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/i20705948v11n2p506

Proposed methods in estimating the ridge regression parameter in Poisson regression model

By Alanaz, Algamal

Published: 14 October 2018

Electronic Journal of Applied Statistical Analysis
Vol. 11, Issue 02, October 2018, 506-515
DOI: 10.1285/i20705948v11n2p506

Proposed methods in estimating the ridge regression parameter in Poisson regression model

Mazin M. Alanaz[a] and Zakariya Yahya Algamal*[b]

[a] *Department of Operation Research and Intelligence Techniques, University of Mosul*
[b] *Department of Statistics and Informatics, University of Mosul*

Published: 14 October 2018

Poisson regression model is considered as an important model among the linear logarithm models. It is usually used to model the count dependent variable. However, as in linear regression model, the multicollinearity problem may be present leading to negatively affect the model parameter estimation. In this study, several methods are proposed to estimate the ridge parameter. Monte-Carlo simulation studies with different factors were conducted to evaluate the performance of the used estimators. The results demonstrate the better performance of the proposed estimator compared to other used estimators in terms of mean squared error (MSE).

keywords: Multicollinearity; ridge estimator; Poisson regression model; shrinkage; Monte Carlo simulation.

1 Introduction

Poisson regression model is widely applied for studying several real data problems, such as in mortality studies where the aim is to investigate the number of deaths and in health insurance where the target is to explain the number of claims made by the individual (Algamal, 2012; Cameron and Trivedi, 2013; De Jong and Heller, 2008). In dealing with the Poisson regression model, it is assumed that the problem of multicollinearity does not exist. However, when this problem exists, the maximum likelihood (ML) estimation of the coefficients are become unstable with a high variance (Algamal, 2018a,b).

*Corresponding author: Tel.: +964 7701640834, E-mail address: zakariya.algamal@uomosul.edu.iq

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

Numerous procedures have been proposed to to deal with multicollinearity. The ridge regression method (RR) (Hoerl and Kennard, 1970), among them, has been consistently demonstrated to be a well-known procedures. RR tries to shrink the regression coefficients toward zero to decrease the large variance (Asar and Genç, 2017). This can be done by adding a positively constant amount to the diagonal of $\mathbf{X}^T\mathbf{X}$.

In classical linear regression models the following relationship is usually adopted

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

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, $\beta = (\beta_1, ..., \beta_p)$ is a $p \times 1$ vector of unknown regression coefficients, and ε is an $n \times 1$ vector of random errors with mean 0 and variance σ^2. In linear regression, the RR is defined as

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

where $\mathbf{I}$ is the identity matrix with dimension $p \times p$ and $k \geq 0$ represents the ridge parameter (shrinkage parameter). The ridge parameter, k, controls the shrinkage of β toward zero.

2 Statistical methodology

2.1 Poisson ridge regression model

Count data often arise in epidemiology, social, and economic studies. This type of data consists of positive integer values. Poisson distribution is a well-known distribution that fit to such type of data. Poisson regression model is used to model the relationship between the counts as response variable and potentially explanatory variables (Algamal and Lee, 2015; KaÇiranlar and Dawoud, 2017).

$$\hat{\beta}_{Ridge} = (\mathbf{X}^T\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}, \tag{3}$$

Let y_i be the response variable and follows a Poisson distribution with mean θ_i, then the probability density function is defined as

$$f(y_i) = \frac{e^{-\theta_i}{\theta_i}^{y_i}}{y_i!}, \quad y_i = 0, 1, \ldots; \; i = 1, 2, \ldots, n. \tag{4}$$

In a Poisson regression model, $\ln(\theta_i) = \mathbf{x}_i^T\beta$ is expressed as a linear combination of explanatory variables $\mathbf{x}_i = (x_{i1}, ..., x_{ip})^T$. The $\ln(\theta_i)$ is called as canonical link function which making the relationship between explanatory variables and response variable linear. The most common method of estimating the coefficients of Poisson regression model is to use the maximum likelihood method. Given the assumption that the observations are independent, the log-likelihood function is defined as

$$\ell(\beta) = \sum_{i=1}^{n} \left\{ y_i\mathbf{x}_i^T\beta - \exp(\mathbf{x}_i^T\beta) - \ln y_i! \right\}. \tag{5}$$

The ML estimator is then obtained by computing the first derivative of the Eq. (5) and setting it equal to zero, as

$$\frac{\partial \ell(\beta)}{\partial \beta} = \sum_{i=1}^{n} \left[y_i - \exp(\mathbf{x}_i^T \beta) \right] \mathbf{x}_i = 0. \tag{6}$$

