http://communications.science.ankara.edu.tr

Commun.Fac.Sci.Univ.Ank.Ser. A1 Math. Stat.
Volume 73, Number 3, Pages 712–723 (2024)
DOI:10.31801/cfsuasmas.1397263
ISSN 1303-5991 E-ISSN 2618-6470

Research Article; Received: November 28, 2023; Accepted: April 17, 2024

TWO PARAMETER RIDGE ESTIMATOR FOR THE BELL REGRESSION MODEL

Melike IŞILAR[1] and Y. Murat BULUT[2]

[1]Graduate School of Natural and Applied Sciences, Eskişehir Osmangazi University, Eskişehir, TÜRKİYE
[2]Department of Statistics, Eskişehir Osmangazi University, Eskişehir, TÜRKİYE

Abstract. One solution to the multicollinearity problem in the Bell regression model, which is utilized for over-dispersion issues, is biased estimators. In recent years, some biased estimators have been proposed in the Bell regression model that can be used in modelling correlated count data. In this article, Bell two-parameter ridge estimator (BTPRE) is proposed. This two-parameter estimator has some advantages over the previously proposed estimators. More efficient results are obtained than the Maximum Likelihood estimator (MLE) and Bell Ridge estimator (BRE) in the case of multicollinearity by using BTPRE. Monte Carlo simulation study and real data results are obtained to show that the proposed estimator is better. Estimators have been compared according to the Mean Squared Error (MSE) criterion. BTPRE is superior to other estimators.

1. Introduction

In count data modelling, the key distribution is the Poisson distribution because of its simplicity. It has only one parameter, the location parameter, to be estimated. However, the main drawback of the Poisson distribution is that the mean and variance of the Poisson distribution are equal, which is called equidispersion. But, in many real datasets, this assumption does not hold since the variance is greater than the mean of the data. This situation is called an overdispersion problem. When the variability of the data is greater than the mean, an overdispersion problem arises. The most popular overdispersed model is the Negative Binomial regression model (NBRM). NBRM is a mixture model which obtains a mixture of Poisson and Gamma distributions.

2020 *Mathematics Subject Classification.* 62J07, 62J12.
Keywords. Regression, count data, overdispersion, multicollinearity.
melikeisilar@gmail.com; 0000-0001-6821-1064
ymbulut@ogu.edu.tr-Corresponding author; 0000-0002-0545-7339.

The advantage of Poisson regression model (PRM) over NBRM is that it has one parameter. In response, the advantage of NBRM is that it can be used to model overdispersed data. As an alternative to this model, the Bell regression model (BRM), which has a single parameter, has been proposed by Castellars *et al.* for modelling overdispersed count data [8]. BRM has the advantages of both PRM and NBRM; it has been widely preferred recently. As compared with the NBRM, BRM is more flexible than the NBRM.

One of the general assumptions of regression analysis is that the independent variables are not collinear. But often, in real-life datasets, the independent variables are correlated. This problem is called multicollinearity. If the assumptions are met, a maximum likelihood estimator (MLE) efficiently estimates the parameter. Highly correlated independent variables affect the performance of MLE. In the case of multicollinearity problems, the variance of MLE increases, and the confidence intervals widen. There are many studies on biased estimators to solve this problem. The variance of MLE, which is an unbiased estimator, is very high in the case of multicollinearity problems. In this case, alternative estimators with a bias value and a smaller variance than the variance of the MLE can be used. Thus, the MSE of the biased estimators is smaller than that of the MLE.

One of the most widely used biased estimators is the Ridge estimator (RE) proposed by Hoerl and Kennard [9]. This estimator depends on the k biased parameter. As with many biased estimators, RE was first proposed in a linear regression model (LRM). There are many studies on RE in the literature regarding both its definitions in different regression models and the estimation of the biased parameter. The logistic ridge estimator was defined by Schaefer *et al.*, the Gamma ridge estimator was defined by Algamal, and the inverse Gaussian ridge estimator was defined by Algamal and its performances were examined [2, 3, 20]. Regarding modeling of counting data, RE studies were carried out by Månsson and Shukur, Månsoon and Amin *et al.* for PRM, NBRM and BRM, respectively [4,12,13].There are alternative estimators to the RE in the literature. Many of these estimators have also been identified in modeling count data [1, 17, 18].

