Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v18n1p213

A new shrinkage estimator in negative binomial regression model

By Ismael, Algamal

15 March 2025

Electronic Journal of Applied Statistical Analysis
Vol. 18, Issue 01, march 2025, 213-224
DOI: 10.1285/i20705948v18n1p213

A new shrinkage estimator in negative binomial regression model

Hasnaa Ahmed Ismael[a] and Zakariya Yahya Algamal*[b]

[a] *University of Mosul, Mosul, Iraq.*
[b] *Department of Statistics and Informatics, University of Mosul, Mosul, Iraq.*

15 March 2025

Studies have proven that the ridge estimator proves itself as a highly desirable shrinkage tool for addressing multicollinearity issues. A widely used model called negative binomial regression model (NBRM functions efficiently when count data contains overdispersion properties. Maximum likelihood estimator (MLE) produces coefficients whose variance becomes affected negatively by multicollinearity issues. The proposed paper introduces the generalized ridge estimator to resolve the shortcomings of ridge estimator. Various approaches to estimate the shrinkage matrix have been developed. Monte Carlo simulation findings demonstrate that the proposed estimation technique produces superior MSE results than traditional MLE estimates and ridge estimates regardless of the selected shrinkage matrix estimation methodology. The estimating methods used for shrinkage matrices result in different levels of performance enhancement.

keywords: Negative binomial regression model; ridge estimator; multicollinearity; generalized ridge estimator; Monte Carlo simulation.

1 Introduction

The negative binomial regression model (NBRM) serves as a common statistical approach to study numerous real-world problems particularly focused on mortality studies that examine fatalities and health insurance problems that analyze individual claims Algamal (2012); Cameron and Trivedi (2013); De Jong (2008); Kandemir Çetinkaya and Kaçıranlar (2019); Lukman et al. (2021); Salih et al. (2024).

*Corresponding authors: E-mail: zakariya.algamal@uomosul.edu.iq

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

When using NBRM it is assumed that no connection exists between the explanatory variables. Actual situations reveal that this common misconception leads to the multi-collinearity issue. The maximum likelihood (ML) estimation approach produces unstable coefficients with high variance that result in low statistical significance when used to estimate NBRM regression coefficients during multicollinearity Kibria et al. (2015); Månsson (2012, 2013); Türkan and Özel (2018); Hoerl and Kennard (1970).

In classical linear regression models the following relationship is usually adopted

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations of the response variable, $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_p)$ is an $n \times p$ known design matrix of explanatory variables, $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)$ is a $p \times 1$ vector of unknown regression coefficients, "and $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of random errors with mean 0 and variance σ^2.

Ridge regression reduces the large variance by shrinking all regression coefficients toward zero Asar and Genç (2016); Batah et al. (2008a). This done by adding a positive amount to the diagonal of $\mathbf{X}^T\mathbf{X}$. As a result, the ridge estimator is biased, but it guaranties a smaller mean squared error than the ML estimator.

In linear regression, the ridge estimator is defined as

$$\hat{\boldsymbol{\beta}}_{Ridge} = (\mathbf{X}^T\mathbf{X} + q\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}, \tag{2}$$

where $\mathbf{I}$ is the identity matrix with dimension $p \times p$ and $q \geq 0$ represents the ridge parameter (shrinkage parameter). The ridge parameter, q, controls the shrinkage of $\boldsymbol{\beta}$ toward zero. For larger value of q, the $\hat{\boldsymbol{\beta}}_{Ridge}$ estimator yields greater shrinkage approaching zero Batah et al. (2008a).

2 Negative binomial regression model (NBRM)

The NBRM is very popular in applied research when the dependent variable y_i becomes non-negative integers or counts distributed as NB($\mu_i, \mu_i + \delta\mu_i^2$) where $\mu_i = \exp(x_i\beta)$ such that x_i is the i^{th} row of the data matrix X which is a n*(p+1) data matrix with p explanatory variables, β is the coefficient vector of order (p+1)*1 with intercept and z_i is a random variable following the gamma distribution such that $z_i \sim \Gamma\ (\delta,\delta)$, i=1, 2, 3, ... , n Hilbe (2011); Massaro and Bozdogan (2015). The density function of the dependent variable y_i is

$$pr\left(y_i \mid x_i\right) = \frac{\Gamma\left(\alpha^{-1} + y_i\right)}{\Gamma\left(\alpha^{-1}\right)\Gamma\left(1 + y_i\right)}\left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i}\right)^{\alpha^{-1}}\left(\frac{\mu_i}{\alpha^{-1} + \mu_i}\right)^{y_i} \tag{3}$$

where the over dispersion parameter α is define as α=1/δ. The conditional mean and variance are given as follows

$$E\left(y_i \mid x_i\right) = \mu_i \qquad , \qquad V\left(y_i \mid x_i\right) = \mu_i(1 + \alpha\mu_i) \tag{4}$$