Because Eq. (4) is nonlinear in β, the iteratively weighted least squares (IWLS) algorithm can be used to obtain the ML estimators of the Poisson regression parameters (PR) as

$$\hat{\beta}_{PRM} = (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{W}} \hat{\mathbf{v}}, \tag{7}$$

where $\hat{\mathbf{W}} = \mathbf{diag}(\hat{\theta}_i)$ and $\hat{\mathbf{v}}$ is a vector where ith element equals to $\hat{v}_i = \ln(\hat{\theta}_i) + ((y_i - \hat{\theta}_i)/\hat{\theta}_i)$. The ML estimator is asymptotically normally distributed with a covariance matrix that corresponds to the inverse of the Hessian matrix

$$\mathbf{cov}(\hat{\beta}_{PRM}) = \left[-E\left(\frac{\partial^2 \ell(\beta)}{\partial \beta_i \, \partial \beta_k} \right) \right]^{-1} = (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X})^{-1}. \tag{8}$$

The mean squared error (MSE) of Eq. (5) can be obtained as

$$\begin{aligned} \mathbf{MSE}\,(\hat{\beta}_{PRM}) &= E(\hat{\beta}_{PRM} - \hat{\beta})^T (\hat{\beta}_{PRM} - \hat{\beta}) \\ &= tr[(\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X})^{-1}] \\ &= \sum_{j=1}^{p} \frac{1}{\lambda_j}, \end{aligned} \tag{9}$$

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

When the multicollinearity problem exist, the matrix $\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X}$ becomes ill-conditioned leading to high variance and instability of the ML estimator of the Poisson regression parameters. As a remedy, Månsson and Shukur (2011) proposed the Poisson ridge regression model (PRRM) as

$$\begin{aligned} \hat{\beta}_{PRRM} &= (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^T \hat{\mathbf{W}} \mathbf{X} \hat{\beta}_{PRM} \\ &= (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^T \hat{\mathbf{W}} \hat{\mathbf{v}}, \end{aligned} \tag{10}$$

where $k \geq 0$. The ML estimator can be considered as a special estimator from Eq. (8) with $k = 0$. Regardless of $k$value, the MSE of the $\hat{\beta}_{PRRM}$ is smaller than that of $\hat{\beta}_{PRM}$ because the MSE of $\hat{\beta}_{PRRM}$ is equal to Kibria et al. (2015)

$$\mathbf{MSE}(\hat{\beta}_{PRRM}) = \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \sum_{j=1}^{p} \frac{\alpha_j}{(\lambda_j + k)^2}, \tag{11}$$

where α_j is defined as the jth element of $\gamma \hat{\beta}_{PRM}$and γ is the eigenvector of the $\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X}$ matrix. Comparing with the MSE of Eq. (5), $\mathbf{MSE}(\hat{\beta}_{PRRM})$ is always small for $k > 0$.

3 Estimating the shrinkage parameter k

The efficiency of RR is depending on k which controls the amount of the shrinkage. When $k = 0$, then ML estimates can be obtained. On the other hand, when k increasing, the influence of k increases on the coefficient estimates. In our paper, a large number of methods are considered to estimate the value of k in Poisson ridge regression model. The idea behind all these methods is obtained from the work by Hoerl and Kennard (1970), Kibria (2003), and Kibria et al. (2015).

1. Hoerl and Kennard (1970) (HK1 and HK2), which are, respectively, defined as

$$HK1 = \frac{p\hat{\sigma}^2}{\hat{\alpha}^T\hat{\alpha}}, \quad j = 1, 2, ..., p, \tag{12}$$

$$HK2 = \frac{\hat{\sigma}^2}{\hat{\alpha}^2_{\max}}, \tag{13}$$

Where $\hat{\alpha}$ is defined as the jth element of $\gamma\,\hat{\beta}_{GRM}$and γ is the eigenvector of the $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$ matrix, $\hat{\alpha}_{\max}$is the maximum value of $\hat{\alpha}$, and $\hat{\sigma}^2 = (y_i - \hat{\mu}_i)^2/n - p - 1$.

