DERIVATION OF PRIOR AND POSTERIOR DISTRIBUTIONS OF REGRESSION PARAMETERS BASED ON ZELLNER'S G-PRIORS

Supported by tetfund

Rotimi K. Ogundeji[1]* and Ismaila A. Adeleke[2]
[1]Department of Mathematics, University of Lagos, Lagos, Nigeria
[2]Department of Actuarial Science and Insurance, University of Lagos, Lagos, Nigeria
*Corresponding author: rogundeji@unilag.edu.ng

Received: June 10, 2020 **Accepted:** October 03, 2020

Abstract: In Bayesian analysis, posterior distribution summarises what we know about uncertain quantities. It is a combination of the prior distribution and the likelihood function. Inference proceeds from the posterior distribution where all required posterior quantities were generated analytically. Informative prior distributions related to a natural conjugate prior specification are studied under a limited choice of a single scalar hyper parameter called g-prior which corresponds to the degree of prior uncertainty on regression coefficients. This research identified a set of nine candidate default priors (called Zellner's g-priors) prominent in literature and applicable in Bayesian model averaging (BMA). The methods adopted are theoretical and literature based and can be applied to derive the prior and posterior distributions of the regression parameters of multiple regression models. Results obtained include the respective prior distributions and posterior distributions based on the set of g-prior structures prominent in Bayesian Model Averaging (BMA).

Keywords: Bayesian model averaging, multiple regression models, prior elicitation

Introduction

Research on Bayesian methodology and applications has progressed remarkably in the past few decades and issues of the choice of prior distribution and ultimately the determination of the posterior distribution have been quite delicate in data analysis. Procedures for assessing informative prior and posterior distributions for the parameters in Bayesian regression models have been put forward by Zellner (1983, 1986); Agliari & Parisetti (1988); Raftery *et al.* (1997); Fernandez *et al.* (2001a) and Eicher *et al.* (2011).

Prior and posterior distributions play very crucial roles in Bayesian probability Theory as it is attractive to have conditional distributions that have a closed form under sampling; (Okafor, 1999 and Rossi *et al.*, 2005). Zellner (1983, 1986) proposed a procedure for evaluating a conjugate prior distribution referred to as Zellner's informative g-prior, or simply g-prior. The g-prior has been vastly used in Bayesian analysis in multiple regression models, due to the verity that analytical results are more readily available, better computational efficiency and its simple interpretation (Davison, 2008; Zhang *et al*., 2008; Rafetry *et al*., 2010 and Ogundeji *et al*., 2018). In linear regression model analysis in which g-prior is used, it has been noted that the choice of a scalar hyperparameter g is crucial for the behaviour of Bayesian model averaging (BMA) procedures. The use of Bayesian model averaging provides a natural solution to model uncertainty that leads to better predictions than simply selecting and using one model (Clyde & George, 2004). The Zellner's g-prior structure has proven universally popular in Bayesian model averaging, since it leads to simple closed form expressions of posterior quantities and because it reduces prior elicitation to the choice of a single hyperparameter g. The approach to prior specification in multiple regression models presented here draws inspiration from the work of Feldkircher *et al.* (2012), Fouskakis & Ntzowfras (2013), Hanson *et al.* (2014) and Li & Clyde (2015).

The aim of this study is to derive both prior distributions and posterior distributions of the regression parameters in Bayesian model averaging using the respective identified g-prior from literature and obtain posterior quantities for inferences. This research identified a set of nine candidate default priors (Zellner's informative g-prior that is based on a sample of n observations and k regression coefficients of independent variables) advocated in literature (Eicher *et al.*, 2011), as shown in Table 1.

