Al-Qadisiyah Journal of Pure Science
ISSN(Printed): 1997-2490 ISSN(Online): 2411-3514
DOI: 10.29350/jops
Al-Qadisiyah Journal of Pure Science

Count Data in Normal-Gamma Exponential Prior distribution

Noor H. Hussein
University of AL- Qadisiyah,
College of Education,
Department of Mathematics,
noorhassn564@gmail.com

.Gorgees Sh. Mohammad
University of AL- Qadisiyah,
College of Education,
Department of Mathematics, Iraq
gorgees.alsalamy@qu.edu.iq

Abstract
In recent years, Count Data methods have become very popular in many uses in various disciplines. The Count Data (CD) in the Normal-Gamma Exponential (NGE) prior. Where used the correct numbers.by method Jittering process Which converts a discrete variable to a continuous variable.   And when comparing this method with other methods showing that this method performs reasonably well

Keywords: Count Data, Jittering process, Normal-Gamma Exponential.

1-**Introduction:**

deem linear model:

$$y_i = \boldsymbol{x}_i^T \beta + \epsilon_i, \;\; \epsilon_i \sim N(0, \sigma^2), \tag{1}$$

where $\boldsymbol{\beta}$ is the $p \times 1$ vector of unknown regression variables and ϵ_i is the error with σ^2 the unknown variance. In this work, we assume that this model operates on count data, Bayesian regression of Count data has become a broad and important field. Many studies have been proposed on the subject, such as studying university accreditation and its importance in the evaluation tests that precede registration [5], and economic and social factors and their impact on those infected with tuberculosis in the census tract with this disease [13] As well as the application of these data in biological epidemiology [2]. These various examples illustrate the importance and necessity of developing numerical data in Bayesian regression, which occurred in the development processes in the form of successive stages starting from the 1970s until the modern era [2,7,10], especially when dealing with numerical data [7], where the dependent variable y takes count values, Such as 0,1,2,3,…. While the response variable is normally distributed in cases where the number of observations is very large, in small sample sizes, they become very discrete and sometimes skewed [7] and thus a normal distribution may not be the best choice. There are alternatives to ordinary regression analysis, the most important of which is the Poisson regression model for count-valued responses

$$\boldsymbol{y_i|\mu_i \sim Pois(v_i)} \tag{2}$$

.

and

$$\boldsymbol{v_i \sim exp\{x_i^T \beta\}} \tag{3}$$

The next step is to convert it back to a continuous variable using what is known as the jittering process [8]. This process works by adding an additional random variable $\boldsymbol{b_i}$distributed uniformly to the response variable $\boldsymbol{y_i}$ by

$$\boldsymbol{b_i \sim U(0,1)}, \tag{4}$$

where $\boldsymbol{U}$ is the uniform distribution. The last step, is to use the logarithm function to achieve the desired continuous variable

$$\boldsymbol{y_i^* = ln\{y_i + b_i\}}$$

In this work, we will analyze the Bayesian regression framework in the present of count data with a normal scale-mixture (exponential- gamma)

$$\beta_j|\sigma^2, v_i^2 \sim \mathcal{N}(0, \sigma^2 v_i^2),\ i = 1, \dots, p \tag{5}$$

$$v_i^2 \sim \mathcal{G}(\alpha, \delta_i^2), \tag{6}$$

$$\delta_i^2 \sim Exp(\lambda).$$

2- The normal-mixture (exponential- gamma) prior

2-1 The compound exponential- gamma prior can be written as

$$\pi(x) = \int_0^\infty Gamma(x; \alpha, r) Exponential(r; \lambda) dr$$

$$= \frac{\alpha\, \lambda\, x^{\alpha-1}}{(x+\lambda)^{\alpha+1}} \tag{7}$$

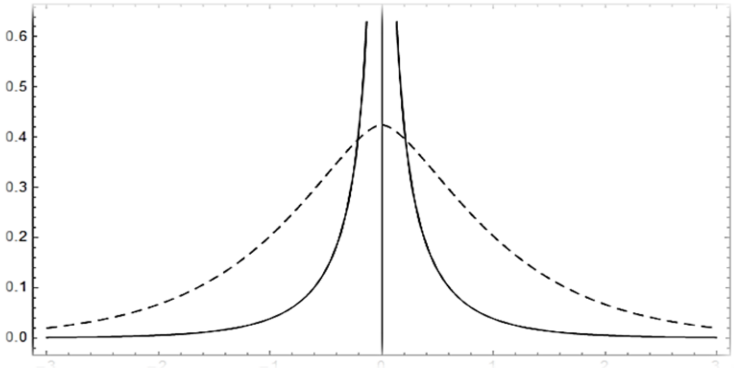


Figure 2. An illustration of the F(x) (7) with the solid line ($\alpha = 0.2, \lambda = 0.65$) and dashed line ($\alpha = 2.5, \lambda = 1.1$)