2. Kibria et al. (2015) used several methods (K1-K12). They are, respectively, defined as

$$\mathbf{K1} = \mathbf{max}\left\{\frac{1}{m_j}\right\}, \tag{14}$$

$$\mathbf{K2} = \mathbf{max}\left\{m_j\right\}, \tag{15}$$

$$\mathbf{K3} = \prod_{j=1}^{p}\left\{\frac{1}{m_j}\right\}^{\frac{1}{p}}, \tag{16}$$

$$\mathbf{K4} = \prod_{j=1}^{p}\left\{m_j\right\}^{\frac{1}{p}}, \tag{17}$$

$$\mathbf{K5} = \mathbf{median}\left\{\frac{1}{m_j}\right\}, \tag{18}$$

$$\mathbf{K6} = \mathbf{median}\left\{m_j\right\}, \tag{19}$$

$$\mathbf{K7} = \mathbf{max}\left\{\frac{1}{q_j}\right\}, \tag{20}$$

$$\mathbf{K8} = \mathbf{max}\left\{q_j\right\}, \tag{21}$$

$$\mathbf{K9} = \prod_{j=1}^{p}\left\{\frac{1}{q_j}\right\}^{\frac{1}{p}}, \tag{22}$$

$$\mathbf{K10} = \prod_{j=1}^{p}\left\{q_j\right\}^{\frac{1}{p}}, \tag{23}$$

$$\mathbf{K11} = \mathbf{median}\left\{\frac{1}{q_j}\right\}, \tag{24}$$

$$\mathbf{K12} = \mathbf{median}\left\{q_j\right\}, \tag{25}$$

where $m_j = \sqrt{\hat{\sigma}^2/\hat{\alpha}_j^2}$ and $q_j = \lambda_{\max}/(n-p)\hat{\sigma}^2 + \lambda_{\max}\hat{\alpha}_j^2$.

4 the Proposed Methods

In this section, we extend the idea of Asar et al. (2014) and Bhat (2016) to the Poisson ridge estimator.

1. Asar et al. (2014) proposed five modifications of ridge parameter. They are defined as, respectively

$$\mathbf{A1} = \frac{p^2}{\lambda_{\max}^2}\frac{\hat{\sigma}^2}{\sum_{j=1}^{p}\hat{\alpha}_j^2}, \tag{26}$$

$$\mathbf{A2} = \frac{p^3}{\lambda_{\max}^3}\frac{\hat{\sigma}^2}{\sum_{j=1}^{p}\hat{\alpha}_j^2}, \tag{27}$$

$$\mathbf{A3} = \frac{p}{(\lambda_{\max})^{1/3}}\frac{\hat{\sigma}^2}{\sum_{j=1}^{p}\hat{\alpha}_j^2}, \tag{28}$$

$$\mathbf{A4} = \frac{p}{(\sum_{j=1}^{p}\sqrt{\lambda_i})^{1/3}}\frac{\hat{\sigma}^2}{\sum_{j=1}^{p}\hat{\alpha}_j^2}, \tag{29}$$

$$\mathbf{A5} = \frac{2p}{\sqrt{\lambda_{\max}}}\frac{\hat{\sigma}^2}{\sum_{j=1}^{p}\hat{\alpha}_j^2}, \tag{30}$$

2. Bhat (2016) proposed two modifications of HK1. They are defined as, respectively

$$\mathbf{B1} = \frac{p\hat{\sigma}^2}{\hat{\alpha}^T\hat{\alpha}} + \frac{1}{\lambda_{\max}\hat{\alpha}^T\hat{\alpha}}, \tag{31}$$

$$\mathbf{B2} = \frac{p\hat{\sigma}^2}{\hat{\alpha}^T\hat{\alpha}} + \frac{1}{2\left(\sqrt{\lambda_{\max}/\lambda_{\max}}\right)^2}, \tag{32}$$

5 Simulation study

In this section, a Monte Carlo simulation experiment is used to examine the performance of these methods in Poisson ridge with different degrees of multicollinearity.

5.1 Simulation design

The response variable of n observations is generated from Poisson regression model by

$$\theta_i = \exp(\mathbf{x}_i^T \beta), \tag{33}$$

where $\beta = (\beta_0, \beta_1, ..., \beta_p)$ with $\sum_{j=1}^{p} \beta_j^2 = 1$ and $\beta_1 = \beta_2 = ... = \beta_p$ (Månsson and Shukur, 2011; Kibria, 2003). In addition, because the value of intercept, β_0, has an effect on θ_i, three values are chosen $\beta_0 \in \{1, 0, -1\}$, where decreasing the value of β_0 leads to lower average value of θ_i, which leads to less variation (Månsson and Shukur, 2011). The explanatory variables $\mathbf{x}_i^T = (x_{i1}, x_{i2}, ..., x_{in})$ have been generated from the following formula

