emptyset}[A_k\{A^{-1} - A^{-1}R'(RA^{-1}R')^{-1}RA^{-1}\}A_k{}']$$
$$= \widehat{\emptyset}[W(k)^{-1}A^{-1}R'(RA^{-1}R')^{-1}RA^{-1}W(k)^{-1}] \qquad (29)$$

$since \quad W(k) = \left(A + kI_p\right), \quad A_k = W(k)A, \quad W(k) = A_kA^{-1}$, We showed that the matrix $[Cov(\hat{\beta}(k)) - Cov(\hat{\beta}_R(k))]$ It is a matrix psd ∀ k≥0, and this is sufficient to prove that the built-in estimator RGRRE is better than the gamma regression estimator GRRE (Qasim et al. 2021).

3. Discussion of Results:

In this section the simulation method is dealt with to generate data with different sizes to discuss the results of estimating the parameters of the restricted gamma ridge regression model in different methods to know the performance of these methods:

1- Generate a random variable following a gamma distribution using method:

$x_{ij} = \sqrt{1-\rho^2}u_{ij} + \rho u_{i(j+1)}$ i=1,2,…n; j=1,2,…p

2- Choose different sample sizes (n=25, 50, 100, 250) and selected default values for the parameter β so that $\sum_{j=1}^{p+1}\beta_j{}^2 = 1$

3- Generating the response variable for the gamma ridge regression model using a random sample of the gamma distribution G(α,β) and the values of the response variable Y_i were calculated:

$$\mu_i = E(y_i) = \left(\beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} + \cdots + \beta_p x_{ip}\right)^{-1} \qquad i=1,2,\ldots n$$

4- Estimating the parameters of the gamma regression model with four estimation methods

5- Generate mean square error criterion to compare between methods for estimating restricted gamma ridge regression model according to the equation (30):

$$MSE\left(\hat{\beta}\right) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\beta}_{(i)} - \beta)'\left(\hat{\beta}_i - \beta\right) \qquad (30)$$

6- The constraint value for the number of explanatory variables (p=4,6,8) will be as follows...

$$R = \begin{bmatrix} 1 & 0 & -2\ 1 & -3 & 1 & 1 & 1 \\ 1 & -1 & 1\ -1 & -3 & 1 & -2 & 1 \end{bmatrix} \qquad r = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

N: The number of times the experiment will be repeated, which will be equal to 2000

In Tables (1), (2), and (3), the results of estimating the mean square error (MSE) will be presented

Table 1 : Value of MSE for all estimation methods for number of explanatory variables p=4

					RGRRE					
N	ρ	ORR	RMLE	GRRE	K1	K2	K3	K4	K5	Best
25	0.91	0.026778	0.021055	0.040693	0.018004	0.022042	0.017416	0.019661	0.018746	K3
	0.93	0.043030	0.033834	0.065391	0.028931	0.035419	0.027985	0.031593	0.030124	K3
	0.95	0.075260	0.059176	0.114369	0.050600	0.061948	0.048947	0.055256	0.052687	K3
	0.97	0.099531	0.078260	0.151252	0.066919	0.081926	0.064732	0.073076	0.069678	K3
	0.99	0.131630	0.103499	0.200031	0.088500	0.108348	0.085608	0.096644	0.092149	K3
50	0.91	0.019413	0.015264	0.030770	0.013052	0.015980	0.012626	0.014253	0.013590	K3
	0.93	0.033954	0.026697	0.053817	0.022829	0.027948	0.022083	0.024929	0.023770	K3
	0.95	0.044904	0.035307	0.071173	0.030191	0.036962	0.029204	0.032969	0.031436	K3
	0.97	0.054561	0.042900	0.086479	0.036683	0.044910	0.035485	0.040059	0.038196	K3
	0.99	0.059385	0.046694	0.094126	0.039927	0.048882	0.038623	0.043601	0.041574	K3
100	0.91	0.009284	0.007300	0.015361	0.006242	0.007642	0.006038	0.006816	0.006499	K3
	0.93	0.012278	0.009654	0.020315	0.008255	0.010106	0.007985	0.009014	0.008595	K3
	0.95	0.016237	0.012767	0.026867	0.010917	0.013365	0.010560	0.011921	0.011367	K3
	0.97	0.021474	0.016884	0.035532	0.014438	0.017675	0.013966	0.015766	0.015033	K3
	0.99	0.028399	0.022330	0.046991	0.019094	0.023376	0.018470	0.020851	0.019881	K3
250	0.91	0.003779	0.002682	0.004610	0.001008	0.002409	0.000885	0.002005	0.001050	K3
	0.93	0.004997	0.003929	0.008641	0.003360	0.004114	0.003250	0.003669	0.003499	K3
	0.95	0.006609	0.005197	0.011427	0.004444	0.005440	0.004298	0.004852	0.004627	K3
	0.97	0.008741	0.006873	0.015113	0.005877	0.007195	0.005685	0.006417	0.006119	K3
	0.99	0.011559	0.009089	0.019987	0.007772	0.009515	0.007518	0.008487	0.008092	K3