In the ordinary least squares estimator (OLSE), there is an orthogonality between the residuals and the dependent variable. The orthogonality of this estimator is not available in the RE. In the RE, the aim is to reduce the variance, and model fit is not considered. A two-parameter ridge estimator (TPRE) was proposed by Lipovetsky and Conklin [10, 11] as a generalized version of the ridge estimator to increase the regression fit. The TPRE consists of k and q parameters. With the added parameter q, orthogonality between the dependent variable and residuals is provided. In addition, more efficient estimates are obtained from MLE and RE estimators. The TPRE for the linear model was compared with the OLSE and RE by Toker and Kaçıranlar according to the matrix MSE criterion [21]. Asar and Genç proposed TPRE for the logistic regression model [5]. Then, TPRE was defined for the inverse Gaussian regression model by Bulut and Işılar [7].

In this study, we propose the BTPRE for the Bell regression model used in modelling the count data. For this purpose, BRM and BRE are given in Section 2. BTPRE has been defined. In Section 3, the Monte Carlo Simulation study and actual data results are given to examine the performance of the proposed estimator. In the Section 4, the results of the studies are examined.

2. Methodology

A discrete random variable Y is said to be Bell distribution with the parameter $\theta > 0$, $Y \sim Bell(\theta)$, if its probability mass function (pmf) is given as

$$P(Y = y) = \frac{\theta^y e^{-e^\theta + 1} B_y}{y!}, \; y = 0, 1, 2, \ldots \tag{1}$$

where $B_y = \frac{1}{n}\sum_{q=0}^{\infty} \frac{q^y}{q!}$ is called the Bell number [6]. Since the Bell distribution is a member of the exponential family, the Bell regression model can be written as a special case of the generalized linear models (GLM's), which are widespread to model the mean of the response variable. Using the reparametrization given by Castellares *et al*. [8], the pmf can be rewritten as follows:

$$P(Y = y) = exp\Big\{1 - exp\big\{W_0(\mu)\big\}\Big\} \frac{W_0(\mu)^y B_y}{y!}, \; y = 0, 1, 2, \ldots \tag{2}$$

where $\theta = W_0(\mu)$ and $W_0(.)$ is the Lambert function. The mean and variance can be written using this parametrization as follows

$$E(y) = \mu, \tag{3}$$
$$Var(y) = \mu[1 + W_0(\mu)]. \tag{4}$$

The BRM is a good alternative to NBRM to model count data with overdispersion. The response variable distributed as $y_i \sim Bell(W_0(\mu_i))$ where $\mu_i = exp\{x_i^T\beta\}exp\{exp\{x_i^T\beta\}\}$ for $i = 1, 2, \ldots, n$. Using the Eq. (2), the log-likelihood function is given as follows

$$\begin{aligned} \ell(\mu_i; y_i) &= n - \sum_{i=1}^{n} exp\big\{W_0(\mu_i)\big\} + \sum_{i=1}^{n} y_i log\big(W_0(\mu_i)\big) + \sum_{i=1}^{n} log\big(B_{y_i}\big) - \sum_{i=1}^{n} log\big(y_i!\big) \\ &\propto \sum_{i=1}^{n} y_i log\big(exp\{x_i^T\beta\}exp\{exp\{c_i^T\beta\}\}\big) - exp\big\{exp\{x_i^T\beta\}exp\{exp\{x_i^T\beta\}\}\big\} \end{aligned} \tag{5}$$

Taking the derivative of the log-likelihood function concerning β parameter, we can obtain the following score function

$$S(\beta) = \frac{d\ell(\mu_i; y_i)}{d\beta} = \sum_{i=1}^{n} \left[x_i \Big(1 + exp\big\{x_i^T\beta\big\} \Big)(y_i - \mu_i) \right] \tag{6}$$