2-2- Posterior Inference:

Since

$$\beta|X, y, \zeta_i^2, \eta_i^2, \sigma^2 \sim \mathcal{N}\left(\mu_\beta, V^{-1}\sigma^2\right).$$

$$\sigma^2|X, y, \zeta_i^2, \eta_i^2 \sim \mathcal{IG}\left(\frac{n+p}{2}, \frac{(y-X\beta)^T(y-X\beta)+\beta^T \boldsymbol{M}^{-1}\beta}{2}\right). \qquad (8)$$

$$P(\beta\big|X, y, \zeta_i^2, \eta_i^2, \sigma^2) \propto P(y|\beta, \sigma^2)\,\pi(\beta),$$

$$\propto exp\left\{-\frac{1}{2\sigma^2}(y^* - X\beta)^T(y^* - X\beta) - \frac{1}{2\sigma^2}\beta^T \boldsymbol{M}^{-1}\beta\right\},$$

$$\propto exp\left\{-\frac{1}{2\sigma^2}(-2y^T X\beta + \beta^T X^T X\beta + \beta^T \boldsymbol{M}^{-1}\beta)\right\},$$

$$= exp\left\{-\frac{V}{2\sigma^2}\left(-2{\mu_\beta^*}^T\beta + \beta^T\beta\right)\right\}, \qquad (9)$$

with $\boldsymbol{\mu_\beta^*} = \boldsymbol{V^{-1}X^T y^*}$ and thus we again obtain the normal distribution

$$\beta|X, y^*, \zeta_i^2, \eta_i^2, \sigma^2 \sim \mathcal{N}\left(\mu_\beta^*, V^{-1}\sigma^2\right) \qquad (10)$$

and

$$P(\sigma^2 \big| X, y, \zeta_i^2, \eta_i^2) \propto P(y^* \big| \beta, \zeta_i^2, \eta_i^2, \sigma^2)\pi(\beta_i \big| \zeta_i^2, \eta_i^2, \sigma^2)\pi(\sigma^2)$$
$$\propto (\sigma^2)^{-\frac{n}{2}} exp\left\{-\frac{(y^* - X\beta)^T(y^* - X\beta)}{2\sigma^2}\right\}$$
$$\times (\sigma^2)^{-\frac{p}{2}} exp\left\{-\frac{\beta^T \boldsymbol{M}^{-1}\beta}{2\sigma^2}\right\} \times (\sigma^2)^{-1} \tag{11}$$
$$\propto (\sigma^2)^{-\left(\frac{n+p}{2}\right)-1}$$
$$\times\, exp\left\{-\frac{(y^* - X\beta)^T(y^* - X\beta) + \beta^T \boldsymbol{M}^{-1}\beta}{2\sigma^2}\right\}$$

resulting in the inverse-gamma distribution

$$\sigma^2 | X, y^*, \zeta_i^2, \eta_i^2 \sim \mathcal{IG}\left(\frac{n+p}{2}, \frac{(y^* - X\beta)^T(y^* - X\beta) + \beta^T \boldsymbol{M}^{-1}\beta}{2}\right) \tag{12}$$

As for count MCEM algorithm, we get

$$\psi(\alpha) = \frac{1}{p}\sum_{i=1}^{p} \mathbb{E}_{\alpha^{k-1}}[log\ (\zeta_i^2)\ |y^*] \tag{13}$$

and

$$\lambda^{-1} = \frac{1}{p}\sum_{i=1}^{p} \mathbb{E}_{\lambda^{k-1}}[log\ (\eta_i^{-2})\ |y^*] \tag{14}$$

3- Simulation Studies

3-1 simulation 1

In this section we will use simulated data to test how our model performs with respect to other models such as the least absolute shrinkage and selection operator(LASSO).[**11**],the elastic net penalty (ENet).[**8**],the horseshoe estimator(HS).[**2**],In addition,we will use three parameters of comparison such as the mean squared error(MSE),false positive rate (FPR), falls negative rate (FNR)we will use data size of 1500 with 2000 burn-ins For this example, we will set $\boldsymbol{\beta} = (\mathbf{5}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{3}, \mathbf{0}, \mathbf{0}), \boldsymbol{\sigma^2} \in \{\mathbf{1}, \mathbf{4}, \mathbf{9}\}$ and $\boldsymbol{n} = (\mathbf{100})$. This takes the case considers the spare case with only two active covariates out of eight. The results are summarized in Table. This shows that none of the methods considered performed better than our model. Moreover, we can see that fitting the prior exponential kamma gives the smallest MSE compared to all other models foot. The hyperparameters are updated every 100 iterations. As well as the proposed model It is very well designed in terms of the FPRs and FNRs needed to choose the best model for variable selection.

