Iraqi Journal of Statistical Sciences

http://stats.uomosul.edu.iq

Shrinkage Estimators in Bell Regression Model: Subject Review

Hutheyfa Hazem Taha

Department of Operations Research and Intelligent Techniques, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq.

Article information

Article history:
Received: June 25,2025
Revised: August 31,2025
Accepted: September 14,2025
Available online: December 1,2025

Keywords:
Regression; Bell Regression; Estimator; Shrinkage.

Correspondence:
Hutheyfa Hazem taha
hutheyfa71@uomosul.edu.iq

Abstract

Bell regression model has become a very versatile model that replaced the conventional count data models and helps to resolve the problem of over-dispersion where the variance of data points surpasses the mean. Nevertheless, in practice, the classical maximum likelihood estimators (MLE) of the parameters of a model are frequently affected by multicollinearity among the explanatory variables, and they yield highly unstable estimates and inflated variances. To address these difficulties, estimation methods developed to estimate shrinkage, such as ridge or Liu estimators have been applied to the Bell regression model. In the subject review, new developments in estimators of shrinkage of Bell regression models are proportionate in discussing their theoretical background in knowledge, estimation process, and asymptotic characteristics. The results on Monte Carlo simulation studies always show that shrinkage estimators outweigh MLEs in that they minimize mean squared error and bias more than MLEs, especially in cases of multicollinearity. Both overall, estimation methods of shrinkage is a considerable improvement in Bell regression modeling that offers certainty and effectiveness of analysis of complicated counts information utilizing problematic distribution attributes.

DOI:10.33899/iqjoss.V22I2.54082

1. Introduction

As statistical modeling aids in describing the slope of the functionality between the variable of interest that is the response variable and the many explanatory variables, it is significant in most spheres of scientific studies. In the model, it is also assumed that the observations made in the dependent variable were independent and identically distributed. The assumption however might not have a lot of use in the real world especially in the usage of these technologies in day-to-day life. An example is that the response variable in the medical sciences is a variable where one can have a positive skew. It is therefore not quite sensible to utilize a linear regression model at all in some way. The GLM linear regression models are slowly creeping into other models to be a statistical modeling technique that can be used in both continuous and discrete dependent variable Algamal, Lukman, Golam, Taofik and Mathematics (2023).

Multicollinearity is a problem regarding econometric modeling. It reveals that there is a strong relationship between the explanatory variables. The covariance matrix of ML estimator is infamous to be ill-conditioned in case of severe multicollinearity. Among the adverse results of such a situation, the variance of regression estimates is overestimated. The significance and the size of coefficients are altered as a consequence. Most of the classical steps made in resolving this problem include deleting correlated variables, gathering additional data or re-specifying the model.

In order to overcome multicollinearity problem in the linear regression model, one different alternative method to MLE is the ridge, Liu, Liu type and other estimators based on the other authors Hoerl and Kennard, (1970), K.

Liu(1993). These estimators have been extended to the GLMs (Akram, Amin, and Amanullah(2020), B. M. G. Kibria (2003) G. Kibria, Månsson and Shukur (2012), Kurtoğlu and Özkale(2016), Mackinnon and Puterman(1989), Månsson and Shukur (2011), Nyquist(1991) Segerstedt(1992), Shamany, Alobaidi and Algamal (2019).

2. Bell regression model

Count data modeling Statistical methods intended to model dependent variables which are counts, i.e. non-negative integers indicating the number of times that an event has happened in a fixed time or space. The data are commonly observed in disciplines like economics, health sciences, ecology and social sciences, where the number of visits to a hospital, accidents, amount of purchase by customers are among results which are observed. The models of count data are usually formulated in the framework of generalized linear model (GLM), which is estimated by maximum likelihood. In them, they can add covariates to model variability in counts, can fit one or more offset terms to reflect the exposure time or population at risk.