Bayesian Model Averaging and Zellner's g-Prior

Bayesian model averaging

Bayesian Model Averaging (BMA) is a technique designed to help account for the uncertainty inherent in the model selection process, BMA focuses on which regressors to include in the analysis. By averaging across a large set of models one can determine those variables which are relevant to the data generating process for a given set of priors used in the analysis (Hoeting *et al.*, 1999). Given a linear regression model with constant term β_0 and k potential explanatory variables $x_1, x_2, \dots, x_k$ of the form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

This gives rise to 2^k possible sampling models (indexed M_j, $j = 1, 2, \dots, 2^k$), depending on whether we include or exclude each of the regressors. Once the model space has been determined, the posterior distribution of any coefficient of interest (say β_h), given the data D is:

$$P(\beta_h|D) = \sum_{j=1}^{2^k} P(\beta_h|M_j)\, P(M_j|D) \quad (2)$$

BMA uses each model's posterior probability, $P(M_j \mid D)$ as weights. Each model (a set of variables) receives a weight and the final estimates are constructed as a weighted average of the parameter estimates from each of the models. BMA includes all of the variables within the analysis, but shrinks the impact of certain variables towards zero through the model weights. These weights are the key feature for estimation via BMA and will depend upon a number of key features of the averaging exercise including the choice of prior specified (Montgomery & Nyhan, 2010).

The posterior model probability of M_j is given by Raftery *et al.* (2010):

$$P(M_j|D) = P(D|M_j)\frac{P(M_j)}{P(D)} = P(D|M_j)\frac{P(M_j)}{\sum_{i=1}^{2^k} P(D|M_i)P(M_i)} \quad (3)$$

where $P(D|M_j) = \int P(D|\beta^j, M_j)P(\beta^j|M_j)d\beta^j \quad (4)$

and β^j is the vector of parameters from model M_j, $P(\beta^j \mid M_j)$ is a prior probability distribution assigned to the

parameters of model M_j and $P(M_j)$ is the prior probability that M_j is the true model.
The estimated posterior means and standard deviations of $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ for model M_j are then constructed as (García-donato *et al.*, 2013):

$$E[\hat{\beta} \mid D] = \sum_{j=1}^{2^k} \hat{\beta} P(M_j|D) \qquad (5)$$

$$V[\hat{\beta} \mid D] = \sum_{j=1}^{2^k} \left(Var[\beta \mid D, M_j] + \hat{\beta}^2\right) P(M_j|D) - E[\beta \mid D]^2 \qquad (6)$$

Zellner's g-Priors

Zellner's g-priors applied in BMA analysis fixes a constant $g > 0$ and specifies the Gaussian prior for the regression coefficients β, conditional on σ^2. Thus, Zellner's g reduces the elicitation of the covariance structure by simply choosing the scalar g (Agliari & Parisetti, 1988);

Assumed model: $Y = X\beta + \varepsilon$ (7)

with $\varepsilon \sim N(0, \sigma^2 I_n)$, I_n is an identity matrix of order n.

The likelihood:

$$P(Y \mid X, \beta, \sigma^2) = (2\pi)^{\frac{-n}{2}} (\sigma^2)^{\frac{-n}{2}} exp\left(-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right). \qquad (8)$$

The Prior: $\beta \mid \sigma^2 \sim N(\beta_0, \text{g}\Omega)$. (9)

The Posterior:

$$\beta \mid \sigma^2, X \sim N(\beta_0, \sigma^2 \text{g}(X^T X)^{-1}) \qquad (10)$$

$$\beta \mid Y, \sigma^2, X \sim N\left(\frac{1}{\text{g}+1}(\beta_0 + \text{g}\hat{\beta}), \frac{\sigma^2 \text{g}}{\text{g}+1}(X^T X)^{-1}\right) \qquad (11)$$

$$E[\beta \mid Y, \sigma^2] = \left(\frac{1}{\text{g}\sigma^2} X^T X + \frac{1}{\sigma^2} X^T X\right)\left(\frac{1}{\text{g}\sigma^2} X^T X \beta_0 + \frac{1}{\sigma^2} X^T Y\right) \qquad (12)$$

$$E[\beta \mid Y, \sigma^2] = \frac{1}{1+\text{g}}\beta_0 + \frac{\text{g}}{1+\text{g}}(X^T X)^{-1} X^T Y \qquad (13)$$

$$= \frac{1}{1+\text{g}}\beta_0 + \frac{\text{g}}{1+\text{g}}\hat{\beta} \qquad (14)$$

Thus, the parameter g allows for direct weighting of the prior, β_0, and data, $\hat{\beta}$. This prior is known as Zellner's informative g-prior, or often referred to simply as the, g-prior. The hyper parameter g embodies how certain a researcher is that the coefficients are indeed zero. The value of g corresponds to the degree of prior uncertainty, (Hanson *et al.*, 2014). The g-prior is not only intuitive to use in the model and prior definition, but also leads to familiar posterior results, (Zhang *et al.*, 2008).