Table 1. Results for model with count data and few active covariates.

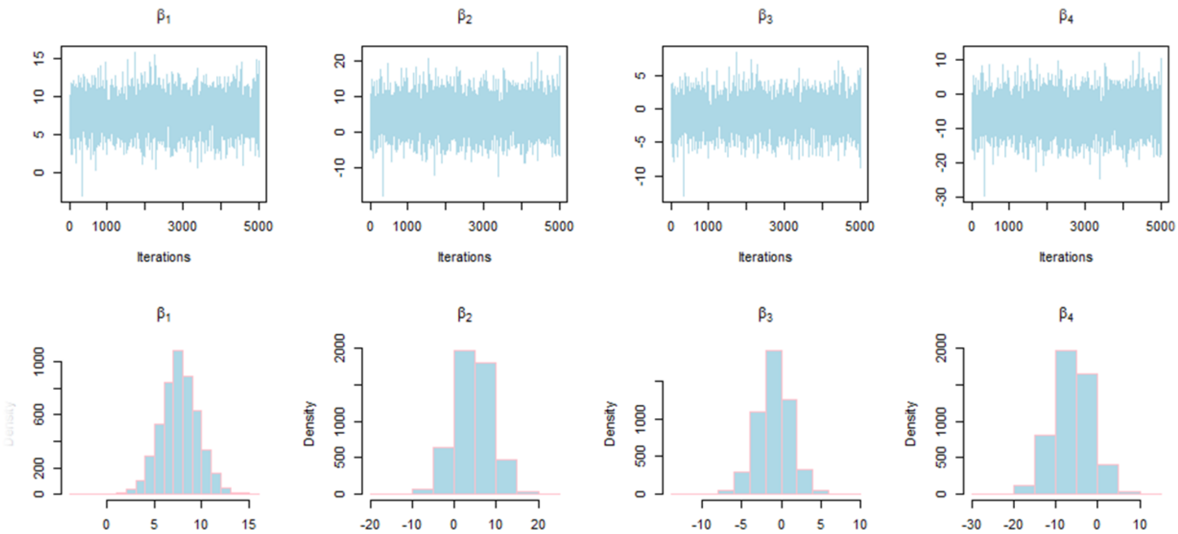
Methods	n	**MSE (sd)**	**FPR (FPRsd)**	**FNR (FNRsd)**
NGE	100	0.0075 (0.0090)	0.1000 (0.3162)	0.0000 (0.0000)
HS	100	0.0091 (0.0093)	0.1000 (0.3162)	0.0000 (0.0000)
Lasso	100	0.0296 (0.0214)	0.0000 (0.0000)	0.0000 (0.0000)
aLasso	100	0.0082 (0.0112)	0.0000 (0.0000)	0.0000 (0.0000)
ENet	100	0.0461 (0.0346)	0.0000 (0.0000)	0.0000 (0.0000)

3-2 Simulation 2

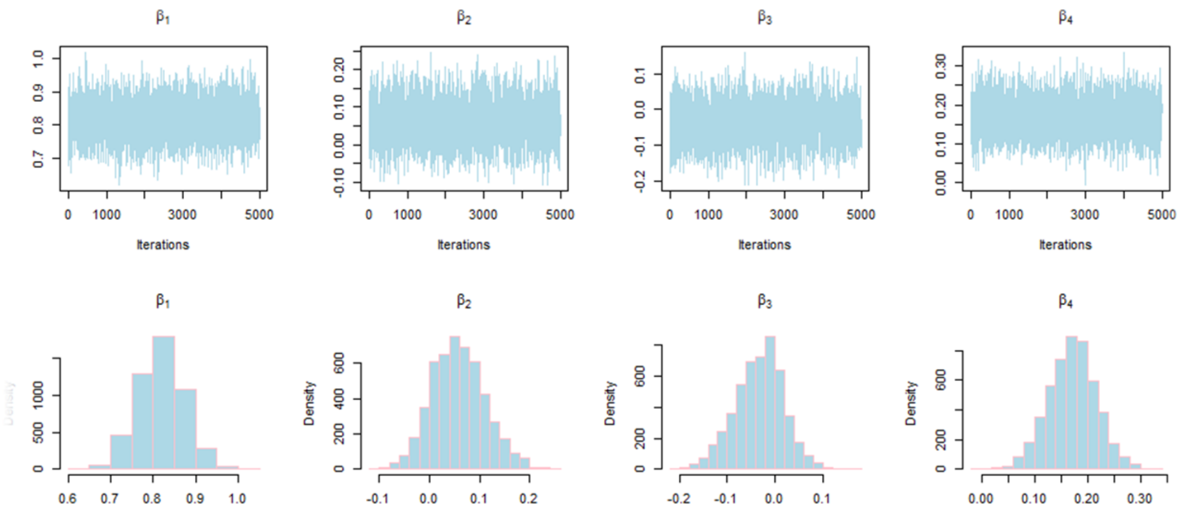
To study the proposed model further and obtain a more precise study of its behavior with different types of data, similar size covariates as the simulation above were used. While reducing the sparsity of the model by changing the parameter values $\beta = (3,0,4,1,7,0,1,2)$with a size of n = 100 repeated simulations with 10,000 iterations. The results appear in Table 2. The results show that with increasing Sparsity, The model gives better performance at higher values of, N, and conversely, models with smaller values of,N are better candidates for data with lower sparsity.