Let (w_i, z_i), $i = 1,2,\ldots,n$is independent observed data with the predictor vector $z_i \in R^{p+1} w_i \in R$ belongs to the Bell distribution. Then the pdf of w_i is

$$P(W = w) = \frac{\alpha^w e^{-e\alpha+1} B_w}{w!}, \; w = 0,1,2,\ldots, \quad (1)$$

where $\alpha > 0$ and $B_w = (1/e) \sum_{d=0}^{\infty} (d^w/d!)$ is the Bell numbers (Eric T Bell (1934), Eric Temple Bell(1934), Castellares, Ferrari and Lemonte(2018)). Then:

$$E(w) = \alpha e^{\alpha}, \quad (2)$$

$$Var(w) = \alpha(1+\alpha)e^{\alpha}. \quad (3)$$

Assuming $\Omega = \alpha e^{\alpha}$ and $\alpha = Q_{\circ}(\Omega)$ where $Q_{\circ}(.)$ is the Lambert function. Eq. (1) can be as

$$P(W = w) = \exp\left(1 - e^{Q_{\circ}(\Omega)}\right) \frac{Q_{\circ}(\Omega)^w B_w}{w!}, \; w = 0,1,2,\ldots, \quad (4)$$

The linear function is stated as $\eta_i = \beta_0 + \sum_{j=1}^{p} z_{ij}\beta_j = z_i^T \beta$ with $z_i^T = (1, z_{i2}, z_{i3}, \ldots, z_{ip})$ and $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$. The link function is as $\mu_i = g^{-1}(\eta_i) = g^{-1}(z_i^T \beta)$. The Bell regression model (BRM) can be modeled by assuming $\Omega_i = \exp(z_i^T \beta) \; \exp\left(\exp(z_i^T \beta)\right)$ and $\log\psi_i = x_i^T \beta \; \exp(x_i^T \beta)$ as
$w_i \sim \text{Bell}\left(Q_{\circ}(\Omega_i)\right)$. The log-likelihood is defined

$$\ell(\beta, \Omega) = \sum_{i=1}^{n} w_i \log\left(\exp(z_i^T \beta) \exp\left(e^{(z_i^T \beta)}\right)\right) + \sum_{i=1}^{n} (1 - e^{e^{(z_i^T \beta)} e^{e^{(z_i^T \beta)}}}) + \log B_w - \log(\textstyle\prod_{i=1}^{n} w_i!). \quad (5)$$

Then, the MLE is

$$\hat{\beta}_{\text{MLE}} = (Z^T \widehat{M} Z)^{-1} Z^T \widehat{M} \hat{u}, \quad (6)$$

where $\widehat{M} = \text{diag}[(\partial\mu_i/\partial\eta_i)^2 / V(w_i)]$ and $\hat{u}$ is a vector where i$^{\text{th}}$ element equals to $\hat{u}_i = \log\widehat{\Omega}_i + [(w_i - \hat{\mu}_i)/\sqrt{\text{var}(\widehat{\Omega}_i)}]$.

In the presence of multicollinearity, the near singularity of $Z^T \hat{M} Z$ makes the estimation unstable and enlarges the variance G. Liu and Piantadosi (2016). The ridge estimator (RE) Hoerl & Kennard (1970), Liu estimator (LE) K. Liu (1993) are alternative to the MLE, when multicollinearity exists. RR and LE have been suggested in Bell regression model by Amin, Akram, and Majid (2021) and Majid, Amin, and Akram (2021), respectively.

3. Shrinkage Estimators

3.1 RE Estimator

Månsson and Shukur (2011) proposed the RE as

$$\hat{\beta}_{RE} = (Z^T\widehat{M}Z + kI)^{-1}Z^T\widehat{M}\hat{u}, \tag{7}$$

where $k \geq 0$ (B. M. Golam Kibria, Månsson, & Shukur, 2015).