The most commonly used estimation method in the GLM's is the maximum likelihood estimation (MLE) method. To obtain the MLE of the BRM, we have to solve Eq. (6). Since the Eq. (6) is a non-linear according to the β, we can use the method of scoring:

$$\beta^{(m+1)} = \beta^{(m)} + I^{-1}\beta^{(m)} S\Big(\beta^{(r)}\Big) \tag{7}$$

where $S\big(\beta^{(m)}\big)$ is the score function evaluated at $\beta^{(m)}$, and

$$I^{-1}\big(\beta^{(m)}\big) = E\left[\frac{d^2\ell(\mu_i; y_i)}{d\beta d\beta^T}\right] = X^T W\big(\beta^{(m)}\big)X,$$

where $W\big(\beta^{(m)}\big) = diag\left\{\frac{\mu_i(\beta^{(m)})}{1+exp\left\{x_i^T\beta^{(m)}\right\}}\right\}$ evaluated at $\beta^{(m)}$. The final step of the Eq. (7) can also be written as

$$\widehat{\beta}_{MLE} = (X^T\widehat{W}X)^{-1}X'\widehat{W}\widehat{z}, \tag{8}$$

where $\widehat{z} = log(\widehat{\mu}) + W^{-\frac{1}{2}}V^{-\frac{1}{2}}(y-\mu)$, and $V = Var(y)$. The covariance matrix of the MLE can be computed as

$$Cov\big(\widehat{\beta}_{MLE}\big) = \big(X^T\widehat{W}X\big)^{-1}, \tag{9}$$

which equals the inverse of the Hessian matrix. The matrix mean square error (MMSE) and scaler mean square error (SMSE) of the MLE are given by

$$MMSE\big(\widehat{\beta}_{MLE}\big) = D^{-1}, \tag{10}$$

$$SMSE\big(\widehat{\beta}_{MLE}\big) = \sum_{j=1}^{l}\frac{1}{\lambda_j}, \tag{11}$$

where $D = X^T\widehat{W}X$, λ_j are the eigenvalues of D matrix and l is a total number of parameter.

When the multicollinearity exits, the MLE inflates. So, Amin *et al*. [4] proposed the Ridge estimator for the BRM to handle the multicollinearity problem as given in the following subsection.

2.1. **Ridge Estimator in the BRM.** Amin *et al*. [4] introduced the Bell Ridge estimator (BRE) to cope with the multicollinearity problem's adverse effects. BRE is given as follows

$$\widehat{\alpha}_k = D_k^{-1}D\alpha \tag{12}$$

where $D_k = (X^T\widehat{W}X + kI_l)$ and $k > 0$ is a biasing parameter. $\alpha = Z^T\beta_{MLE}$ where Z is a eigenvector of D. The MMSE and SMSE of the BRE are given as

$$MMSE\Big(\widehat{\alpha}_k\Big) = D_k D D_k + k^2 D_k^{-1}\alpha\alpha^T D_k^{-1}, \quad (13)$$

$$SMSE\Big(\widehat{\alpha}_k\Big) = \sum_{j=1}^{l}\frac{\lambda_j}{(\lambda_j+k)^2} + k^2\sum_{j=1}^{l}\frac{\alpha_j^2}{(\lambda_j+k)^2}. \quad (14)$$

In this study, the biased parameter has estimated as follows

$$\widehat{k} = \frac{l}{\widehat{\alpha}^T\widehat{\alpha}}. \quad (15)$$

2.2. **Two Parameter Ridge Estimator in BRM.** Lipovetsky and Conklin [11] has proposed an objective function for the TPRE as follows

$$S^2 = ||Y - X\beta||^2 + q_1||\beta||^2 + q_2||X^TY - \beta||^2 + q_3||Y^T(Y - X\beta)||^2. \quad (16)$$

The generalization of the TPRE in the BRM obtained from the objective function given in Eq. (16) is given below

$$\widehat{\alpha}_{qk} = qD_k^{-1}D\widehat{\alpha} \quad (17)$$

where $k > 0$ and $q > 0$. This estimator is the Bell two-parameter Ridge estimator (BTPRE) in which BRE and MLE are special cases of it. For example, if $q = 1$ is taken in Eq. (17), we can obtain $\widehat{\alpha}_k$. If we takes $q = 1$ and $k = 0$, $\widehat{\alpha}_{MLE}$ can be obtained. The coefficient of determination for the BTPRE is given in Eq. (18).

