Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 37–45. Research Paper

Bayes factor: hypothesis testing in restricted Stein-rule regression analysis with autocorrelated error

Oloyede Isiaka*

Department of Statistics, University of Ilorin, Ilorin, Nigeria
Email(s): oloyede.i@unilorin.edu.ng

Abstract. Regression is a generic tool used to make inferences so as to draw logical conclusions and judgments about a particular problem. It has been widely used in engineering, sciences, education, technology, and social sciences for a long time. Breaking down of the underlying estimation mar the judgment and decision of the study. Chiefly among the error structure that can lead to the drawback of the inferences of regression is the autocorrelation of error term of both AR and MA processes. Restricted Stein-rule regression analysis was used with data injected with autocorrelated error; H_1 was modeled with autocorrelated error whereas H_0 was modeled without, alternative approach in Bayes factor of $AR(1)$ and $MA(1)$ processes were introduced and compared with Bayesian information criterion approach in both negative and positive ρ (symmetrical). The choice of the Bayes factor (Bf) is due to the probabilistic nature of Bayesian inference, which over the years has been performed better than the classical approach. The datasets with five covariates was set at 25 to capture the error structure and project the property of a small sample. The study, therefore, concluded that Bayesian inference being probabilistic about the uncertainty of the parameters should be adopted to verify the presence or absence of autocorrelated error in the data before estimation.

Keywords: Autocorrelated Error; Bayesian; Bayes factor; Hypothesis; Restricted; Stein-Rule.

1 Introduction

Han and Carlin (2001) attempted to compare several methods of estimations under the categories of simple regression, hierarchical longitudinal model, and binary data latent. They thought that the joint model parameters search space technique performs well but is only complicated in terms

*Corresponding author

Received: 31 October 2023 / Accepted: 7 March 2024
DOI: 10.22067/smps.2024.45179

 https://smps.um.ac.ir

of computing algorithm whereas marginal likelihood approach is of less difficult. They therefore suggested from their study that the marginal likelihood approach is appropriate for researchers with model choice setting while the joint space method can be used for comparing models of varying dimensions. It was learned from the literature that the origin of the Bayesian hypothesis testing was credited to Jeffreys (1961), and ordinarily it was referred to as a significant test.

Kass and Raftery (1995) reviewed the work of Jeffreys (1961) and added that the Bayesian technique to hypothesis testing meets the test of time with moderate computation technique. They found out that the Bayes factor is sensitive to the assumption of the parametric model and choice of prior. Giampaoli et al. (2015) examined the Bayes factor in a restricted simple linear regression and pointed out that their approach considered restricted parameter space to be more informative than unrestricted parameter space, which permits evidence for null hypothesis. Bayes factor provided a means through which two competing hypothesis may be compared; see Johnson et al. (2023). The Bayes factor $B10$ is used to compute the probability of the observed data under the alternative hypothesis H_1 versus the null hypothesis H_0, which is expressed as

$$B_{10} = \frac{H_1}{H_0} = \frac{\int P(Y|\beta)\,\pi(\beta|H_1)\,d\beta}{\int P(Y|\beta)\,\pi(\beta|H_0)\,d\beta}. \tag{1}$$

Jarosz and Wileyr (2014) in their presentation, showcased the similarities between the Bayes factor and p-value. They believe that bf should have similar information as p-value. The only difference is that bf allowed the researcher to give evidence in both the alternative and null hypothesis unlike p-value that will solely give evidence about null hypothesis. Roozbeh and Hamzah (2020) pointed out that the restriction of parameter in partial linear regression can overcome the problem of multicollinearity. They believed that restricted estimator outperformed traditional estimator. They added that restriction can be hypothesis that may have to be tested.

Over the years, many studies have been carried out in Bayesian inference for linear and nonlinear models on the premise of model averaging, parameter estimation, and hypothesis testing, but most importantly restricted stem-rule are not well studied in the Bayesian framework. Against this backdrop, this paper seems to examine nonspherical disturbance (autocorrelated error) in the Bayesian restricted stem-rule paradigm.