This is the most commonly applied NBRM and the estimate of β is usually found by maximizing the log likelihood

$$l(\alpha,\beta)=\sum_{i=1}^{n}\left\{\sum_{j=0}^{y_i-1}\log\left(j+\alpha^{-1}\right)-\log\left(y_i!\right)\right.$$
$$\left.-\left(y_i-\alpha^{-1}\right)\log\left(1+\alpha\exp\left(x_i\beta\right)\right)+y_i\log\left(\alpha\right)+y_i\log\left(\exp\left(x_i\beta\right)\right)\right\} \quad (5)$$

Since $\ln\left[\frac{\Gamma(\alpha^{-1}+y_i)}{\Gamma(\alpha^{-1})}\right]=\sum_{i=1}^{n}\left(j+\alpha^{-1}\right)$. The vector of coefficients using maximum likelihood estimation by solving the equation

$$S(\beta)=\frac{\partial l(\mu,y)}{\partial\beta}=\sum_{i=1}^{n}\frac{(y_i-\mu_i)}{1+\alpha\mu_i}x_i=0 \quad (6)$$

Since the Eq. (6) is nonlinear in β the solution of S(β) equal to zero is found by using the method of scoring

$$\beta_r=\beta_{r-1}+I^{-1}\left(\beta_{r-1}\right)S(\beta_{r-1}) \quad (7)$$

where $S(\beta_{r-1})$ is the first derivative of the log-likelihood evaluated at β_{r-1} and

$$I^{-1}\left(\beta_{r-1}\right)=E\left(\frac{\partial^2 l\left(X,\beta\right)}{\partial\beta\partial\beta'}\right)=X^TWX \quad (8)$$

where $W=diag\left[\frac{(\mu_i(\beta_{r-1}))}{1+\alpha_{r-1}\mu_i(\beta_{r-1})}\right]$. The final part of Eq. (8) may be written as

$$X^TWX\beta_r=X^TWz(\beta_{r-1}) \quad (9)$$

By define $z(\beta_{r-1})$ as a vector where the ith value equals $\log\left(\mu_i\left(\beta_{r-1}\right)\right)+\frac{y_i-\mu_i(\beta_{r-1})}{\mu_i(\beta_{r-1})}$. This method is known as iteratively weighted least squares IWLS and in the final the maximum likelihood estimate of β denoted as β_{ML} is obtained (Rashad et al., 2021; Al-Taweel and Algamal, 2020; Algamal, 2025). The covariance matrix of this estimator given by

$$\mathrm{Cov}\left(\widehat{\beta}_{ML}\right)=(X^TWX)^{-1} \quad (10)$$

The MSE is given by

$$\begin{aligned}MSE\left(\widehat{\beta}_{ML}\right)&=E(\widehat{\beta}_{ML}-\beta)^T(\widehat{\beta}_{ML}-\beta)\\&=tr[(X^T\hat{W}X)^{-1}]\\&=\sum_{j=1}^{J}\frac{1}{\lambda_j(\widehat{\beta}_{ML})}\end{aligned} \quad (11)$$

where λ_j is the jth eigenvalue of the $(X^T\hat{W}X)$ matrix.

3 Generalized ridge estimator in NBRM

The weighted matrix of cross-products is used when there is a substantial correlation between the explanatory factors, $X^T\hat{W}X$ is ill-conditioned which leads to instability and high variance of the maximum likelihood estimator. To avoid this problem the Negative Binomial ridge regression proposed (NBRR) by (Månsson, 2012). By minimizes the weighted sum of squares error (WSSE). Hence $\widehat{\beta}_{ML}$ is given by:

$$\widehat{\beta}_{NBRR} = (X^{'}\hat{W}X + qI)^{-1} X^{'}\hat{W}X\ \widehat{\beta}_{NBML}$$
$$= (X^{'}\hat{W}X + qI\ )^{-1} X^{'}\hat{W}\hat{S} \tag{12}$$