$$R^2 = 2qr^TD_k^{-1}r - q^2r^TD_k^{-1}DD_k^{-1}r \quad (18)$$

where $r = X^T\widehat{W}\widehat{z}$. In order to maximize the model fit, optimal q is as follow

$$q = \frac{r^TD_k^{-1}r}{r^TD_k^{-1}DD_k^{-1}r}. \quad (19)$$

MMSE and MSE are computed as

$$MMSE(\widehat{\alpha}_{qk}) = q^2D_k^{-1}DD_k^{-1} + (qD_k^{-1}D - I)\alpha\alpha^T(qD_k^{-1}D - I), \quad (20)$$

$$MSE(\widehat{\alpha}_{qk}) = q^2\sum_{j=1}^{l}\frac{\lambda_j}{(\lambda_j+k)^2} + \sum_{j=1}^{l}\frac{\alpha_j^2(q\lambda_j - \lambda_j - k)^2}{(\lambda_j+k)^2}. \quad (21)$$

In the literature related to biased estimators, there are different estimation equations for the parameters of the estimators. In order to minimize the MSE in the BTPRE, the derivatives of Eq. (12) for k and q, respectively, were calculated. The optimal parameter estimates obtained by equating the equations to zero are given below.

$$k = \frac{\sum_{j=1}^{l} q\lambda_j + (q-1)\lambda_j^2\alpha_j^2}{\sum_{j=1}^{l}\lambda_j\alpha_j^2} \quad (22)$$

$$q = \frac{\sum_{j=1}^{l} \frac{\lambda_j \alpha_j^2}{\lambda_j + k}}{\sum_{j=1}^{l} \frac{\lambda_j + \lambda_j \alpha_j^2}{(\lambda_j + k)^2}} \tag{23}$$

Two methods were used for the proposed BTPRE in this study. First, the parameters for the BTPRE symbolized as $\widehat{\alpha}_{qk_1}$ were calculated by following the steps below.

Step 1. The initial value is determined so that $\widehat{k}^0 > \frac{1}{\widehat{\alpha}^T \widehat{\alpha}}$.
Step 2. Calculate $\widehat{q}$ using Eq. (23) with $\widehat{k}^0$ given in Step 1.
Step 3. It is calculated as $\widehat{k} = \frac{1}{l} \sum_{j=1}^{l} \frac{\widehat{q}\lambda_j + (\widehat{q}-1)\lambda_j^2 \widehat{\alpha}_j^2}{\lambda_j \widehat{\alpha}_j^2}$.

Secondly, the TPRE calculated with the following steps is given as $\widehat{\alpha}_{qk_2}$.

Step 1. Calculate the initial value as $q^0 > \sum_{j=1}^{l} \frac{\lambda_j \widehat{\alpha}_j^2}{1 + \lambda_j \widehat{\alpha}_j^2}$.
Step 2. Eq. (22) using q^0 yields k^0.
Step 3. $\widehat{q}$ is calculated from Eq. (19).
Step 4. Using Eq. (23), $\widehat{q}$ is updated.

Theorem 1. *Let $k > 0$, BTPRE is superior to MLE if $k > \lambda_j(q-1)$ where $j = 1, ..., l$.*

Proof. The difference between MSE's of the MLE and BTPRE is obtained by

$$\begin{aligned} \boldsymbol{\delta} &= MSE(\widehat{\alpha}) - MSE(\widehat{\alpha}_{qk}) \\ &= \sum_{j=1}^{l} \frac{1}{\lambda_j} - q^2 \sum_{j=1}^{l} \frac{\lambda_j}{(\lambda_j + k)^2} - \sum_{j=1}^{l} \frac{(q\lambda_j - \lambda_j - k)^2 \alpha_j^2}{(\lambda_j + k)^2}. \end{aligned} \tag{24}$$

The difference between MSE's is pozitif definite, if $\frac{1}{\lambda_j} - \frac{\lambda_j}{(\lambda_j + k)^2}$ is pozitif. The fact that $\boldsymbol{\delta}$ is a p.d. iff $k > \lambda_j(q-1)$. The proof is finished. □