Table 1: Summary of identified g-prior structures examined

S/N	Structure of g-prior	Comments/Sources	S/N	Structure of g-prior	Comments/Sources
1	$\mathbf{g = n}$	Unit Information Prior (UIP) based on number of observations Kass & Wasserman (1996).	6	$\mathbf{g = \sqrt{\frac{1}{n}}}$	This is an intermediate case, where we choose a smaller asymptotic penalty term for large models than in the Schwarz criterion.
2	$\mathbf{g = max(n,\ k^2)}$	Corresponds to the benchmark prior suggested by Fernandez *et al.*, (2001b).	7	$\mathbf{g = \sqrt{\frac{k}{n}}}$	The prior information increases with the Number of regressors in the model. (Fernandez *et al*., 2001a)
3	$\mathbf{g = k^2}$	Conforms to the risk inflation criterion by Foster & George (1994).	8	$\mathbf{g = \frac{1}{k^2}}$	This prior is suggested by the risk inflation criterion (RIC). Foster & George (1994).
4	$\mathbf{g = \frac{1}{n}}$	It is in the spirit of the" unit information priors" of Kass & Wasserman (1996).	9	$\mathbf{g = \frac{n}{\sqrt{k}}}$	Conforms with the spirit of "Unit Information Prior and increased number of regressors (Ogundeji *et al*., 2018a, 2018b)
5	$\mathbf{g = \frac{k}{n}}$	Here, we assign more information to the prior as we have more regressors in the model			

Methodology

The methods and framework used by Zellner (1986) to obtain both the prior distributions and posterior distributions for Multiple Regression Models are adopted.

Given a regression model:

$$Y = X\beta + \varepsilon \qquad (15)$$

with $\varepsilon \sim N(0, \sigma^2 I_n)$.

The likelihood function for the model is given by

$$l(\beta, \sigma \mid Y, X) \ \alpha \ \sigma^{-n} exp\left(-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right). \qquad (16)$$

$$\alpha \ \sigma^{-n} exp\left\{-\left[vs^2 + (\beta - \hat{\beta})' X'X(\beta - \hat{\beta})\right]/2\sigma^2\right\}. \qquad (17)$$

where $\hat{\beta} = (X'X)^{-1}X'Y$, $vs^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})$ *and* $v = n\text{-}k$.

Given anticipated values of β and σ^2 denoted by β_a and σ_a^2 respectively from a conceptual or imaginary sample:

$$Y_0 = X\beta + \varepsilon_0 \qquad (18)$$

The joint informative g-prior distribution is:

$$P(\beta, \sigma^2 \mid \eta_0) \ \alpha \ \sigma^{-(v+1)} exp\left\{-\frac{v\overline{\sigma}_a^2}{2\sigma^2}\right\} \text{ x } \sigma^{-k} exp\{-\text{g}(\beta - \beta_a)'X'X(\beta - \beta_a)/2\sigma^2\} \qquad (19)$$

where $\eta_0' = (\beta_a', \overline{\sigma}_a, \text{g}, v)$.

The marginal prior distributions for β and σ are respectively:

$$P(\beta \mid \beta_a,\ \mathrm{g}, v)\ \alpha\ \{v\bar{\sigma}_a^2 + \mathrm{g}(\beta - \beta_a)'X'X(\beta - \beta_a)\}^{-(v+k)/2} \tag{20}$$

and $$P(\sigma \mid \bar{\sigma}_a, V)\ \alpha\ \sigma^{-(v+1)}exp\{-v\bar{\sigma}_a^2/2\sigma^2\} \tag{21}$$

The posterior distributions for β and σ given a g-prior:

$$P(\beta,\sigma \mid D)\ \alpha\ P(\beta,\sigma)\, l(\beta,\sigma \mid Y) \tag{22}$$

$$\alpha\ \sigma^{-(n+k+1)}exp\{-[(Y - X\beta)'(Y - X\beta) + \mathrm{g}(\beta - \bar{\beta})'X'X(\beta - \bar{\beta})]/2\sigma^2\} \tag{23}$$

D denotes the data and $\bar{\beta}$ is the prior mean vector for regression coefficient vector β.