In Månsson [6] it is shown that the MSE of this estimator equals:

$$MSE(\widehat{\beta}_{NBRR}) = \sum_{j=1}^{J} \frac{\lambda_j}{(\lambda_j + q)^2} + q^2 \sum_{j=1}^{J} \frac{\alpha_j{}^2}{(\lambda_j + q)^2} = \gamma_1\,(q) + \gamma_2(q) \tag{13}$$

where $\gamma_1\,(q)$ is the variance and $\gamma_2\,(q)$ is the bias part of $\widehat{\beta}_{NBRR}$

The MSE of $\widehat{\beta}_{NBRR}$ is lower than $\hat{\beta}_{ML}$ estimate such that when we found q (where $q$may take on value between zero and infinity) such that the reduction in the variance part is greater than the increase of the squared part, for this reason NBRR estimation is better than ML, furthermore NBRR is simple method since it dose not require any changes of the negative binomial regression.

Researchers recommend the generalized ridge estimator (GRE) as an extension of ridge estimator which distinguishes itself from the regular ridge estimator is there are p values of q, such that Hoerl and Kennard (1970)

$$\hat{\boldsymbol{\beta}}_{GRE} = (\mathbf{X}^T\mathbf{X} + \mathbf{Q})^{-1}\mathbf{X}^T\mathbf{y}, \tag{14}$$

where $\mathbf{Q} = \mathrm{d}iag(q_1, q_2, ..., q_p)$. The good thing where using GRE is find the best values of $q_i$so as to obtain the MSE which is less than when we using the ridge estimator and OLS.

The generalized ridge estimator for NBRM (GRNBRM) is defined as

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{GRNBRM} &= (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + \mathbf{Q})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}\hat{\boldsymbol{\beta}}_{ML} \\ &= (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + \mathbf{Q})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\hat{\mathbf{u}}.\end{aligned} \tag{15}$$

The selection of the matrix $\mathbf{Q}$ is very important. In this paper, several methods are adapted to estimate $\mathbf{Q}$, such as Hocking et al. (1976), Nomura (1988), Troskie and Chalton (1996), Firinguetti (1999), Alkhamisi and Shukur (2007), Batah et al. (2008b), Al-Hassan (2010), Dorugade and Kashid (2010), Månsson et al. (2010), Dorugade (2014), Asar et al. (2014) and Bhat and Raju (2017). These methods are given below, respectively

$$\hat{q}_{i(HK)} = \frac{1}{\alpha_i^2}, \tag{16}$$

where α_j is defined as the jth element of $\gamma\hat{\boldsymbol{\beta}}_{ML}$and γ is the eigenvector of the$\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$.

$$\hat{q}_{i(N)} = \frac{1}{\hat{\alpha}_i^2}\left\{1+\left[1+\lambda_i(\hat{\alpha}_i^2)^{1/2}\right]\right\} [15] \tag{17}$$

$$\hat{q}_{i(TC)} = \frac{\lambda_i}{\lambda_i\hat{\alpha}_i^2} \tag{18}$$

$$\hat{q}_{i(F)} = \frac{\lambda_i}{\lambda_i\hat{\alpha}_i^2+(n-p)} \tag{19}$$

$$\hat{q}_{i(HSL)} = \frac{\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2)^2}{(\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2))^2} \tag{20}$$

$$\hat{q}_{i(AH)} = \frac{\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2)^2}{(\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2))^2}+\frac{1}{\lambda_{\max}} \tag{21}$$

$$\hat{q}_{i(D)} = \frac{1}{\lambda_{\max}\hat{\alpha}_i^2} \tag{22}$$

$$\hat{q}_{i(SB)} = \frac{\lambda_i}{\lambda_i\hat{\alpha}_i^2}+\frac{1}{\lambda_{\max}} \tag{23}$$

$$\hat{q}_{i(SV1)} = \frac{p}{\hat{\alpha}_i^2}+\frac{1}{\lambda_{Max}\hat{\alpha}_i^2} \tag{24}$$

$$\hat{q}_{i(SV2)} = \frac{p}{\hat{\alpha}_i^2}+\frac{1}{2\left(\sqrt{\lambda_{Max}/\lambda_{Min}}\right)^2} \tag{25}$$

$$\hat{q}_{i(M)} = \frac{1}{\frac{\lambda_{Max}\hat{\alpha}_i^2}{(n-p)+\lambda_{Max}\hat{\alpha}_i^2}} \tag{26}$$