2 Materials and methods

Let $y = X\beta + u$ be the linear regression model, where y is an $n \times 1$ set of observations on the regressand, X is a set of $n \times p$ full column rank of regressors, and β is $p \times 1$ vectors of unknown parameters while u is an $n \times 1$ vectors of disturbance error not necessarily well behave, which is characterized by autocorrelated errors. Let there be m linearly independent restriction that constrains the regression coefficients such that

$$r = R\beta, \tag{2}$$

where r is an $m\times 1$ vector and R is an $m\times p$ matrix of rank $m<p$ Chaturvedi et al. (2001). The $AR(1)$ and $MA(1)$ processes are expressed as

$$\widehat{\Omega}_{AR}=\frac{\sigma_u^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}$$

The $MA(1)$ with autocorrelated error is expressed as

$$\widehat{\Omega}_{MA}=\sigma_u^2\begin{bmatrix} (1+\phi^2) & \phi & \cdot & \cdot & \cdots \\ \phi & (1+\phi^2) & \phi & \cdot & \cdots \\ \cdot & \phi & (1+\phi^2) & \phi & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \cdot & \cdot & \cdot & \cdots & (1+\phi^2) \end{bmatrix},$$

where ϕ ranges between -1 and 1. The generalized restricted least squares estimates are obtained as follows: Adopting the criterion of minimizing the sum of squares $(y-X\beta)'\widehat{\Omega}(y-X\beta)$ subject to the condition that $R\beta = r$. This leads to the Lagrangian function

$$\widehat{\beta}_R = \widehat{\beta}+\left(X\widehat{\Omega}X\right)^{-1}\left[R\left(X\widehat{\Omega}X\right)^{-1}R\right]^{-1}R'\left(r- R\widehat{\beta}\right), \tag{3}$$

$$\widehat{\beta}_R=\widehat{\beta}+\left(X'\widehat{\Omega}X\right)^{-1}R'\left[R\left(X'\widehat{\Omega}X\right)^{-1}R'\right]^{-1}\left(r-R\widehat{\beta}\right). \tag{4}$$

Thus β_R is a constrained estimates, following Chaturvedi et al. (2001), the restricted stein-rule version of intertwined disturbance errors can be expressed as

$$\widehat{\beta}_S=\left[1-\frac{a}{n}\frac{\left(y-X\widehat{\beta}\right)'\widehat{\Omega}\left(y-X\widehat{\beta}\right)}{\widehat{\beta}X\widehat{\Omega}X\widehat{\beta}}\right]\widehat{\beta}. \tag{5}$$

We have

$$\widehat{\beta}_{RS}=\widehat{\beta}_S+\left(X'\widehat{\Omega}X\right)^{-1}R'\left[R\left(X'\widehat{\Omega}X\right)^{-1}R'\right]^{-1}\left(r-R\widehat{\beta}_S\right), \tag{6}$$

where $\widehat{\Omega}=\Omega(b)$, in which b is a consistent and efficient estimator of β; thus $b=\left(X'\widehat{\Omega}X\right)^{-1}\left(X'\widehat{\Omega}y\right)$. Following Chaturvedi and Shukla (1990), we have modified the Stein rule estimator of beta.

3 Bayesian restricted least squares estimator

The restricted posterior density of restricted β and σ is obtained by marginalizing the conjugate of normal-inverse gamma and restricted likelihood since both are of the same family of distribution Oloyede (2023). The linear model is expressed as

$$y = X\beta_R + u. \tag{7}$$

The likelihood function of β, X, and y, where $\theta = (\beta, \lambda)$ given sample vectors X_1, $X_2 = (1, 2, \ldots, n)'$ and $y = (y_1,\ y_2, \ldots, y_n)'$ and incorporating restricted β_R, is expressed as

$$L\left(\beta_R, \sigma^2 | X, y\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\left(y - X\beta_R\right)' \widehat{\Omega}^{-1}\left(y - X\beta_R\right)\right] \to H_1, \tag{8}$$

$$L\left(\beta_R, \sigma^2 | X, y\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\left(y - X\beta_R\right)'\left(y - X\beta_R\right)\right] \to H_0. \tag{9}$$