Theorem 2. *Let $k > 0$, $MSE(\widehat{\alpha}_k) - MSE(\widehat{\alpha}_{qk}) > 0$, if only $q > 1$.*

Proof. The difference between MSE's of the BRE and BTPRE is obtained by

$$\begin{aligned} \boldsymbol{\delta} &= MSE(\widehat{\alpha}_k) - MSE(\widehat{\alpha}_{qk}) \\ &= \sum_{j=1}^{l} \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \sum_{j=1}^{l} \frac{\alpha_j^2}{(\lambda_j + k)^2} - q^2 \sum_{j=1}^{l} \frac{\lambda_j}{(\lambda_j + k)^2} - \sum_{j=1}^{l} \frac{(q\lambda_j - \lambda_j - k)^2 \alpha_j^2}{(\lambda_j + k)^2} \\ &= (1 - q^2) \sum_{j=1}^{l} \frac{\lambda_j}{(\lambda_j + k)^2} + \sum_{j=1}^{l} \frac{[k^2 - (q\lambda_j - \lambda_j - k)^2]\alpha_j^2}{(\lambda_j + k)^2}. \end{aligned} \tag{25}$$

For the δ to be positive, the difference between variances must be positive. If only $q < 1$ then $(1 - q^2) > 0$. The proof is completed. □

3. Simulation Study and Real Data Example

The performances of the estimators are compared according to the MSE criterion in the simulation and a real data example.

3.1. **Simulation Study.** The $X_{n\times p}$ independent variable matrix formed in the studies on biased estimators was created by McDonald and Galarneau [14] using the equation given in Eq. (26).

$$x_{ij} = (1-\rho^2)^{1/2} z_{ij} + \rho z_{ip}, \tag{26}$$

where ρ is the correlation coefficient. The z_{ij} are pseudo random numbers. $y_{n\times 1}$ is generated as

$$y_i \sim Bell(W_0(exp\{\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}\})), \tag{27}$$

where $\beta_{p\times 1}$ was selected using the method given in [16]. In the simulation study, the sample size is chosen as $n = 50, 100, 150, 200, 250,$ and 300 , and the correlation coefficient is $\rho = 0.90, 0.95,$ and 0.99, and finally, the number of independent variables is taken as $p = 3, 5, 7$.

This study was done in R program [19] with 2000 repetitions. The results obtained by calculating the performances of the estimators with the MSE equation given in Eq. (28) are given in Table (1)-(3).

$$MSE(\widehat{\beta}) = \frac{1}{2000} \sum_{r=1}^{2000} (\widehat{\beta}_r - \beta)'(\widehat{\beta}_r - \beta). \tag{28}$$

When the Tables (1)-(3) are examined, the MSE of all estimators is decreasing as the sample size increases. As the correlation coefficient increases, the MSE values of all estimators increase in all scenarios. Similarly, increasing the number of independent variables negatively affects the performance of the estimators. The BTPRE has the smallest MSE value for each sample size and correlation coefficient in all designs. The result is that the proposed estimator is superior to MLE and BRE. In addition, two different parameter selection methods were used in the study. It is seen from the MSE values that the method mentioned as $BTPRE_2$ is better than $BTPRE_1$. It is seen that the smallest MSE value belongs to $BTPRE_2$ for all cases.

3.2. **Real Data Example.** In this subsection, an application study is given to support the simulation study. Mine fracture dataset provided by Myers *et al.*, consisting of n=44 observations, was used as the real dataset [15]. Dependent variable comprises the number of injuries in coal mines in the Appalachian region. Models used in modelling the dependent variable are PRM, NBRM and BRM. Akaike Information Information (AIC) value has been used to select the best model from the Poisson, Negative Binomial and Bell distributions. The results of the AIC are given in the Table (4). According to the results from the Table (4), the