$$\hat{q}_{i(AS)} = \frac{1}{\hat{\alpha}_i^2}+\frac{1}{\lambda_i} \tag{27}$$

4 Simulation results

This section demonstrates how Monte Carlo simulation examines the new estimator with varying levels of multicollinearity through experimental tests. The response variable of observations is generated from negative binomial regression as $NB(\mu_i, \mu_i + \theta\mu_i^2)$ with $\mu_i = \exp(x_i^T\beta)$ (Algamal, 2020; Algamal and Abonazel, 2022; Abonazel et al., 2022). Here, $\beta = (\beta_0, \beta_1, ..., \beta_p)$ with and Månsson and Shukur (2011). n= 30, 50 and 150 and p=3 and 7. The generated data is repeated 1000 times and the averaged mean squared errors (MSE) is calculated as

$$MSE(\hat{\beta}) = \frac{1}{1000}\sum_{i=1}^{1000}(\hat{\beta}-\beta)^T(\hat{\beta}-\beta), \tag{28}$$

where $\hat{\beta}$ is the estimated coefficients for the used estimator.

The averaged MSE all the combination of n, p, and ρ, are respectively summarized in Tables 1 – 6. The best value of the averaged MSE is highlighted in bold. The following are some observations that can be made:

1. The MSE of GRNBRM is often lower than that of MLE.
2. Regardless of the kind of **K** matrix estimation technique used, it is clear that GRNBRM achieved a lower MSE than NBRR.
3. It is clear that there is a negative effect on MSE with respect to the number of explanatory variables, as their values increase as p increases.
4. The MSE performance of SV2 method exceeded alternative methods when applied to the NBRM on all experimental conditions. All testing conditions demonstrated that HK as well as SB methods yielded inferior performance compared to alternative methods used.
5. When the correlation degree grows, independent of the values of n and p, the MSE values increase in terms of ρ values, demonstrating the superiority of the SV2 approach.

Table 1: Average MSE values when $n = 30$ and $p = 3$

Methods			
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.6741	5.8372	6.0518
NBRR	4.2001	4.2214	4.1991
F	2.9714	3.0257	3.0403
HSL	3.3838	3.4381	3.4527
SV1	3.3424	3.3967	3.4113
SV2	2.7977	2.852	2.8666
D	3.2043	3.2151	3.2207
AH	3.309	3.3633	3.3779
HK	3.7548	3.8091	3.8237
N	3.4182	3.4264	3.4288
TC	3.4968	3.5511	3.5657
SB	3.5015	3.5558	3.5704
M	3.3065	3.3603	3.3749
AS	3.3674	3.4217	3.4363

Table 2: Average MSE values when $n = 30$ and $p = 7$

Methods			
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.8841	6.0472	6.2618
NBRR	4.4101	4.4314	4.4091
F	3.1814	3.2357	3.2503
HSL	3.5938	3.6481	3.6627
SV1	3.5524	3.6067	3.6213
SV2	2.9977	3.052	3.0666
D	3.4143	3.4251	3.4307
AH	3.519	3.5733	3.5879
HK	3.9648	4.0191	4.0337
N	3.6282	3.6364	3.6388
TC	3.7068	3.7611	3.7757
SB	3.7115	3.7658	3.7804
M	3.5165	3.5703	3.5849
AS	3.5774	3.6317	3.6463

Table 3: Average MSE values when $n = 50$ and $p = 3$

Methods			
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.7831	5.9462	6.1608
NBRR	4.3091	4.3304	4.3081
F	3.0804	3.1347	3.1493
HSL	3.4928	3.5471	3.5617
SV1	3.4514	3.5057	3.5203
SV2	2.8967	2.951	2.9656
D	3.3133	3.3241	3.3297
AH	3.418	3.4723	3.4869
HK	3.8638	3.9181	3.9327
N	3.5272	3.5354	3.5378
TC	3.6058	3.6601	3.6747
SB	3.6105	3.6648	3.6794
M	3.4155	3.4693	3.4839
AS	3.4764	3.5307	3.5453

Table 4: Average MSE values when $n = 50$ and $p = 7$

Methods			
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.7501	5.9462	6.1608
NBRR	4.2761	4.3304	4.3081
F	3.0474	3.1347	3.1493
HSL	3.4598	3.5471	3.5617
SV1	3.4184	3.5057	3.5203
SV2	2.8637	2.951	2.9656
D	3.2803	3.3241	3.3297
AH	3.385	3.4723	3.4869
HK	3.8308	3.9181	3.9327
N	3.4942	3.5354	3.5378
TC	3.5728	3.6601	3.6747
SB	3.5775	3.6648	3.6794
M	3.3825	3.4693	3.4839
AS	3.4434	3.5307	3.5453