Note that normal-inverse gamma priors are conjugate priors and selected because the prior and posterior densities are of the same family of distributions; see Oloyede (2023). Moreover,

$$p\left(\beta_R|\sigma^2\right) p\left(\sigma^2\right) = (2\pi)^{-\frac{k}{2}} \left|\widehat{\Omega}\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\beta_R - \mathbb{B}\right)' \widehat{\Omega}^{-1}\left(\beta - \mathbb{B}\right)\right] \times \sigma^{-(a-k)} \exp\left[-\frac{b}{\sigma^2}\right], \tag{10}$$

$$p\left(\beta_R|\sigma^2\right) p\left(\sigma^2\right) = \sigma^{-(a-k)} \exp\left[-\frac{1}{2\sigma^2}\left(\beta_R - \mathbb{B}\right)' \widehat{\Omega}^{-1}\left(\beta_R - \mathbb{B}\right) + 2b\right], \tag{11}$$

where $X$is an $n \times k$ matrix,
β_R is unknown parameter,
B is a prior mean vector of β (true value),
σ^2 is a prior variance for β,
$\widehat{\sigma}^2 = \frac{(y - X\beta_R)\widehat{\Omega}(y - X\beta_R)}{n-k}$,
$a - k$ is first hyper-parameter,
b is second hyper-parameter.

Since σ^2 is known, normal-inverse gamma conjugate prior is adopted. For details of the posterior, the reader is advised to see Oloyede (2023).

Posterior density function is

$$\widehat{\beta}_R \sim MVN\left(\widehat{\beta}_R\ ,\ \widehat{\sigma}^2\left(X'\widehat{\Omega}X\right)^{-1}\left[1 - \left(X'\widehat{\Omega}X\right)^{-1} R'\left[R\left(X'\widehat{\Omega}X\right)^{-1} R'\right]^{-1} R\ \right]\ \right), \tag{12}$$

$$\breve{\beta}_{RS} \sim MVN\left(\widehat{\beta}_{RS}\ ,\ \widehat{\sigma}^2\left(X'\widehat{\Omega}X\right)^{-1}\left[1 - \left(X'\widehat{\Omega}X\right)^{-1} R'\left[R\left(X'\widehat{\Omega}X\right)^{-1} R'\right]^{-1} R\ \right]\ \right), \tag{13}$$

for restricted the Stein-rule estimator.

4 Bayesian hypothesis testing

Let $M_i = X\beta + u*$ be the regression model with autocorrelated error, setting up dual distinct model M_1 and M_0, where 1 and 0 represent the model with autocorrelated error and without autocorrelated error, respectively. The prior belief of parameter estimates of both models is the same. Let D be the set of observed datasets corrupted with autocorrelated error of $AR(1)$ and $MA(1)$ independently. Thus, the posterior of M_1 and M_0 can be expressed as follows:

$$P(\beta|\mathscr{D},\ M_0\ ) \propto P(\beta|\ M_0) * P(D|\beta, M_0\ ) \rightarrow H_0, \tag{14}$$

$$P(\beta|\mathscr{D},\ M_1\ ) \propto P(\beta|\ M_1) * P(D|\beta, M_1\ ) \rightarrow H_1. \tag{15}$$

Bayes factor is therefore expressed as

$$BF_{10} = \frac{Posterior\ model\ odds}{prior\ model\ odds}, \tag{16}$$

$$\frac{P(D|\beta, M_1\ )}{P(D|\beta, M_0\ )} = \frac{P(M_1|D\ )}{P(M_0|D\ )} \div \frac{P(M_1)}{P(M_0)}, \tag{17}$$

$$\frac{P(M_1|D\ )}{P(M_0|D\ )} = \frac{P(M_1)}{P(M_0)} \times \frac{P(D|\beta, M_1\ )}{P(D|\beta, M_0\ )}. \tag{18}$$

5 Data generation processes and simulation experiment

The Markov Chain Monte Carlo simulation algorithm was adopted to examine the small sample properties of the family of restricted least squares estimators with autocorrelated error in Bayesian frameworks. Data were generated based on the following parameters of the model: $P = 6, n = 25,\ y_t = X_t\beta + u_t,\ t = 1, \ldots, 25$, where ε_t assumed to be generated by the $AR(1)$ process $u_t = \rho u_{t-1} + \varepsilon_t$ or the $MA(1)$ process $u_t = \varepsilon_t - \rho\varepsilon_{t-1},\ _t\ N(0,1)$. Moreover, $\hat{\beta}$ was set as $(1.2, 2, 0.8, 0.3, 2.1, 1.1)$ while seed was set at 1234. Also, 10000 iterations were set for Bayesian Monte Carlo simulation, and ρ was set at $-0.8, -0.5, -0.3, 0, 0.3, 0.5, 0.8$ for both $AR(1)$ process and $MA(1)$ process. The restriction of parameters was set as

$$R = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{19}$$

$$r = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \tag{20}$$

where $\beta_1 - \beta_3 = 0,\ \beta_2 + \beta_4 = 1$ and $\beta_5 = 0$. All computations were carried out using Statistical software R-Core (2022). The datasets class contained the posterior sample for the model parameters.