3.2 Liu Estimator

The Liu estimator is a biased regression procedure, which is used when it remains that there may be a higher inter-correlation in predictor variables in a linear model that presents multicollinearity. Proposed by Liu (1993), it conforms a middle ground between the OLS estimator and ridge regression as a means to balance bias and variance to offer more stable estimations. The Liu estimator (LE) is defined as

$$\hat{\beta}_{LE} = (Z^T\widehat{M}Z + I)^{-1}(Z^T\widehat{M}Z + d\,I)\hat{\beta}_{MLE}, \tag{8}$$

where $0 < d < 1$. Regardless of d value, the MSE of the $\hat{\beta}_{LE}$is smaller than that of $\hat{\beta}_{MLE}$because the MSE of $\hat{\beta}_{LE}$ B. M. Golam Kibria et al.(2015).

3.3 Liu-type Estimator

Introduced by Kejian Liu (2003), the Liu-type estimator is a biased estimator of the regression, which tries to solve problems of multicollinearity in linear and generalized regression. It generalizes both OLS and ridge regressions by adding one or two shrinkage (biasing) parameters that allow one to manage the bias-variance trade-off when estimating the parameters. The Liu-type estimator (LT) is defined as

$$\hat{\beta}_{LT} = (Z^T\widehat{M}Z + k\,I)^{-1}(Z^T\widehat{M}Z - d\,I)\hat{\beta}_{MLE}, \tag{9}$$

where $-\infty < d < \infty$ and $k \geq 0$ (Alheety and Golam Kibria(2013) , B. M. Golam Kibria and Saleh (2004) ,Norouzirad and Arashi (2017), Wu (2014, 2016).

3.4 Two-parameter Estimator

Asar and Genç (2017) and Huang and Yang (2014) proposed the two-parameter estimator (TP) is defined as:

$$\hat{\beta}_{TP} = (Z^T\widehat{M}Z + k\,I)^{-1}(Z^T\widehat{M}Z + k\,d\,I)\hat{\beta}_{MLE}. \tag{10}$$

The $\hat{\beta}_{TP}$ is a combination of two different estimators generalized RE and generalized LE.

4. Simulation Study

In this part we will simulate a pair of collinear explanatory variables and a response variable w which is bell distributed. The following variables are provided as the explanatory ones according to the research of Kibria (2003) and Lukman et al. (2019a, b):

$$z_{ij} = \sqrt{(1-r^2)}x_{ij} + rx_{i(j+1)}, i = 1, \dots, n; j = 1, \dots p \tag{11}$$

where x_{ij} are independent standard normal pseudo-random numbers and ρ^2 denotes the correlation between the explanatory variables such that $r =$ 0.9, 0.95 and 0.99. We assumed that $w_i \sim bell\big(Q_o(\mu_i)\big)$, where

$$\log(\mu_i) = \eta_i = \beta_1 z_{i1} + \beta_2 z_{i2} + \cdots + \beta_p z_{ip} \tag{12}$$

Then n=30, 50, 100, and 200 while p is taken to be 4, 8 and 12. The real values of β are chosen such that $\sum_{i=1}^{p}\hat{\beta}_i^2 = 1$, Kibria and Lukman (2020), Lukman, Dawoud, Kibria, Algamal and Aladeitan (2021)). The MSE was employed to evaluate the estimators' performance.

$$MSE(\beta^*) = \frac{1}{1000}\sum_{j=1}^{100}(\beta_{ij}^* - \beta_i)'(\beta_{ij}^* - \beta_i) \tag{13}$$

where β_{ij}^* is the estimator and β_i is the parameter.

The Tables 1-4 present the MSE values for different estimators: MLE, RR, LE, LT, and TP across varying numbers of predictors (p = 4, 8, 12) and correlation levels (r = 0.90, 0.95, 0.99). MSE measures the average squared difference between estimated and true parameter values, combining both bias and variance; thus, lower MSE indicates better estimator performance. For each fixed number of predictors, as the correlation among predictors increases from 0.90 to 0.99, the MSE for all estimators generally increases. This reflects the known difficulty in parameter estimation under stronger multicollinearity, which inflates variance and estimation error.