Table 5: Average MSE values when $n = 150$ and $p = 3$

Methods			
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.7961	5.9592	6.1738
NBRR	4.3221	4.3434	4.3211
F	3.0934	3.1477	3.1623
HSL	3.5058	3.5601	3.5747
SV1	3.4644	3.5187	3.5333
SV2	2.9097	2.964	2.9786
D	3.3263	3.3371	3.3427
AH	3.431	3.4853	3.4999
HK	3.8768	3.9311	3.9457
N	3.5402	3.5484	3.5508
TC	3.6188	3.6731	3.6877
SB	3.6235	3.6778	3.6924
M	3.4285	3.4823	3.4969
AS	3.4894	3.5437	3.5583

Table 6: Average MSE values when $n = 150$ and $p = 7$

Methods			
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.7811	5.9442	6.1588
NBRR	4.3071	4.3284	4.3061
F	3.0784	3.1327	3.1473
HSL	3.4908	3.5451	3.5597
SV1	3.4494	3.5037	3.5183
SV2	2.8947	2.949	2.9636
D	3.3113	3.3221	3.3277
AH	3.416	3.4703	3.4849
HK	3.8618	3.9161	3.9307
N	3.5252	3.5334	3.5358
TC	3.6038	3.6581	3.6727
SB	3.6085	3.6628	3.6774
M	3.4135	3.4673	3.4819
AS	3.4744	3.5287	3.5433

5 Real data application

We use the suggested estimator to the Spanish La Liga football season of 2016–2017 in order to further examine the utility of our new estimator. There are 20 teams in this data. The number of matches won is represented by the response variable. The six considerable explanatory variables included the number of yellow cards (x1), the number of red cards (x2), the total number of substitutions (x3), the number of matches with 2.5 goals on average (x4), the number of matches that ended with goals (x5), and the ratio of the goal scores to the number of matches (x6)". As stated in (Alobaidi et al., 2021), the data follows NBRM and there is high colinearity.

The estimated MSE values for the used estimators are listed in Table 7. According to Table 7, it is clearly seen that the SV2 estimator shrinks the value of the estimated coefficients efficiently.

6 Conclusions

NBRM serves as a statistical modeling approach which handles count data that shows overdispersion by having a variance exceeding its mean. The MLE computations produce unreliable coefficient estimates when multicollinearity occurs because they become characteristically uncertain. The generalized ridge estimator was presented in this study as a solution to the NBRM multicollinearity issue. A number of matrix estimation techniques

Table 7: MSE values of the real data application

Methods	MSE
MLE	10.7091
NBRR	9.2351
F	8.0064
HSL	8.4188
SV1	8.3774
SV2	7.8327
D	8.2393
AH	8.344
HK	8.7898
N	8.4532
TC	8.5318
SB	8.5365
M	8.3415
AS	8.4024

have been modified. Monte Carlo simulation experiments show that the GRNBRM estimator performs better than MLE and NBRR in terms of MSE, independent of the type of estimating method used for the **Q** matrix.

References

Abonazel, M. R., Algamal, Z. Y., Awwad, F. A., and Taha, I. M. (2022). A new two-parameter estimator for beta regression model: method, simulation, and application. *Frontiers in Applied Mathematics and Statistics*, 7:780322.

Al-Hassan, Y. M. (2010). Performance of a new ridge regression estimator. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 9(1):23–26.

Al-Taweel, Y. and Algamal, Z. (2020). Some almost unbiased ridge regression estimators for the zero-inflated negative binomial regression model. *Periodicals of Engineering and Natural Sciences (PEN)*, 8(1):248–255.

Algamal, Z. Y. (2012). Diagnostic in poisson regression models. *Electronic Journal of Applied Statistical Analysis*, 5(2):178–186.

Algamal, Z. Y. (2020). Shrinkage parameter selection via modified cross-validation approach for ridge regression model. *Communications in Statistics-Simulation and Computation*, 49(7):1922–1930.

Algamal, Z. Y. (2025). A modified ridge estimator in cox regression model. *Communications in Statistics-Simulation and Computation*, pages 1–8.

Algamal, Z. Y. and Abonazel, M. R. (2022). Developing a liu-type estimator in beta regression model. *Concurrency and Computation: Practice and Experience*, 34(5):e6685.

Alkhamisi, M. A. and Shukur, G. (2007). A monte carlo study of recent ridge parameters. *Communications in Statistics—Simulation and Computation®*, 36(3):535–547.