$$x_{ij} = (1 - \rho^2)^{1l2} w_{ij} + \rho w_{ip}, \quad i = 1, 2, ..., n, \quad j = 1, 2, ..., p, \tag{34}$$

where ρ represents the correlation between the explanatory variables and w_{ij}s are independent standard normal pseudo-random numbers. Because the sample size has direct impact on the prediction accuracy, three representative values of the sample size are considered: 30, 50 and 100. In addition, the number of the explanatory variables is considered as $p = 3$ and $p = 7$ because increasing the number of explanatory variables can lead to increase the MSE. Further, because we are interested in the effect of multicollinearity, in which the degrees of correlation considered more important, three values of the pairwise correlation are considered with $\rho = \{0.90, 0.95, 0.99\}$. For a combination of these different values of n, p, β_0, and ρ the generated data is repeated 1000 times and the averaged mean squared errors (MSE) is calculated as

$$\mathbf{MSE}(\hat{\beta}) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\beta} - \beta)^T (\hat{\beta} - \beta), \tag{35}$$

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

5.2 Simulation results

The estimated MSE of Eq. (35) for all the different selection methods of k and the combination of n, p, and ρ, are respectively summarized in Tables 1-3. Several observations can be obtained as follows: "

1. In terms of ρ values, there is increasing in the MSE values when the correlation degree increases regardless the value of n and p.

2. Regarding the number of covariates, it is easily seen that there is a negative impact on MSE , where there is increasing in its values when the p increasing from three covariates to seven covariates.

3. With respect to the value of n, the MSE values decrease when n increases, regardless the value of $\rho$and p.

4. All the selection methods of k are superior to the ML estimator in terms of MSE.

5. Clearly, in terms of MSE, $\mathbf{K}_2$ and $\mathbf{K}_8$ improved the performance of the Poisson ridge regression compared to ML estimator in all the cases without any domination. In contrast, $\mathbf{A}_2$ estimator attained poor results comparing with the other used estimators in all cases.

6. For comparisons between the modification estimators of $\mathbf{HK}_1$, i.e. $\mathbf{B}_1$ and $\mathbf{B}_2$, it is seen that $\mathbf{B}_2$ achieves the lowest MSE compared to $\mathbf{B}_1$.

"

Table 1: MSE values when $n = 30$

Method	$p = 3$			$p = 7$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	1.703	3.064	3.713	9.806	12.633	13.65
HK1	1.571	3.131	3.304	4.688	5.088	6.118
HK2	1.622	3.224	3.371	4.484	10.014	11.031
K1	1.637	1.652	3.174	5.516	4.238	5.255
K2	1.349	2.549	3.228	1.438	1.476	2.493
K3	1.681	2.881	2.941	9.012	9.705	10.722
K4	1.523	2.723	2.967	3.17	3.21	4.234
K5	1.658	2.858	3.028	8.86	8.994	10.011
K6	1.565	2.765	3.203	3.441	3.922	4.939
K7	1.694	2.894	3.129	2.599	3.722	4.739
K8	1.341	2.541	3.308	1.969	2.045	3.062
K9	1.7	2.9	3.099	9.018	9.674	10.691
K10	1.374	2.574	2.991	3.161	3.183	4.2
K11	1.698	2.898	2.945	9.241	9.658	10.675
A1	1.704	2.904	3.713	9.806	12.632	13.649
A2	1.706	2.906	3.715	9.809	12.633	13.65
A3	1.672	2.872	3.52	8.538	10.533	11.55
A4	1.644	2.844	3.299	7.358	8.571	9.588
A5	1.678	2.878	3.55	8.879	11.088	12.105
B1	1.57	2.77	2.796	4.687	5.076	6.093
B2	1.565	2.765	2.793	4.629	4.996	6.013