6 Hypothesis metrics

Bayesian Information Criterion (BIC) and Quadratic weight loss and risk function approaches that incorporated autocorrelated error structures were used to evaluate the Bayes factor of both the Bayes estimate and posterior mean. Let $\left(\widehat{\beta}_R - \beta\right) = \left(\widehat{\beta}_R - \beta\right) Q \left(\widehat{\beta}_R - \beta\right)$ be a quadratic or square error loss function, where $\hat{\beta}_R$ is an estimator of β and Q is the $\sum_{i=1}^{\beta_R} \widehat{\beta}_R$ weight of loss function. For obtaining the Bayes factor for $\hat{\beta}_R$ and $hat\beta_{RS}$, ratio of quadratic loss function of both H_1 and H_0 was computed as well as ratio of BIC for both H_1 and H_0.

7 Data analysis and discussion

Both null and alternative hypotheses were subjected to the same datasets having autocorrelation error. Two types of autocorrelation error $AR(1)$ and $MA(1)$ processes were examined. The null hypothesis was modeled in a natural setting without being generalized, but the alternative hypothesis was modeled in a generalized pattern by incorporating autocorrelation error structure into the model. Bayesian restricted least squares and Bayesian restricted stem-rule were the predominant estimators adopted in the study. The study looked into the evidence that alternative hypothesis have over null hypothesis due to the presence of autocorrelated error in the datasets in Bayesian paradigm. The threshold of 1 is the decision value.

Table 1: Bayes factor with loss and risk function approach with $AR(1)$ process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	1.001761	1.001954	1.000197	0.999665
-0.6	0.947783	0.947851	0.94699	0.946795
-0.3	0.964658	0.964679	0.964214	0.964158
0	1.000008	1.000014	0.999745	0.999731
0.3	1.039884	1.039891	1.03971	1.039685
0.6	1.073562	1.07359	1.073367	1.073277
0.8	1.094908	1.094954	1.094644	1.094493

Table 2: Bayes factor with BIC approach with $AR(1)$ process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	1.172955243	1.188872	1.010712	1.010932
-0.6	2.273047888	2.286148	1.021469	1.021607
-0.3	2.005146615	2.008686	1.013486	1.01364
0	0.998911484	0.999406	1.000062	1.000056
0.3	0.380557759	0.380816	1.035379	1.035065
0.6	0.18084545	0.181293	1.253861	1.253553
0.8	0.33050558	0.332143	2.335213	2.336675

Tables 1 and 2 provide information about the outcome of study about the Bayes factor with autocorrelated error of autoregressive process of order 1. Table 1 shows that the autocorrelated error of magnitude -0.8, 0 to 0.8 provided evidence in favour of the alternative hypothesis H_1 for both the Bayes estimates and posterior mean considering Bayesian restricted least squares whereas the magnitude of -0.6 and -0.3 provided evidence in support of accepting null hypothesis H_0. In the same vein, the Bayes factor for Bayesian restricted Stein-rule provided information in favor of alternative hypothesis considering the magnitude of 0.3 to 0.8 autocorrelated errors. In Table 2 where BIC approach was adopted to compute the Bayes factor, the magnitudes of -0.8 through -0.3 portend evidences in favour of alternative hypothesis whereas other magnitudes were otherwise. This is in Bayesian restricted least squares for both Bayes estimates and posterior mean. Interestingly, Bayesian restricted stein rule has Bayes factor that portends evidence in favour of alternative hypothesis for both Bayes estimates and posterior mean for all the magnitudes.