Alobaidi, N. N., Shamany, R. E., and Algamal, Z. Y. (2021). A new ridge estimator for the negative binomial regression model. *Thailand Statistician*, 19(1):115–124.

Asar, Y. and Genç, A. (2016). New shrinkage parameters for the liu-type logistic estimators. *Communications in Statistics-Simulation and Computation*, 45(3):1094–1103.

Asar, Y., Karaibrahimoğlu, A., and Genç, A. (2014). Modified ridge regression parameters: A comparative monte carlo study. *Hacettepe Journal of Mathematics and Statistics*, 43(5):827–841.

Batah, F. S. M., Ramanathan, T. V., and Gore, S. D. (2008a). The efficiency of modified jackknife and ridge type regression estimators: a comparison. *Surveys in Mathematics and its Applications*, 3:111–122.

Batah, F. S. M., Ramanathan, T. V., and Gore, S. D. (2008b). The efficiency of modified jackknife and ridge type regression estimators: a comparison. *Surveys in Mathematics and its Applications*, 3:111–122.

Bhat, S. and Raju, V. (2017). A class of generalized ridge estimators. *Communications in Statistics-Simulation and Computation*, 46(7):5105–5112.

Cameron, A. C. and Trivedi, P. K. (2013). *Regression analysis of count data*. Number 53. Cambridge university press.

De Jong, P. (2008). *Generalized linear models for insurance data*, volume 81. Cambridge University Press.

Dorugade, A. and Kashid, D. (2010). Alternative method for choosing ridge parameter for regression. *Applied Mathematical Sciences*, 4(9):447–456.

Dorugade, A. V. (2014). New ridge parameters for ridge regression. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 15:94–99.

Firinguetti, L. (1999). A generalized ridge regression estimator and its finite sample properties: A generalized ridge regression estimator. *Communications in Statistics-Theory and Methods*, 28(5):1217–1229.

Hilbe, J. (2011). *Negative binomial regression*. Cambridge University Press.

Hocking, R. R., Speed, F., and Lynn, M. (1976). A class of biased estimators in linear regression. *Technometrics*, 18(4):425–437.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67.

Kandemir Çetinkaya, M. and Kaçıranlar, S. (2019). Improved two-parameter estimators for the negative binomial and poisson regression models. *Journal of Statistical Computation and Simulation*, 89(14):2645–2660.

Kibria, B. G., Månsson, K., and Shukur, G. (2015). A simulation study of some biasing parameters for the ridge type estimation of poisson regression. *Communications in Statistics-Simulation and Computation*, 44(4):943–957.

Lukman, A. F., Dawoud, I., Kibria, B. G., Algamal, Z. Y., and Aladeitan, B. (2021). A new ridge-type estimator for the gamma regression model. *Scientifica*, 2021(1):5545356.

Månsson, K. (2012). On ridge estimators for the negative binomial regression model. *Economic Modelling*, 29(2):178–184.

Månsson, K. (2013). Developing a liu estimator for the negative binomial regression model: method and application. *Journal of Statistical Computation and Simulation*, 83(9):1773–1780.

Månsson, K. and Shukur, G. (2011). A poisson ridge regression estimator. *Economic Modelling*, 28(4):1475–1481.

Månsson, K., Shukur, G., and Golam Kibria, B. (2010). A simulation study of some ridge regression estimators under different distributional assumptions. *Communications in Statistics-Simulation and Computation*, 39(8):1639–1670.

Massaro, T. J. and Bozdogan, H. (2015). Variable subset selection via ga and information complexity in mixtures of poisson and negative binomial regression models. *arXiv preprint arXiv:1505.05229*.

Nomura, M. (1988). On the almost unbiased ridge regression estimator. *Communications in Statistics-Simulation and Computation*, 17(3):729–743.

Rashad, N. K., Hammood, N. M., and Algamal, Z. Y. (2021). Generalized ridge estimator in negative binomial regression model. In *Journal of Physics: Conference Series*, volume 1897, page 012019. IOP Publishing.

Salih, A. M., Algamal, Z., and Khaleel, M. A. (2024). A new ridge-type estimator for the gamma regression model. *Iraqi Journal for Computer Science and Mathematics*, 5(1):85–98.

Troskie, C. G. and Chalton, D. . (1996). *Detection of outliers in the presence of multicollinearity*, pages 273–292. De Gruyter, Berlin, Boston.

Türkan, S. and Özel, G. (2018). A jackknifed estimators for the negative binomial regression model. *Communications in Statistics-Simulation and Computation*, 47(6):1845–1865.