TABLE 1. MSE values for p=3

n	ρ	$\widehat{\beta}_{MLE}$	$\widehat{\beta}_{k}$	$\widehat{\beta}_{qk_1}$	$\widehat{\beta}_{qk_2}$
	0.90	6.73214	5.87381	2.68612	1.68484
50	0.95	9.28130	7.71959	4.12875	2.35018
	0.99	33.36669	24.60985	17.43750	7.44138
	0.90	5.42670	4.97601	2.19685	1.37202
100	0.95	6.61325	5.79349	2.87992	1.72569
	0.99	16.08049	12.06670	7.10425	3.62146
	0.90	5.08468	4.78100	2.12913	1.29036
150	0.95	5.83406	5.27756	2.54194	1.50144
	0.99	12.69978	9.89156	6.30262	3.24474
	0.90	4.93791	4.71320	2.10739	1.25240
200	0.95	5.35580	4.92607	2.41291	1.39102
	0.99	10.10888	8.12280	4.82236	2.56908
	0.90	4.86601	4.68486	2.02834	1.19642
250	0.95	5.21363	4.87818	2.38204	1.35741
	0.99	9.04588	7.37682	4.44209	2.37815
	0.90	4.75503	4.60505	2.01804	1.19031
300	0.95	4.89462	4.62580	2.27716	1.31450
	0.99	7.99842	6.65926	3.86502	2.15689

appropriate model is chosen as the BRM since the Bell distribution has the smallest AIC value. Independent variables used in the dataset are as follows

X1: inner burden thickness in feet,
X2: percent extraction of the lower previously pricked mined seam,
X3: the lower seam height
X4: the time that the mine

The number of conditions used to determine whether the multicollinearity occurred in the data set is 296.5585. The correlation chart showing the correlation between the independent variables is given in Figure 1.

Because of existing multicollinearity, we calculate the MLE, RE and BTPRE coefficients for the data set. Then, the estimated coefficients, the standard errors and the square root of MSE values are given in Table (5).

When the Table (5) is examined, it is seen that BTPRE has the smallest MSE. The real data results show that the performance of the proposed BTPRE is superior to the MLE and RE, like the simulation studies. In addition, the method used to estimate k and q parameters in $TPRE_2$ is more effective than that of $TPRE_1$.

TABLE 2. MSE values for p=5

n	ρ	$\widehat{\beta}_{MLE}$	$\widehat{\beta}_{k}$	$\widehat{\beta}_{qk_1}$	$\widehat{\beta}_{qk_2}$
	0.90	9.82742	8.67058	2.91611	1.74571
50	0.95	14.70160	12.30335	5.20670	2.66722
	0.99	58.62601	43.50621	31.47012	10.27462
	0.90	7.71904	7.18999	2.27929	1.36482
100	0.95	9.53292	8.44894	3.16134	1.75353
	0.99	30.45618	23.44846	13.20561	5.28299
	0.90	7.00468	6.66178	1.99375	1.21469
150	0.95	8.02216	7.27704	2.60181	1.51511
	0.99	21.49925	16.84574	8.92770	3.77031
	0.90	6.65705	6.41170	1.96885	1.18348
200	0.95	7.47940	6.92626	2.35127	1.33888
	0.99	17.04084	13.75045	6.44487	3.01546
	0.90	6.51391	6.31932	1.88913	1.15233
250	0.95	6.94140	6.51257	2.27448	1.31927
	0.99	14.35469	11.78576	5.82564	2.69328
	0.90	6.43463	6.27732	1.87978	1.12367
300	0.95	6.87342	6.33572	2.00512	1.29277
	0.99	12.73392	10.52779	4.84047	2.41665

4. CONCLUSION

PRM and NBRM have been generally used in the modelling of count data. BRM has been widely preferred as an alternative to these models in recent years. BRM may be more suitable for modelling overdispersed count data. As seen in the real data set discussed in the study, the Bell distribution is more convenient than the alternative distributions. Considering this situation, alternative biased estimators are proposed for the Bell regression model to handle the multicollinearity problem.

In this article, we propose BTPRE as an alternative to these estimators. It is concluded from the simulation study and a real data example that the performance of the proposed estimator is superior to MLE and BRE.

Author Contribution Statements This article was produced from the PhD. thesis prepared by Melike Işılar under the supervision of Assoc. Prof. Dr. Y. Murat Bulut. All authors have read and approved the article.

Declaration of Competing Interests This work does not have any conflict of interest.