Table 3: Bayes factor with Loss and risk function with MA(1) Process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	1.052104	1.052086	1.047152	1.047216
-0.6	1.002316	1.002319	1.000276	1.000283
-0.3	0.98147	0.981478	0.980911	0.980893
0	1.000008	1.000014	0.999745	0.999731
0.3	1.058752	1.05874	1.058577	1.058607
0.6	1.069787	1.069688	1.069916	1.07014
0.8	0.799912	0.799774	0.799965	0.800259

Table 4: Bayes factor with Bayesian Information Criteria with MA(1) Process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	0.428503544	0.428626	1.294917	1.294175
-0.6	1.728149432	1.729083	1.137145	1.137505
-0.3	1.770699531	1.771931	1.022389	1.022589
0	0.998911484	0.999406	1.000062	1.000056
0.3	0.373687701	0.373339	1.023952	1.023549
0.6	0.792633093	0.786735	1.050708	1.051685
0.8	41.83140879	41.63028	1.511668	1.523933

Tables 3 and 4 show Bayes factor Loss and risk function with $MA(1)$ Process . Tables 3 and 4 provide information about the outcome of the study about the Bayes factor with autocorrelated error of moving average process of order 1. Table 3 shows that autocorrelated error of magnitude -0.8,-0.6, 0, 0.3, and 0.6 provided evidence in favour of alternative hypothesis H_1 for both Bayes estimates and posterior mean considering Bayesian restricted least squares whereas other magnitudes provided evidence in support of accepting null hypothesis H_0. In the same vein, the

Bayes factor for Bayesian restricted the Stein-rule provided information in favour of alternative hypothesis considering the magnitude of -0.8, -0.6, 0.3, and 0.6 autocorrelated errors. In Table 4 where the BIC approach was adopted to compute the Bayes factor, the magnitudes of -0.6, -0.3, and 0.8 portend evidences in favour of alternative hypothesis whereas other magnitudes were otherwise. This is in Bayesian restricted least squares for both Bayes estimates and posterior mean. Interestingly, the Bayesian restricted Stein rule has Bayes factors that portend evidence in favour of alternative hypothesis for both Bayes estimates and posterior mean for all the magnitudes.

8 Conclusion

This study introduced Bayes factor as an alternative to probability value, which has an age long drawback in science, technology, education, and so on, with the view of having probability, not only in the value of data, but also in the uncertainty of unknown parameters. The study established and compared the BIC approach and loss/risk function approach in determining the Bayes factor. The small sample behaviour of regression parameter with uncertainty was studied via the violation of non-serial correlation of error terms in a restricted least squares and restricted stein-rule least squares. The study allays fears in the minds of researchers in engineering, science, technology, education, and so on, who have over the years unsatisfied with p-value or using it wrongly. The Bayesian test hypothesis now provides alternative means through which they can verify the presence/absence of autocorrelated errors in their data.

References

Chaturvedi, A. and Shukla, G. (1990). Stein rule estimation in linear model with non-scalar error covariance matrix. *Sankhyä: The Indian Journal of Statistics*, **Series B**, 293–304.

Chaturvedi, A., Wan, A. T. K. and Singh, S. P. (2001). Stein-rule restricted estimator in a linear regression model with nonspherical disturbances. *Communications in Statistics - Theory and Methods*, **30**, 55–68.

Giampaoli, V., Pereira, C., Bolfarine, H., Singer, J. (2015). *Bayes Factors for Comparison of Restricted Simple Linear Regression Coefficients*. In: Polpo, A., Louzada, F., Rifo, L., Stern, J., Lauretto, M. (eds) Interdisciplinary Bayesian Statistics. Springer Proceedings in Mathematics & Statistics, 118. Springer, Cham.

Han, C. and Carlin, B. P. (2001). Markov chain Monte Carlo methods for computing Bayes factors: A comparative review. *Journal of the American Statistical Association,* **96**, 1122–1132.

Jarosz, A. F. and Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving,* **7**, 2.

Jeffreys, H. (1961). *Theory of Probability.* (3rd ed) Oxford University Press, Oxford.

Johnson, V. E., Pramanik, S. and Shudde, R., (2023). Bayes factor functions for reporting outcomes of hypothesis tests. *Statistics Psychological And Cognitive Sciences*, **120**, e2217331120.

Kass, R. E. and Raftery, A. E., (1995). Bayes factors. *Journal of the American Statistical Association,* **90**, 773–795.

Oloyede, I. (2023). Bayesian linear restricted least squares estimators. *Journal of Indian Society Probability and Statistics*, **24** 65–74.

R-Core Team (2022). R: A language and environment for statistical computing, *R Foundation for Statistical Computing*, Vienna, Austria.

Roozbeh, M., and Hamzah, N. A. (2020). Uncertain stochastic ridge estimation in partially linear regression models with elliptically distributed errors. *Statistics*, **54**, 494–523.