Table 2 : Value of MSE for all estimation methods for number of explanatory variables p=6

					RGRRE					
N	ρ	ORR	RMLE	GRRE	K1	K2	K3	K4	K5	Best
25	0.91	0.088523	0.058487	0.036910	0.019536	0.023917	0.020341	0.021333	0.018897	K5
	0.93	0.142248	0.093983	0.059311	0.031392	0.038432	0.032686	0.034281	0.030366	K5
	0.95	0.248792	0.164377	0.103736	0.054905	0.067218	0.057169	0.059957	0.053111	K5
	0.97	0.329028	0.217388	0.137190	0.072611	0.088896	0.075605	0.079293	0.070239	K5
	0.99	0.435139	0.287496	0.181434	0.096029	0.117565	0.099988	0.104865	0.092891	K5
50	0.91	0.064176	0.042401	0.027909	0.014163	0.017339	0.014747	0.015466	0.013700	K5
	0.93	0.180366	0.119168	0.078439	0.039804	0.048731	0.041445	0.043467	0.038503	K5
	0.95	0.112244	0.074160	0.048814	0.024771	0.030326	0.025792	0.027050	0.023961	K5
	0.97	0.148443	0.098076	0.064556	0.032759	0.040106	0.034110	0.035773	0.031689	K5
	0.99	0.196316	0.129706	0.085375	0.043324	0.053040	0.045110	0.047310	0.041908	K5
100	0.91	0.030690	0.020277	0.013933	0.006773	0.008292	0.007052	0.007396	0.006551	K5
	0.93	0.040587	0.026816	0.018427	0.008957	0.010966	0.009326	0.009781	0.008664	K5
	0.95	0.053676	0.035464	0.024369	0.011846	0.014502	0.012334	0.012936	0.011459	K5
	0.97	0.070987	0.046901	0.032228	0.015666	0.019179	0.016312	0.017107	0.015154	K5
	0.99	0.093880	0.062027	0.042622	0.020718	0.025364	0.021572	0.022624	0.020041	K5
250	0.91	0.012492	0.007449	0.004182	0.001094	0.002614	0.001139	0.002175	0.000960	K5
	0.93	0.016521	0.010915	0.007837	0.003646	0.004463	0.003796	0.003981	0.003527	K5
	0.95	0.021848	0.014435	0.010365	0.004822	0.005903	0.005020	0.005265	0.004664	K5
	0.97	0.028895	0.019091	0.013708	0.006377	0.007807	0.006640	0.006963	0.006168	K5
	0.99	0.038213	0.025247	0.018128	0.008433	0.010324	0.008781	0.009209	0.008157	K5

Table 3: Value of MSE for all estimation methods for number of explanatory variables p=8

					RGRRE					
N	ρ	ORR	RMLE	GRRE	K1	K2	K3	K4	K5	Best
25	0.91	0.154349	0.101978	0.053797	0.034062	0.041702	0.035467	0.037197	0.032949	k5
	0.93	0.096053	0.063462	0.033479	0.021198	0.025951	0.022072	0.023148	0.020505	k5
	0.95	0.269957	0.178360	0.094091	0.059575	0.072936	0.062032	0.065057	0.057629	k5
	0.97	0.357018	0.235882	0.124436	0.078788	0.096458	0.082037	0.086038	0.076214	k5
	0.99	0.472156	0.311953	0.164566	0.104198	0.127566	0.108494	0.113786	0.100793	k5
50	0.91	0.072630	0.047987	0.025315	0.016028	0.019623	0.016689	0.017503	0.015505	k5
	0.93	0.204126	0.134866	0.071147	0.045048	0.055150	0.046905	0.049193	0.043576	k5
	0.95	0.127030	0.083929	0.044275	0.028034	0.034321	0.029190	0.030613	0.027118	k5
	0.97	0.167998	0.110996	0.058554	0.037075	0.045389	0.038603	0.040486	0.035863	k5
	0.99	0.222177	0.146792	0.077438	0.049031	0.060027	0.051053	0.053543	0.047429	k5
100	0.91	0.036259	0.023956	0.012638	0.008002	0.009796	0.008332	0.008738	0.007740	k5
	0.93	0.047953	0.031682	0.016713	0.010582	0.012956	0.011019	0.011556	0.010237	k5
	0.95	0.063417	0.041900	0.022104	0.013995	0.017134	0.014572	0.015283	0.013538	k5
	0.97	0.083869	0.055413	0.029232	0.018509	0.022660	0.019272	0.020212	0.017904	k5
	0.99	0.110917	0.073283	0.038659	0.024478	0.029967	0.025487	0.026730	0.023678	k5
250	0.91	0.015422	0.009196	0.003793	0.001351	0.003227	0.001407	0.002685	0.001185	k5
	0.93	0.020396	0.013475	0.007109	0.004501	0.005510	0.004687	0.004915	0.004354	k5
	0.95	0.026973	0.017821	0.009401	0.005953	0.007288	0.006198	0.006500	0.005758	k5
	0.97	0.035672	0.023569	0.012433	0.007872	0.009638	0.008197	0.008597	0.007615	k5
	0.99	0.047177	0.031170	0.016443	0.010411	0.012746	0.010840	0.011369	0.010071	k5