Let $w' = \left(Y' : \mathrm{g}^{\frac{1}{2}}\, \bar{\beta}'\, X'\right)$ and $Z' = \left(X' :\ \mathrm{g}^{\frac{1}{2}}\ X'\right)$

Then the terms in square brackets in the exponential can be expressed as:

$$(w - Z\beta)'(w - Z\beta) = \left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right) \tag{24}$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w.$

Thus (23) can be expressed as:

$$P(\beta,\sigma \mid D)\ \alpha\ \sigma^{-(n+k+1)}exp\left\{-\left[\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right]/2\sigma^2\right\} \tag{25}$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w = \left(\hat{\beta} + \mathrm{g}\bar{\beta}\right)/(1 + \mathrm{g}).$

This is the mean of the posterior distribution with $\hat{\beta} = (X'X)^{-1}X'Y.$

The covariance matrix of the conditional normal posterior distribution for β given σ, denoted by $V(\beta \mid \sigma,\ D)$, is

$$V(\beta \mid \sigma,\ D) = (Z'Z)^{-1}\sigma^2 \tag{26}$$

$$= (X'X)^{-1}\sigma^2/(1 + \mathrm{g}) \tag{27}$$

$$P(\beta \mid D)\ \alpha\ \left\{\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right\}^{-(n+k)/2} \tag{28}$$

with covariance matrix $$V(\beta \mid D) = (Z'Z)^{-1}\sigma^2 = (X'X)^{-1}\sigma^2/(1 + \mathrm{g}) \tag{29}$$

where $$(n-2)\sigma^2 \equiv \left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) = \left(Y - X\bar{\bar{\beta}}\right)'\left(Y - X\bar{\bar{\beta}}\right) + \mathrm{g}\left(\bar{\bar{\beta}} - \beta\right)'X'X\left(\bar{\bar{\beta}} - \beta\right). \tag{30}$$

Also, the marginal posterior distribution for σ obtained from (25) by integrating with respect to β, is

$$P(\sigma \mid D)\ \alpha\ \sigma^{-(n+1)}exp\left\{-\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right)/2\sigma^2\right\}. \tag{31}$$

Results and Discussion

Based on the methods in section three and relying on the results of Zellner (1986) on prior and posterior distributions for multiple regression models, the respective prior and posterior distributions for β and σ are obtained using each of the g-prior structures identified (Table 1).

The Prior and Posterior distributions for β and σ using the proposed g-prior: $g = \mathbf{n}$.

The joint informative g-prior distribution using $\mathbf{g = n}$ is:

$$P(\beta, \sigma^2 \mid \eta_0)\ \alpha\ \sigma^{-(v+1)}exp\left\{-\frac{v\bar{\sigma}_a^2}{2\sigma^2}\right\} \text{ x } \sigma^{-k}\ exp\{-\mathrm{n}(\beta - \beta_a)'X'X(\beta - \beta_a)/2\sigma^2\} \tag{32}$$

where $\eta_0' = (\beta_a', \bar{\sigma}_a, \mathrm{n}, v).$

The marginal prior distributions for β and σ are respectively:

$$P(\beta \mid \beta_a, \mathrm{n}, v)\ \alpha\ \{v\bar{\sigma}_a^2 + \mathrm{n}(\beta - \beta_a)'X'X(\beta - \beta_a)\}^{-(v+k)/2} \tag{33}$$

and $$P(\sigma \mid \bar{\sigma}_a, V)\ \alpha\ \sigma^{-(v+1)}exp\{-v\bar{\sigma}_a^2/2\sigma^2\}. \tag{34}$$

Similarly, the Posterior distribution for β and σ**:**

$$P(\beta,\sigma \mid D)\ \alpha\ P(\beta,\sigma)\, l(\beta,\sigma \mid Y) \tag{35}$$

$$\alpha\ \sigma^{-(n+k+1)}exp\{-[(Y - X\beta)'(Y - X\beta) + n(\beta - \bar{\beta})'X'X(\beta - \bar{\beta})]/2\sigma^2\}. \tag{36}$$

D denotes the data and $\bar{\beta}$ is the prior mean vector for regression coefficient vector β.