Table 2: MSE values when $n = 50$

Method	$p = 3$			$p = 7$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	1.6159	2.9769	3.6259	9.7189	12.5459	13.5629
HK1	1.4839	3.0439	3.2169	4.6009	5.0009	6.0309
HK2	1.5349	3.1369	3.2839	4.3969	9.9269	10.9439
K1	1.5499	1.5649	3.0869	5.4289	4.1509	5.1679
K2	1.2619	2.4619	3.1409	1.3509	1.3889	2.4059
K3	1.5939	2.7939	2.8539	8.9249	9.6179	10.6349
K4	1.4359	2.6359	2.8799	3.0829	3.1229	4.1469
K5	1.5709	2.7709	2.9409	8.7729	8.9069	9.9239
K6	1.4779	2.6779	3.1159	3.3539	3.8349	4.8519
K7	1.6069	2.8069	3.0419	2.5119	3.6349	4.6519
K8	1.2539	2.4539	3.2209	1.8819	1.9579	2.9749
K9	1.6129	2.8129	3.0119	8.9309	9.5869	10.6039
K10	1.2869	2.4869	2.9039	3.0739	3.0959	4.1129
K11	1.6109	2.8109	2.8579	9.1539	9.5709	10.5879
A1	1.6169	2.8169	3.6259	9.7189	12.5449	13.5619
A2	1.6189	2.8189	3.6279	9.7219	12.5459	13.5629
A3	1.5849	2.7849	3.4329	8.4509	10.4459	11.4629
A4	1.5569	2.7569	3.2119	7.2709	8.4839	9.5009
A5	1.5909	2.7909	3.4629	8.7919	11.0009	12.0179
B1	1.4829	2.6829	2.7089	4.5999	4.9889	6.0059
B2	1.4779	2.6779	2.7059	4.5419	4.9089	5.9259

Table 3: MSE values when $n = 100$

Method	$p = 3$			$p = 7$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	1.5538	2.9148	3.5638	9.6568	12.4838	13.5008
HK1	1.4218	2.9818	3.1548	4.5388	4.9388	5.9688
HK2	1.4728	3.0748	3.2218	4.3348	9.8648	10.8818
K1	1.4878	1.5028	3.0248	5.3668	4.0888	5.1058
K2	1.1998	2.3998	3.0788	1.2888	1.3268	2.3438
K3	1.5318	2.7318	2.7918	8.8628	9.5558	10.5728
K4	1.3738	2.5738	2.8178	3.0208	3.0608	4.0848
K5	1.5088	2.7088	2.8788	8.7108	8.8448	9.8618
K6	1.4158	2.6158	3.0538	3.2918	3.7728	4.7898
K7	1.5448	2.7448	2.9798	2.4498	3.5728	4.5898
K8	1.1918	2.3918	3.1588	1.8198	1.8958	2.9128
K9	1.5508	2.7508	2.9498	8.8688	9.5248	10.5418
K10	1.2248	2.4248	2.8418	3.0118	3.0338	4.0508
K11	1.5488	2.7488	2.7958	9.0918	9.5088	10.5258
A1	1.5548	2.7548	3.5638	9.6568	12.4828	13.4998
A2	1.5568	2.7568	3.5658	9.6598	12.4838	13.5008
A3	1.5228	2.7228	3.3708	8.3888	10.3838	11.4008
A4	1.4948	2.6948	3.1498	7.2088	8.4218	9.4388
A5	1.5288	2.7288	3.4008	8.7298	10.9388	11.9558
B1	1.4208	2.6208	2.6468	4.5378	4.9268	5.9438
B2	1.4158	2.6158	2.6438	4.4798	4.8468	5.8638

6 Conclusion

In this paper, several selection methods of the k are investigated in Poisson ridge regression model. According to simulation studies, it has been seen that some of these selection methods can make improvement relative to others, in terms of MSE. In conclusion, the use of K2 and K8 is recommended when multicollinearity is present in the Poisson ridge regression model.

References

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

Algamal, Z. Y. (2018a). Developing a ridge estimator for the gamma regression model. *Journal of Chemometrics*, page e3054.

Algamal, Z. Y. (2018b). Shrinkage estimators for gamma regression model. *Electronic Journal of Applied Statistical Analysis*, 11:253–268.

Algamal, Z. Y. and Lee, M. H. (2015). Penalized poisson regression model using adaptive modified elastic net penalty. *Electronic Journal of Applied Statistical Analysis*, 8(2):236–245.

Asar, Y. and Genç, A. (2017). A new two-parameter estimator for the poisson regression model. *Iranian Journal of Science and Technology, Transactions A: Science*.

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.

Bhat, S. S. (2016). A comparative study on the performance of new ridge estimators. *Pakistan Journal of Statistics and Operation Research*, 12(2):317–325.

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

De Jong, P. and Heller, G. Z. (2008). *Generalized linear models for insurance data*, volume 10. Cambridge University Press Cambridge.

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

KaÇiranlar, S. and Dawoud, I. (2017). On the performance of the poisson and the negative binomial ridge predictors. *Communications in Statistics - Simulation and Computation*, pages 0–0.

Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics - Simulation and Computation*, 32(2):419–435.

Kibria, B. M. 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.

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