TABLE 3. MSE values for p=7

n	$\boldsymbol{\rho}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{MLE}}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{k}}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{qk_1}}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{qk_2}}$
	0.90	14.06553	12.77779	5.05469	2.72267
50	0.95	19.69486	16.76399	6.42533	3.13107
	0.99	82.85781	63.98084	55.16817	15.71594
	0.90	13.21807	12.63627	4.93636	2.41681
100	0.95	15.44038	14.15015	5.53295	2.74144
	0.99	43.16660	34.19774	19.14917	7.47775
	0.90	11.84296	11.46179	4.80369	2.39698
150	0.95	14.01502	13.16318	5.12539	2.67410
	0.99	31.62615	26.13034	13.52178	5.87188
	0.90	11.57151	11.30324	4.51272	2.10850
200	0.95	12.76125	12.16552	4.86460	2.19485
	0.99	25.18756	21.20189	10.95476	5.17243
	0.90	10.40783	10.19874	3.40672	2.03554
250	0.95	11.78194	11.30935	4.13962	2.10319
	0.99	21.31977	18.07933	7.76840	3.92102
	0.90	8.88708	8.71718	2.14829	1.34627
250	0.95	11.21841	10.82050	4.07510	2.07894
	0.99	18.49267	15.92869	7.09717	3.11654

TABLE 4. AIC values of dependent variable

	POISSON	**NEGATIVE BINOMIAL**	**BELL**
AIC	173.2554	172.3399	169.4784

TABLE 5. Results of the Mine fracture dataset

	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{2}}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{3}}$	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{4}}$	**MSE**
$\widehat{\boldsymbol{\beta}}$	0.00293	-0.01126	0.01819	-0.02384	-4.00837	1.38936
	(1.38907)	(0.00106)	(0.01684)	(0.00679)	(0.02179)	
$\widehat{\boldsymbol{\beta}}_{\boldsymbol{k}}$	0.00294	-0.01126	0.01819	-0.02384	-3.57858	1.24044
	(1.31249)	(0.00106)	(0.01599)	(0.00676)	(0.02178)	
$\widehat{\boldsymbol{\beta}}_{\boldsymbol{qk_1}}$	0.00322	-0.00545	0.00117	-0.00021	-0.00001	0.00147
	(4.00788)	(0.00247)	(0.06173)	(0.00514)	(0.02970)	
$\widehat{\boldsymbol{\beta}}_{\boldsymbol{qk_2}}$	-0.00294	-0.00140	0.00020	-0.00003	-0.00001	0.00043
	(4.00784)	(0.00359)	(0.06547)	(0.00367)	(0.02969)	

Acknowledgements Melike Işılar's doctoral education and thesis work are supported by TÜBİTAK Directorate of Science Fellowships and Grant Programmes (BİDEB) in the 2211 National PhD scholarship program.

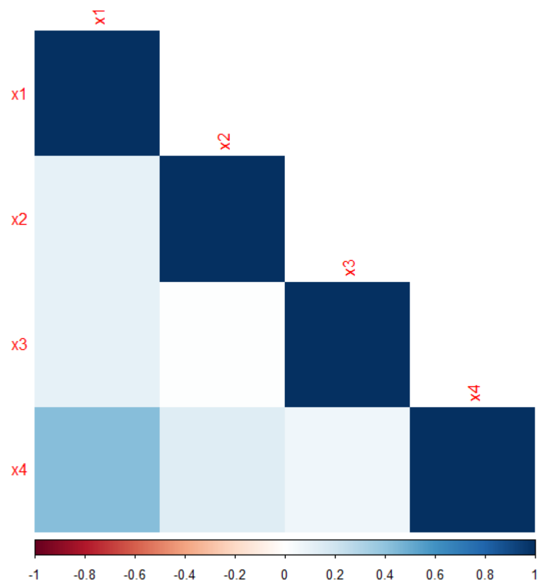

FIGURE 1. Correlation chart between independent variables

REFERENCES

[1] Alheety, M. I., Qasim, M., Månsson, K., Kibria, B. M., Modified almost unbiased two-parameter estimator for the Poisson regression model with an application to accident data, *SORT*, 45 (2021), 121-142. DOI: 10.2436/20.8080.02.112