Let $w' = \left(Y' : \sqrt{\mathrm{n}}\, \bar{\beta}'\, X'\right)$ and $Z' = \left(X' : \sqrt{\mathrm{n}}\ X'\right).$

Then the terms in square brackets in the exponential can be expressed as:

$$(w - Z\beta)'(w - Z\beta) = \left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right) \tag{37}$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w.$ Thus (36) becomes:

$$P(\beta,\sigma \mid D)\ \alpha\ \sigma^{-(n+k+1)}exp\left\{-\left[\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right]/2\sigma^2\right\} \tag{38}$$

where $$\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w = \left(\hat{\beta} + \mathrm{n}\bar{\beta}\right)/(1 + \mathrm{n})$$

is the mean of the posterior distribution with $\hat{\beta} = (X'X)^{-1}X'Y.$

The marginal posterior distribution for β obtained from (38) by integrating with respect to σ, is

$$P(\beta \mid D)\ \alpha\ \left\{\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right\}^{-(n+k)/2} \quad (39)$$

with covariance matrix: $V(\beta \mid D) = (Z'Z)^{-1}\sigma^2 = (X'X)^{-1}\sigma^2/(1+\mathrm{n})$ (40)

The Prior and Posterior distributions for β and σ using the proposed g-prior: $\mathbf{g = k^2}$ **.**

The joint informative g-prior distribution using $\mathbf{g = k^2}$ is:

$$P(\beta, \sigma^2 \mid \eta_0)\ \alpha\ \sigma^{-(v+1)} exp\left\{-\frac{v\bar{\sigma}_a^2}{2\sigma^2}\right\} \text{ x } \sigma^{-k}\ exp\{-\mathrm{k}^2(\beta - \beta_a)'X'X(\beta - \beta_a)/2\sigma^2\} \quad (41)$$

where $\eta_0' = (\beta_a', \bar{\sigma}_a, \mathrm{k}^2, v)$.

The marginal prior distributions for β and σ are respectively:

$$P(\beta \mid \beta_a, \mathrm{k}^2, v)\ \alpha\ \{v\bar{\sigma}_a^2 + \mathrm{k}^2(\beta - \beta_a)'X'X(\beta - \beta_a)\}^{-(v+k)/2} \quad (42)$$

and $P(\sigma \mid \bar{\sigma}_a, V)\ \alpha\ \sigma^{-(v+1)} exp\{-v\bar{\sigma}_a^2/2\sigma^2\}$. (43)

The Posterior distribution for β and σ*:*

$$P(\beta, \sigma \mid D)\ \alpha\ P(\beta, \sigma)\ l(\beta, \sigma \mid Y) \quad (44)$$

$$\alpha\ \sigma^{-(n+k+1)} exp\{-[(Y - X\beta)'(Y - X\beta) + (k)^2(\beta - \bar{\beta})'X'X(\beta - \bar{\beta})]/2\sigma^2\} \quad (45)$$

D denotes the data and $\bar{\beta}$ is the prior mean vector for regression coefficient vector β .

Let $w' = (Y' : \mathrm{k}\,\bar{\beta}'\,\mathrm{X}')$ and $Z' = (X' : \mathrm{k}\ \mathrm{X}')$.

Then the terms in square brackets in the exponential can be expressed as

$$(w - Z\beta)'(w - Z\beta) = \left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right) \quad (46)$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w$. Thus (45) becomes:

$$P(\beta, \sigma \mid D)\ \alpha\ \sigma^{-(n+k+1)} exp\left\{-\left[\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right]/2\sigma^2\right\} \quad (47)$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w = \left(\hat{\beta} + \mathrm{k}^2\bar{\beta}\right)/(1 + \mathrm{k}^2)$.

This is the mean of the posterior distribution with $\hat{\beta} = (X'X)^{-1}X'Y$.