For a fixed correlation, increasing the number of predictors from 4 to 12 tends to increase the MSE for all estimators, indicating that higher model complexity challenges estimation accuracy. The MLE consistently exhibits the highest MSE values across all scenarios, indicating it is the least efficient estimator under multicollinearity and high-dimensional settings.
Further, RR improves upon MLE by shrinking coefficients, thus reducing variance and lowering MSE. The LE further reduces MSE compared to RR, demonstrating better bias-variance trade-off. In addition, The LT shows even lower MSE than LE, suggesting enhanced performance in handling multicollinearity.
The TP estimator achieves the lowest MSE values in all cases, indicating it has the best estimation accuracy among the compared methods. For example, at p=12 and r=0.99, MSE drops from 13.288 (MLE) to 4.9361 (TP), a substantial reduction demonstrating the effectiveness of shrinkage and specialized estimators in complex settings.

Table 1: Values of the MSE when n=30

p	r	MLE	RR	LE	LT	TP
4	0.90	6.172	5.819	5.741	5.332	5.2779
	0.95	7.269	6.304	6.03	5.946	5.8918
	0.99	8.312	6.821	6.088	6.047	5.9929
8	0.90	6.289	5.649	5.421	5.225	5.1709
	0.95	8.334	6.48	5.924	5.775	5.7207
	0.99	9.266	7.205	6.431	5.582	5.5279
12	0.90	7.526	5.277	5.102	4.926	4.8712
	0.95	10.749	6.628	5.911	5.022	4.9653
	0.99	13.288	6.955	5.727	4.991	4.9361

Table 2: Values of the MSE when n=50

p	ρ	MLE	RR	LE	LT	TP
4	0.90	6.136	5.783	5.705	5.296	5.2419
	0.95	7.233	6.268	5.994	5.91	5.8558
	0.99	8.276	6.785	6.052	6.011	5.9569
8	0.90	6.253	5.613	5.385	5.189	5.1349
	0.95	8.298	6.444	5.888	5.739	5.6847
	0.99	9.23	7.169	6.395	5.546	5.4919
12	0.90	7.49	5.241	5.066	4.89	4.8352
	0.95	10.713	6.592	5.875	4.986	4.9293
	0.99	13.252	6.919	5.691	4.955	4.9001

Table 3: Values of the MSE when n=100

p	ρ	MLE	RR	LE	LT	TP
4	0.90	6.008	5.655	5.577	5.168	5.1139
	0.95	7.105	6.14	5.866	5.782	5.7278
	0.99	8.148	6.657	5.924	5.883	5.8289
8	0.90	6.125	5.485	5.257	5.061	5.0069
	0.95	8.17	6.316	5.76	5.611	5.5567
	0.99	9.102	7.041	6.267	5.418	5.3639
12	0.90	7.362	5.113	4.938	4.762	4.7072
	0.95	10.585	6.464	5.747	4.858	4.8013
	0.99	13.124	6.791	5.563	4.827	4.7721

Table 4: Values of the MSE when n=200

p	ρ	MLE	RR	LE	LT	TP
4	0.90	5.66	5.307	5.229	4.82	4.7659
	0.95	6.757	5.792	5.518	5.434	5.3798
	0.99	7.8	6.309	5.576	5.535	5.4809
8	0.90	5.777	5.137	4.909	4.713	4.6589
	0.95	7.822	5.968	5.412	5.263	5.2087
	0.99	8.754	6.693	5.919	5.07	5.0159
12	0.90	7.014	4.765	4.59	4.414	4.3592
	0.95	10.237	6.116	5.399	4.51	4.4533
	0.99	12.776	6.443	5.215	4.479	4.4241

5. Conclusion

The BRM is a statistical method that can be used to model count-based data; explained that it is used to model overdispersion data type where the variance is more than the average. The BRM presupposes that the response is Bell-distributed, which is a discrete distribution with probability mass function dependent on the Bell and Lambert functions. This paper is a thorough literature review of biased estimators in bell regression models where there is multicollinearity. The two-parameter estimator is superior to MSE of the MLE, RR, LT, and LE in simulating. Lastly, in the event of multicollinearity in the beta regression model, two parameter estimators ought to be employed.