4. Conclusions:

Table 1 , 2 and Table 3 show the values of the mean square error MSE for all estimation methods studied at a number of variables (p=4,6,8). It can be noted that the values of the mean square error MSE decrease as the sample size increases, this reflects one of the good characteristics when the estimator approaches the true value of the parameter by increasing the sample size and holding two factors (degree of correlation, explanatory variables) constant. As for the effect of the other factor, which is the degree of correlation, we notice that increasing the degree of correlation between the explanatory variables leads to an increase in the average value. The mean square error MSE for all estimation methods, in addition to the application of the gamma regression method GRR and the combined estimator method RGRR is directly affected by the increase in the degree of correlation, as we notice that the value of the mean square error MSE increases with the increase in the degree of correlation, as shown by the simulation results, We also note that the combined estimator method RGRR with the formula k3 gave the lowest value for the mean square error MSE when the number of explanatory variables is (p=4)... whereas when the number of explanatory variables is (p=6.8), the formula is k5 It is the most appropriate because it gives the lowest value of the mean square error MSE.

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.

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

Reference:

1.Amine M., Qasim M., Yasin A. and Amanullah M. (2020) "Almost unbiased ridge estimator in the gamma regression model" Communications in Statistics – Simulation and Computations,51(7),3830-3850.

2.Amin A., Qasim M., Amanullah M. (2020) "Preformance of some ridge estimators for the gamma regression model "Statistical Papers, 61,997-1026.

3.Asar Y , Arashi M. , Wu J. (2017)" Restricted ridge estimator in the logestic regression model"Commun Stat Simul Comput ;46(8):6538-6544.

4.Al-Gamal ZY. (2018)"Developing a ridge estimator for the gamma regression model" J Chemometer;32(10),3054.

5.Al-Gamal ZY & Yasin A. (2020) "Liu-type estimator for the gamma regression model" Communications in Statistics - Simulation and Computation, 49(8), 2035-2048.

6.Adewale F., Dawoud I, Kibria B. M., Al-Gamal ZY, Aladeitan B.(2021) "A New Ridge-Type Estimator for the Gamma Regression Model", Scientifica, Vol. 2021, Article ID 5545356, 8 pages.

7.Al-Gamal ZY.(2018) "Shrinkage estimators for gamma regression model" Electronic Journal of Applied Statistical Analysis,11(1),253-268.

8.Abdeljabbar L.,(2021)"Comparison Between Maximum Likelihood and Bayesian methods for estimating the Gamma regression with practical application" Journal of Economics and Administrative Sciences,27(125),477-492.

9.Abdeljabbar Z.,Hussain S.(2020)"Comparison of the performance of some r-(k,d)class estimator with the (PCTP)estimator that used in estimating the general linear regression model in the presence of autocorrelation and multicollinearity problems at the same time "Journal of Economics and Administrative Sciences,26(121),397-414.

10. Harbe H.,Saad H.(2020)"Comparsion Branch and Bound Algorithm with Penalty function for solving Non-Linear Bi-level programming "Journal of Economics and Administrative Sciences,26(119),444-457.

11. Hardin and Hilbe, (2007) "Generalized linear models and extensions" stata press .

12. Kazem, A., Muslim B. (2002) "Advanced Economic Measurement Theory and Application" Donia Al-Amal Library, Al-Sina'ah Street, Baghdad,Iraq.