[2] Algamal, Y. A., Developing a ridge estimator for the gamma regression model, *Journal of Chemometrics*, (2018), 32. DOI:10.1002/cem.3054

[3] Algamal, Y. A., Performance of ridge estimato in inverse Gaussian regression model, *Communications in Statistics-Theory and Methods*, 48(15) (2019), 3836–3849. DOI: 10.1080/03610926.2018.1481977

[4] Amin, M., Akram, M. N., Majid, A., On the estimation of Bell regression model using ridge estimator, *Communications in Statistics-Simulation and Computation*, (2021), https://doi.org/10.1080/03610918.2020.1870694

[5] Asar, Y., Genç, A., Two-parameter Ridge estimator in the binary logistic regression, *Comm. Statist. Simulation Comput.*, 46(9) (2017), 7088–7099. DOI: 10.1080/03610918.2016.1224348

[6] Bell, E. T., Exponential numbers, *The American Mathematical Monthly*, 41(7) (1934), 411–419.

[7] Bulut, Y. M., Işılar, M., Two parameter Ridge estimator in the inverse Gaussian regression model, *Hacettepe Journal of Mathematics and Statistics*, 50(3) (2021), 895–910. DOI : 10.15672/hujms.813540

[8] Castellares, F., Ferrari, S. L. P., Lemonte, A. J., On the Bell distribution and its associated regression model for count data, *Applied Mathematical Modelling*, 56 (2018), 172–185. DOI: 10.1016/j.apm.2017.12.014

[9] Hoerl, A. E., Kennard, R. W., Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, 42(1) (1970), 80–86, http://www.jstor.org/stable/1271436

[10] Lipovetsky, S., Two parameter Ridge regression and its convergence o the eventual pairwise model, *Math Comput Model*, 44 (2006), 304–318. DOI: 10.1016/j.mcm.2006.01.017

[11] Lipovetsky, S., Conklin, W. M., Ridge regression in two-parameter solution, *Appl. Stoch. Models Bus. Ind.*, 21(6) (2005), 525–540. DOI: 10.1002/asmb.603

[12] Månsson, K., On ridge estimators for the negative binomial regression model, *Economic Modelling*, 29 (2012), 178–184. DOI: 10.1016/j.econmod.2011.09.009

[13] Månsson, K., Shukur, G., A Poisson ridge regression estimator, *Econ. Model.,* 28 (2011), 1475–1481. DOI:10.1016/j.econmod.2011.02.030

[14] McDonald, G. C., Galarneau, D. I., A monte carlo evaluation of some ridge-type estimators, *Journal of the American Statistical Association*, 70(350) (1975), 407–416.

[15] Myers, R., Montgomery, D., Vining, G., Robinson, T., Generalized Linear Models with Applications in Engineering and the Sciences, Second Edition, Wiley, A John Wiley Sons, Inc., Publication, 2012.

[16] Newhouse, J. P., Oman, S. D., An evaluation of ridge estimators, *Rand Corporation (p-716-PR), Santa Monica*, (1971), 1–16, https://doi.org/10.1080/00949655.2018.1498502

[17] Qasim, M., Kibria, B. M. G., Månsson, K., Sjölander, P., A new Poisson Liu regression estimator: method and application, *Journal of Applied Statistics*, 47(12) (2020), 2258–2271. DOI: 10.1080/02664763.2019.1707485

[18] Qasim, M., Bulut, Y. M., Mansson, K., The Wald-type confidence interval on the mean response function of the Poisson inverse Gaussian Ridge regression, Accepted: October 2023. REVSTAT-Statistical Journal.

[19] R Core Team, R: A Language and Environment for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing, 2014, http://www.R-project.org/

[20] Scahaefer, R. L., Roi, L. D., Wolfe, R. A., A ridge logistic estimator, *Communications in Statistics-Theory and Methods*, 13 (1984), 99–113.

[21] Toker, S., Kaçıranlar, S., On the performance of two parameter ridge estimator under the mean square error criterion, *Appl. Math. Comput.*, 219 (2013), 4718–4728. DOI: 10.1016/j.amc.2012.10.088