References

1. Abdul Kareem, Q., & Algamal, Z. (2020). Generalized ridge estimator shrinkage estimation based on particle swarm optimization algorithm. Iraqi Journal of Statistical Sciences, 17(2), 26-35. DOI: 10.33899/iqjoss.2020.0167387
2. Akram, M. N., Amin, M., & Amanullah, M. (2020). Two-parameter estimator for the inverse Gaussian regression model. *Communications in Statistics - Simulation and Computation*, 1-19. doi:10.1080/03610918.2020.1797797
3. Alheety, M. I., & Golam Kibria, B. M. (2013). Modified Liu-Type Estimator Based on (r − k) Class Estimator. *Communications in Statistics - Theory and Methods, 42*(2), 304-319. doi:10.1080/03610926.2011.577552
4. Amin, M., Akram, M. N., & Majid, A. (2021). On the estimation of Bell regression model using ridge estimator. *Communications in Statistics - Simulation and Computation*, 1-14. doi:10.1080/03610918.2020.1870694
5. Asar, Y., & Genç, A. (2017). A New Two-Parameter Estimator for the Poisson Regression Model. *Iranian Journal of Science and Technology, Transactions A: Science*. doi:10.1007/s40995-017-0174-4
6. Bell, E. T. (1934). Exponential numbers. *The American Mathematical Monthly, 41*(7), 411-419. doi:10.1080/03610926.2011.577552.
7. Bell, E. T. (1934). Exponential polynomials. *Annals of Mathematics, 35*(2), 258-277. doi:10.1080/03610918.2020.1797797
8. Castellares, F., Ferrari, S. L. P., & Lemonte, A. J. (2018). On the Bell distribution and its associated regression model for count data. *Applied Mathematical Modelling, 56*, 172-185. doi:10.1016/j.apm.2017.12.014
9. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55-67. doi: 10.1080/00401706.1970.10488634
10. Huang, J., & Yang, H. (2014). A two-parameter estimator in the negative binomial regression model. *Journal of Statistical Computation and Simulation, 84*(1), 124-134. Doi: 10.1080/00949655.2012.696648
11. Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics - Simulation and Computation, 32*(2), 419-435. doi:10.1081/SAC-120017499
12. Kibria, B. M. G., Månsson, K., & 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. doi:10.1080/03610918.2013.796981
13. Kibria, B. M. G., & Saleh, A. K. M. E. (2004). Preliminary test ridge regression estimators with student?s t errors and conflicting test-statistics. *Metrika, 59*(2), 105-124. doi:10.1007/s001840300273
14. Kibria, G., Månsson, K., & Shukur, G. (2012). Performance of some logistic ridge regression estimators. *Computational Economics, 40*(4), 401–414. Doi: 10.1007/s10614-011-9275-x
15. Kurtoğlu, F., & Özkale, M. R. (2016). Liu estimation in generalized linear models: application on gamma distributed response variable. *Statistical Papers, 57*(4), 911-928. doi:10.1007/s00362-016-0814-3