13. Kamal G.,Khazal S.(2019)"Comparison of some robust methods in the presence of problems of multicollinearity and high leverage points" Journal of Economics and Administrative Sciences,25(114),523-538.

14. Kurtoglu F., Ozkale MR. (2017)"Restricted ridge estimator in generalized linear models: Monte Carlo simulation studies on Poisson and binomial distributed response" Commun Sta tSimul Comput ;48(4),1191-1218.

15. Mahmoudi A., Arabi R., Mandal S. (2020)" A comparsion of perlimniary test,stein-type and penalty estimators in gamma regression model" J Stat Comput Simul ;90(17):3051-3079.

16. Qasim M. , Akram M., Amin M., Mansson K. (2021) "A restricted gamma ridge regression estimator combining the gamma ridge regression and the restricted maximum liklehood methods of estimation" Journal of Statistical Computation and Simulation,92(8),1696-1713.

17. Qasim M., Amin M., Amanullah M. (2018)"On the performance of some new Liu parameters for the gamma regression model" J stat Comput simul;88(16),3065-3080.

18. Qasim M., Kibria B., Mansson K. (2020)" A new Poisson Liu regression estimator: method and application" J Appl Stat;47(12),2258-2271.

19. R.Tibshirani (1996)"Regression shrinkage and selection via the lasso" Journal of the Royal Statistical Sociaty: Series B,58(1),267-288.
20. Sampreet Ch., Price B. (1989) "Regression Analysis with Examples" Higher Education presses in Mosul, Mousl,Iraq.
21. Sarkar N. (1992) "A new estimator combining the ridge regression and the restricted least squares methods of estimation"Commun Stat Theory Methods; 21(7),1987-2000.
22. Seifollahi S., Bevrani H., Kamary K. (2023) “Inequality Restricted Estimator for Gamma Regression: Bayesian approach as a solution to the Multicollinearity” Communications in Statistics - Theory and Methods, 2303-05120.
23. Yasin A. ,Amin M., Qasim M. , Muse A.,Soliman A. (2022) " More on the ridge parameter estimator for the Gamma ridge regression model ; Simulation and Application"Vol. 2022,Article ID 6769421.

Journal of Economics and Administrative Sciences
2024; 30(143), pp. 510-522
P-ISSN 2518-5764
E-ISSN 2227-703X

مقارنة بين طرائق تقدير إنموذج إنحدار جاما الحرف المقيد بإستخدام المحاكاة

ميفان شكور عبدالله
جامعة بغداد/ كلية الإدارة والاقتصاد / قسم الإحصاء
بغداد , العراق
Mivan.shakur2101m@coadec.uobaghdad.edu.iq
https://orcid.org/0009-0002-9567-4624

سهيل نجم عبود
جامعة بغداد/ كلية الإدارة والاقتصاد/ قسم الإحصاء
بغداد , العراق
suhnaj2005@coadec.uobaghdad.edu.iq

Received: 11/2/2024 Accepted: 17/3/2024 Published Online First: 1 /10/ 2024

مستخلص البحث:

نتناول في هذا البحث تقدير معالم إنموذج إنحدار الحرف جاما المقيد عن طريق دمج إنحدار جاما الحرف مع الإمكان الأعظم المقيد, وسيتم التعرف على خصائص المقدر الجديد وتفوقه على مقدر إنحدار جاما الحرف والإمكان الأعظم المقيد , بإستخدام عدة صيغ لمعامل الإنكماش k , إذ تم إستخدام أسلوب محاكاة مونت كارلو لتوليد بيانات تعاني من مشكلة التعدد الخطي بأحجام مختلفة (n=25,50,100,250) في ظل عوامل مؤثرة أخرى (درجة الإرتباط , عدد المتغيرات التفسيرية)، وإخضاع المعلمات الى قيد خطي , للتخلص من مشكلة التعدد الخطي في ظل خضوع معالم الإنموذج الى قيد خطي وسيتم تقدير معالم الإنموذج بإستخدام أربع طرق للتقدير , وتم الإعتماد على متوسط مربعات الخطأ MSE كمعيار للمقارنة بين طرائق التقدير ومن خلال نتائج تجربة المحاكاة تبين إن طريقة المقدر المدمج هي الطريقة الأفضل في تقدير معالم إنموذج إنحدار الحرف جاما المقيد.

نوع البحث : ورقة بحثية
المصطلحات الرئيسة للبحث: انحدار جاما الحرف المقيد (RGRR) , إنحدار جاما الحرف (GRR) ، الإمكان الأعظم المقيد (RMLE) ، معامل الإنكماش k ، متوسط مربعات الخطأ MSE .

- **البحث مستل من رسالة ماجستير**