The marginal posterior distribution for β obtained from (47) by integrating with respect to σ, is:

$$P(\beta \mid D)\ \alpha\ \left\{\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right\}^{-(n+k)/2} \quad (48)$$

with covariance matrix: $V(\beta \mid D) = (Z'Z)^{-1}\sigma^2 = (X'X)^{-1}\sigma^2/(1+\mathrm{k}^2)$ (49)

The Prior and Posterior distributions for β and σ using the proposed g-prior: $\mathbf{g = \frac{1}{n}}$**.**

The joint informative g-prior distribution using $\mathbf{g = \frac{1}{n}}$ is:

$$P(\beta, \sigma^2 \mid \eta_0)\ \alpha\ \sigma^{-(v+1)} exp\left\{-\frac{v\bar{\sigma}_a^2}{2\sigma^2}\right\} \text{ x } \sigma^{-k}\ exp\left\{-\left(\frac{1}{\mathrm{n}}\right)(\beta - \beta_a)'X'X(\beta - \beta_a)/2\sigma^2\right\} \quad (50)$$

where $\eta_0' = \left(\beta_a', \bar{\sigma}_a, \left(\frac{1}{\mathrm{n}}\right), v\right)$.

The marginal prior distributions for β and σ are respectively:

$$P\left(\beta \mid \beta_a, \left(\frac{1}{\mathrm{n}}\right), v\right)\ \alpha\ \left\{v\bar{\sigma}_a^2 + \left(\frac{1}{\mathrm{n}}\right)(\beta - \beta_a)'X'X(\beta - \beta_a)\right\}^{-(v+k)/2} \quad (51)$$

and $P(\sigma \mid \bar{\sigma}_a, V)\ \alpha\ \sigma^{-(v+1)} exp\{-v\bar{\sigma}_a^2/2\sigma^2\}$. (52)

The Posterior distribution for β and σ**:**

$$P(\beta, \sigma \mid D)\ \alpha\ P(\beta, \sigma)\ l(\beta, \sigma \mid Y) \quad (53)$$

$$\alpha\ \sigma^{-(n+k+1)} exp\left\{-\left[(Y - X\beta)'(Y - X\beta) + \left(\frac{1}{n}\right)(\beta - \bar{\beta})'X'X(\beta - \bar{\beta})\right]/2\sigma^2\right\}. \quad (54)$$

D denotes the data and $\bar{\beta}$ is the prior mean vector for regression coefficient vector β .

Let $w' = \left(Y' : \left(\frac{1}{\mathrm{n}}\right)^{\frac{1}{2}} \bar{\beta}'\,\mathrm{X}'\right)$ and $Z' = \left(X' : \left(\frac{1}{\mathrm{n}}\right)^{\frac{1}{2}}\ \mathrm{X}'\right)$.

Then the terms in square brackets in the exponential can be expressed as

$$(w - Z\beta)'(w - Z\beta) = \left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right) \quad (55)$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w$. Thus (54) becomes:

$$P(\beta, \sigma \mid D)\ \alpha\ \sigma^{-(n+k+1)} exp\left\{-\left[\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right]/2\sigma^2\right\} \quad (56)$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w = \left(\hat{\beta} + \left(\frac{1}{\mathrm{n}}\right)\bar{\beta}\right)/\left(1 + \left(\frac{1}{\mathrm{n}}\right)\right)$.

This is the mean of the posterior distribution with $\hat{\beta} = (X'X)^{-1}X'Y$.

The marginal posterior distribution for β obtained from (56) by integrating with respect to σ, is:

$$P(\beta \mid D)\ \alpha\ \left\{\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right\}^{-(n+k)/2} \quad (57)$$

with covariance matrix: $$V(\beta \mid D) = (Z'Z)^{-1}\sigma^2 = (X'X)^{-1}\sigma^2 / \left(1 + \left(\frac{1}{\mathrm{n}}\right)\right) \tag{58}$$

The Prior and Posterior distributions for β and σ using the proposed g-prior: $\mathbf{g} = \frac{\mathbf{k}}{\mathbf{n}}$.