16. Liu, G., & Piantadosi, S. (2016). Ridge estimation in generalized linear models and proportional hazards regressions. *Communications in Statistics - Theory and Methods, 46*(23), 11466-11479. doi:10.1080/03610926.2016.1267767
17. Liu, K. (1993). A new class of biased estimate in linear regression. *Communication in Statistics -Theory and Methods, 22*, 393–402. Doi: 10.1080/03610929308831027
18. Liu, K. (2003). Using Liu-type estimator to combat collinearity. *Communications in Statistics-Theory and Methods, 32*(5), 1009-1020. Doi: 10.1081/STA-120019959
19. Mackinnon, M. J., & Puterman, M. L. (1989). Collinearity in generalized linear models. *Communications in Statistics - Theory and Methods, 18*(9), 3463-3472. doi:10.1080/03610928908830102
20. Majid, A., Amin, M., & Akram, M. N. (2021). On the Liu estimation of Bell regression model in the presence of multicollinearity. *Journal of Statistical Computation and Simulation*, 1-21. doi:10.1080/00949655.2021.1955886
21. Månsson, K., & Shukur, G. (2011). A Poisson ridge regression estimator. *Economic Modelling, 28*(4), 1475-1481. doi:10.1016/j.econmod.2011.02.030
22. Norouzirad, M., & Arashi, M. (2017). Preliminary test and Stein-type shrinkage ridge estimators in robust regression. *Statistical Papers*. doi:10.1007/s00362-017-0899-3
23. Nyquist, H. (1991). Restricted estimation of generalized linear models. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 40*(1), 133-141. doi:10.2307/2347912
24. Othman, F. A. Y. R. A. (2022). Shrinkage Estimators In Inverse Gaussian Regression Model: Subject. Iraqi Journal of Statistical Sciences, 19(1), 46-53. DOI: 10.33899/iqjoss.2022.0174331
25. Segerstedt, B. (1992). On ordinary ridge regression in generalized linear models. *Communications in Statistics - Theory and Methods, 21*(8), 2227-2246. doi:10.1080/03610929208830909
26. Shamany, R., Alobaidi, N. N., & Algamal, Z. Y. (2019). A new two-parameter estimator for the inverse Gaussian regression model with application in chemometrics. *Electronic Journal of Applied Statistical Analysis, 12*(2), 453-464. Doi: 10.1285/i20705948v12n2p453
27. Wu, J. (2014). Modified Restricted Almost Unbiased Liu Estimator in Linear Regression Model. *Communications in Statistics - Simulation and Computation, 45*(2), 689-700. doi:10.1080/03610918.2013.870198
28. Wu, J. (2016). Preliminary test Liu-type estimators based on W, LR, and LM test statistics in a regression model. *Communications in Statistics - Simulation and Computation, 46*(9), 6760-6771. doi:10.1080/03610918.2016.1210170
29. Abdul Kareem, Q., & Algamal, Z. (2020). Generalized ridge estimator shrinkage estimation based on particle swarm optimization algorithm. Iraqi Journal of Statistical Sciences, 17(2), 26-35.

التقديرات المقلصة في انموذج انحدار بيل: مراجعة مقال

حذيفة حازم طه

قسم بحوث العمليات والتقنيات الذكائية، كلية علوم الحاسوب والرياضيات جامعة الموصل، الموصل, العراق

الخلاصة: لقد أصبح نموذج انحدار بيل نموذجًا متعدد الاستخدامات حلّ محل نماذج بيانات العد التقليدية ويساعد على حل مشكلة التشتت المفرط حيث يتجاوز تباين نقاط البيانات المتوسط. ومع ذلك، في الممارسة العملية، غالبًا ما تتأثر مقدرات الامكان الاعظم (MLE) لمعلمات النموذج بتعدد التباينات بين المتغيرات التفسيرية، وتنتج تقديرات غير مستقرة للغاية وتباينات متضخمة. ولمعالجة هذه الصعوبات، تم تطبيق طرق التقدير التي تم تطويرها لتقدير الانكماش، مثل مقدرات التلال أو مقدرات ليو على نموذج انحدار بيل. في استعراض هذا الموضوع، تتناسب التطورات الجديدة في تقديرات انكماش نماذج انحدار بيل مع مناقشة خلفيتها النظرية في المعرفة وعملية التقدير والخصائص التقاربية. تُظهر نتائج دراسات محاكاة مونت كارلو للمحاكاة دائمًا أن مقدرات الانكماش تتفوق على المقدرات المتغيرة من حيث أنها تقلل من متوسط الخطأ المربع والتحيز أكثر من المقدرات المتغيرة خاصة في حالات تعدد التماثلات. وبشكل عام، تعد طرق تقدير الانكماش تحسنًا كبيرًا في نمذجة انحدار بيل التي توفر اليقين والفعالية في تحليل معلومات التعداد المعقدة باستخدام سمات التوزيع الإشكالية.

الكلمات المفتاحية: انحدار ;انحدار الجرس ;مقدر ;انكماش.