Table 2. Results for model with count data and most covariates are active.

Methods	n	**MSE (sd)**	**FPR (FPRsd)**	**FNR (FNRsd)**
NGE	100	0.0016 (0.0006)	0.2000 (0.4216)	0.0000 (0.0000)
HS	100	0.0014 (0.0007)	0.0000 (0.0000)	0.0000 (0.0000)
Lasso	100	0.0039 (0.0059)	0.0000 (0.0000)	0.0000 (0.0000)
aLasso	100	0.0032 (0.0040)	0.0000 (0.0000)	0.0000 (0.0000)
ENet	100	0.0040 (0.0064)	0.0000 (0.0000)	0.0000 (0.0000)



$\boldsymbol{Figure}(\,2).Tracking\ plots\ and\ \ histograms\ for\ simulation\ 1$



$\boldsymbol{Figure}(\,3).Tracking\ plots\ and\ \ histograms\ for\ simulation\ 2$

4- **Conclusions**

The results show that the proposed method of imputation of data using normal exponential gamma before scale mixture under normal mixture, performs well compared to other existing models such as Bayesian cord, Bayesian adaptive cord, and Bayesian elastic network. We hope that this matter will be studied more extensively in the future using different types of data such as controlled data, quantitative data, etc

References

1- Carvalho, C. M., N. G. Polson, and J. G. Scott (2010). The horseshoe estimator for sparse signals. Biometrika 97 (2), 465–480

2- Gourieroux, C., Monfort, A., & Trognon, A. (1984). Pseudo maximum likelihood methods: Applications to Poisson models. Econometrica: Journal of the Econometric Society, 701-720.

3- Jantre, S. (2023). Bayesian quantile regression for longitudinal count data. Journal of Statistical Computation and Simulation, 93(1), 103-127

4- Jorgenson, D. W. (1961). Multiple regression analysis of a Poisson process. Journal of the American Statistical Association, 56(294), 235-245. [10] Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. Econometrica: journal of the Econometric Society, 33-50.

5- Grilli, L., Rampichini, C., & Varriale, R. (2016). Statistical modelling of gained university credits to evaluate the role of pre-enrolment assessment tests: An approach based on quantile regression for counts. Statistical Modelling, 16(1), 47-66

6- Lee, D., & Neocleous, T. (2010). Bayesian quantile regression for count data with application to environmental epidemiology. Journal of the Royal Statistical Society: Series C (Applied Statistics), 59(5), 905-920.

7- Li, Q., N. Lin, et al. (2010). The Bayesian elastic net. Bayesian Analysis 5 (1), 151–170.

8- Machado, J. A. F., & Silva, J. S. (2005). Quantiles for counts. Journal of the American Statistical Association, 100(472), 1226-1237.

9- McCullagh, P., and Nelder, J. A. (1989), Generalized Linear Models (2nd ed.), London: Chapman & Hall.

10- Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. Journal of the Royal Statistical Society: Series A (General), 135(3), 370-384.

11- Sarvi, F., Momenian, S., Khodadost, M., Pahlavanzadeh, B., Nasehi, M., & Sekhavati, E. (2016). The examination of relationship between socioeconomic factors and number of tuberculosis using quantile regression model for count data in Iran 2010-2011. Medical Journal of the Islamic Republic of Iran, 30, 399.
12- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society, Series B 58, 267–288.
13- Wu, H., Gao, L., & Zhang, Z. (2014). Analysis of crash data using quantile regression for counts. Journal of Transportation Engineering, 140(4), 04013025.