The joint informative g-prior distribution using $\mathbf{g} = \frac{\mathbf{k}}{\mathbf{n}}$ is:

$$P(\beta, \sigma^2 \mid \eta_0) \;\alpha\; \sigma^{-(v+1)} exp\left\{-\frac{v\bar{\sigma}_a^2}{2\sigma^2}\right\} \text{ x } \;\sigma^{-k}\; exp\left\{-\left(\frac{\mathrm{k}}{\mathrm{n}}\right)(\beta - \beta_a)'X'X(\beta - \beta_a)/2\sigma^2\right\} \tag{59}$$

where $\eta_0' = \left(\beta_a', \bar{\sigma}_a, \left(\frac{\mathrm{k}}{\mathrm{n}}\right), v\right)$.

The marginal prior distributions for β and σ are respectively:

$$P\left(\beta \mid \beta_a, \left(\frac{\mathrm{k}}{\mathrm{n}}\right), v\right) \;\alpha\; \left\{v\bar{\sigma}_a^2 + \left(\frac{k}{\mathrm{n}}\right)(\beta - \beta_a)'X'X(\beta - \beta_a)\right\}^{-(v+k)/2} \tag{60}$$

and $$P(\sigma \mid \bar{\sigma}_a, V) \;\alpha\; \sigma^{-(v+1)} exp\{-v\bar{\sigma}_a^2/2\sigma^2\}. \tag{61}$$

The Posterior distribution for β and σ**:**

$$P(\beta, \sigma \mid D) \;\alpha\; P(\beta, \sigma)\, l(\beta, \sigma \mid Y) \tag{62}$$

$$\alpha\; \sigma^{-(n+k+1)} exp\left\{-\left[(Y - X\beta)'(Y - X\beta) + \left(\frac{k}{n}\right)(\beta - \bar{\beta})'X'X(\beta - \bar{\beta})\right]/2\sigma^2\right\}. \tag{63}$$

D denotes the data and $\bar{\beta}$ is the prior mean vector for regression coefficient vector β.

Let $w' = \left(Y' : \left(\frac{\mathrm{k}}{\mathrm{n}}\right)^{\frac{1}{2}} \bar{\beta}' \mathrm{X}'\right)$ and $Z' = \left(X' : \left(\frac{\mathrm{k}}{\mathrm{n}}\right)^{\frac{1}{2}} \mathrm{X}'\right)$.

Then the terms in square brackets in the exponential can be expressed as;

$$(w - Z\beta)'(w - Z\beta) = \left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right) \tag{64}$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w$. Thus (63) becomes:

$$P(\beta, \sigma \mid D) \;\alpha\; \sigma^{-(n+k+1)} exp\left\{-\left[\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right]/2\sigma^2\right\} \tag{65}$$

where $\bar{\bar{\beta}} = (Z'Z)^{-1}Z'w = \left(\hat{\beta} + \left(\frac{\mathrm{k}}{\mathrm{n}}\right)\bar{\beta}\right) / \left(1 + \left(\frac{\mathrm{k}}{\mathrm{n}}\right)\right)$.

This is the mean of the posterior distribution with $\hat{\beta} = (X'X)^{-1}X'Y$.

The marginal posterior distribution for β obtained from (65) by integrating with respect to σ, is:

$$P(\beta \mid D) \;\alpha\; \left\{\left(w - Z\bar{\bar{\beta}}\right)'\left(w - Z\bar{\bar{\beta}}\right) + \left(\beta - \bar{\bar{\beta}}\right)'Z'Z\left(\beta - \bar{\bar{\beta}}\right)\right\}^{-(n+k)/2} \tag{66}$$

with covariance matrix: $$V(\beta \mid D) = (Z'Z)^{-1}\sigma^2 = (X'X)^{-1}\sigma^2 / \left(1 + \left(\frac{\mathrm{k}}{\mathrm{n}}\right)\right) \tag{67}$$

Conclusion

In Bayesian Model Averaging, the sensitivity of g-priors to predictive performance of regression models have been demonstrated (Ogundeji *et al.*, 2018). Ultimately, the predictive performance of regression models is sensititve to the resulting posterior distribution based on a specified g-prior. However, the estimation of g-priors is less straigthforward than the estimation of regressiom parameters.

The study investigated nine g-priors identified in literature and based on the framework and methods in sections two and three. The results of prior distributions and posterior distributions of the regression parameters were derived for the g-prior structures numbered 1, 2, 3, 4 and 5 in Table 1. Similarly, the results of prior distributions and posterior distributions of the regression parameters derived for the g-prior structures numbered 6, 7, 8 and 9 are also available. Further work to be done will include the derivation of the sampling properties in term of the expected mean and variance of the posterior distributions for the respective g-prior structures investigated.

Conflict of Interest

Authors have declared that there is no conflict of interest reported in this work.

References

Agliari A & Parisetti C 1988. A-g reference informative prior: A note on Zellner's g-Prior. *Journal of the Royal Statistical Society*, Series D (The Statistician), 37(3): 271-275.

Clyde M & George EI 2004. Model uncertainty. *Statistical Science*, 19: 81–94.

Davison AC 2008. Statistical Models. Cambridge University Press, New York.

Eicher T, Papageorgiou C & Raftery A 2011. Default priors and predictive performance in Bayesian model averaging, with application to growth determinants. *J. Appl. Econometrics*, 26: 30–55.

Feldkircher M 2012. Forecast combination and Bayesian model averaging: A prior sensitivity analysis. *Journal of Forecasting*, 31(4): 361 – 376.

Fernandez C, Ley E & Steel M 2001a. Model uncertainty in cross-country growth regressions. *J. Appl. Econometrics*, 16: 563-576.

Fernandez C, Ley E & Steel M 2001b. Benchmark priors for Bayesian model averaging. *Journal of Econometrics*, 100: 381-427.

Foster DP & George EI 1994. The risk inflation criterion for multiple regression. *Annals of Statistics*, 22: 1947 – 1975.

Fouskakis D & Ntzowfras I 2013. Power Conditional-Expected Prior: Using g-Priors with Random Imaginary Data for Variable Selection.

Hanson TE, Branscum AJ & Johnson WO 2014. Informative g-priors for logistic regression. *Bayesian Analysis*, 9(3): 597- 612.

Hoeting JA, Madigan D, Raftery AE & Volinsky CT 1999. Bayesian model averaging: A tutorial (with discussion). *Statistical Science*, 14: 382–401, corrected version at http://www.stat.washington.edu/www/research/online/hoeting1999.pdf.

Kass R & Wasserman L 1996. The selection of prior distributions by formal rules. *J. Am. Stat. Assoc.*, 91(435): 1343-1370.

Li Y & Clyde M 2015. Mixtures of g-priors in Generalised Linear Models. Cornell University Library, axXiv: 1503.06913.

Montgomery J & Nyhan B 2010. Bayesian model averaging: Theoretical developments and practical applications. *Oxford Journal: Political Analysis*, 18(2): 245-270.

Ogundeji RK, Adeleke I & Okafor RO 2018a. An assessment of predictive performance of Zellner's g-priors in Bayesian model averaging. *Model Assisted Statistics and Applications* (IOS Press), 13: 63–71.

Ogundeji RK, Adeleke I & Okafor RO 2018b. A new Zellner's g-prior for Bayesian model averaging in regression analysis. *J. Scientific Res. and Dev. (JSRD)*, 18(1): 51-60.

Okafor R 1999. Using an empirical Bayes model to estimate currency exchange rates. *J. Appl. Stat.*, 26(8): 973 – 983.

Raftery AE, Karny M & Ettler P 2010. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics*, 52: 52-66.

Raftery AE, Madigan D & Hoeting JA 1997. Bayesian model averaging for linear regression models. *J. Am. Stat. Assoc.*, 92: 179–191.

Rossi PE, Allenby GM & Mcculloch R 2005. Bayesian Statistics and Marketing. Wiley Series in Probability and Statistics.

Zellner A 1983. Applications of Bayesian analysis in econometrics. *J. Royal Stat. Soc.*. Proceedings of the 1982 I.O.S. Annual Conference on Practical Bayesian Statistics, Series D (The Statistician), 32(2): 23-34.

Zellner A 1986. On assessing prior distributions and Bayesian regression analysis with g-prior distributions. *The American Statistician*, 49: 327-335.

Zhang Z, Jordan M & Yeung D 2008. Posterior consistency of the Silverman g-prior in Bayesian model choice. *22nd Annual Conference Neural Information Processing System (NIPS)*, 1969-1976.