International Journal of Advanced Scientific and Technical Research Issue 5 volume 4, July-August 2015
Available online on http://www.rspublication.com/ijst/index.html ISSN 2249-9954

BAYESIAN POSTERIOR ESTIMATES: QUANTILES IN DETERMINING NORMAL HYPERPARAMETERS.

Fasoranbaku Olusoga Akin#1, Salami Taofeek Adeola#2 and Soyombo Ayodeji Olugbenga#3

#1 Department of Statistics, Federal University Of Technology, Akure, Nigeria.
+2348033932972, oafasoranbaku@yahoo.com

#2 Department of Statistics, Federal University Of Technology, Akure, Nigeria.
+2348034182907, taofeekone@yahoo.com

#3 Department of Statistics, Federal University Of Technology, Akure, Nigeria.
+2347032980805, ayo_sho@yahoo.com

ABSTRACT

This paper focuses on obtaining Bayesian posterior estimates using quantiles to determine normal hyperparameters. This is with a view to obtain prior estimates using quantiles of the confidence interval of the Ordinary Least Square (OLS) estimates and determine the adequacy of the models in order to identify the best. The data used is a secondary data, collected from the Ministry of Land and Physical Planning, Osogbo, Osun State, Nigeria. The Confidence Interval (CI) is divided into 100 equal parts to give 99 grid points and each grid point is taken as the hyperparameter of the prior distributions. The values of the prior hyperparameters are chosen to reflect the prior information. The posterior mean is the estimate of the regression coefficient β_0 - β_2 and for each of the 99 models, we obtain the Mean Absolute Deviation (MAD) and choose the model with the least MAD as the corresponding optimal hyperparameter. The adequacy of the models is determined, using mean absolute deviation (MAD). Though the 47th, 48th, and 49th models have smaller MAD compared to that of the OLS estimate, the least MAD corresponds to the 47th model, which is the optimal hyperparameter.

Keywords: Bayesian Posterior Estimates, Quantiles, Normal Hyperparameters, Mean Absolute Deviation.

Corresponding Author: Soyombo .A.O

INTRODUCTION

Bayesian econometrics is based on a few simple rules of probability. This is one of the chief advantages of the Bayesian approach. All of the things that an econometrician would wish to do, such as estimate the parameters of a model, compare different models or obtain predictions from a model, involve the same rules of probability. Bayesian methods are thus, universal and can be used any time a researcher is interested in using data to learn about a phenomenon.

Bayesians treat **p(θ/*y*)** as being of fundamental interest. That is, it directly addresses the question; "Given the data, what do we know about θ?" The treatment of θ as a random variable is controversial among some econometricians. The chief competitor to Bayesian econometrics, often called frequentist econometrics, says that θ is not a random variable. However, Bayesian econometrics is based on a subjective view of probability, which argues that our uncertainty about anything unknown can be expressed using the rules of probability. Econometrics involves learning about something unknown (e.g. coefficients in a regression) given something known (e.g. data) and the conditional probability of the unknown given the known is the best way of summarizing what we have learned.

In addition to learning about parameters of a model, an econometrician might be interested in comparing different models. In cases where many models are being entertained, it is important to be explicit about which model is under consideration.

Ricardo Ehlers (2007) use Markov Chain Monte Carlo (MCMC) methods in order to estimate and compare stochastic production frontier models from a Bayesian perspective. He considers a number of competing models in terms of different production functions and the distribution of the asymmetric error term. In order to compare and select the most appropriate model among those considered, he used the Deviance Information Criterion (DIC), where lower values indicate a good model fit relative to the number of parameters in the model. Additive quantile regression models have gained considerable attention. Yue and Rue (2011) and Oh et al. (2011) propose differentiable approximations to the AWAD criterion that allow to employ different types of smoothing approaches while Li et al. (2010) show ways to incorporate regularization on fixed effects into linear Bayesian quantile regression. Fenske et al. (2011) propose boosting approaches for flexible, additive quantile regression models, where penalized least squares estimates are utilized as base-learners.

This paper focuses on obtaining Bayesian posterior estimates using quantiles to determine normal hyperparameters. This is with a view to obtain prior estimates using quantiles of the confidence interval of the Ordinary Least Square (OLS) estimates and determine the adequacy of the models in order to identify the best. The article is organized as follows. In section 2, we consider the methodology in obtaining the posterior Bayesian estimates and using the mean absolute deviation to get the optimal hyperparameter. Section 3 reports the results and discussion of the experiment.

MATERIALS AND METHODS

Given the model Y = Xβ, where Y is n×1 vector, X = n×p matrix and β = p×1 vector; the Ordinary Least Square (OLS) estimate of β is $\hat{\beta} = (X'X)^{-1}(X'Y)$, where $\hat{\beta} \sim N(\beta, \Sigma)$. Confidence intervals for β will be obtained through $\hat{\beta} \pm t_{\alpha/2} \times SE(\beta)$ where $SE(\beta) = S\sqrt{x_{ii}}$ and x_{ii} = diag ($(X'X)^{-1}$). The Confidence Interval (CI) is divided into 100 equal parts to give 99 grid points. Each grid point is taken as the hyperparameter of the prior distributions:

$\beta / h \sim N(\underline{\beta}, h^{-1}\underline{V})$ and a prior for h of the form $h \sim G(\underline{S}^{-2}, \underline{v})$ The notation for the Normal-Gamma distribution of the natural conjugate prior for β and h is denoted by: $\beta, h \sim NG(\underline{\beta}, \underline{V}, \underline{S}^{-2}, \underline{v})$. The notation for the prior density is:

$$P(\beta, h) = f_{NG}(\underline{\beta}, h / \underline{\beta}, \underline{V}, \underline{s}^{-2}, \underline{v})$$ where:

β is a K-vector containing the prior means for the k regression coefficients $\beta_1 \ldots \beta_k$.

$\underline{V}$ is a K×K positive definite prior covariance matrix. Therefore, the particular values of the so-called prior hyperparameters $\underline{\beta}, \underline{V}, \underline{S}^{-2}$ and $\underline{v}$ are chosen to reflect the prior information. Next, we obtain the posterior Bayesian estimates. The posterior mean is the Bayes estimate which is the estimate of the regression coefficient β_0 - β_2 given as:

$$\frac{\text{Prior Precision}}{\text{Posterior Precision}} \times \text{Prior Mean} + \frac{\text{Observation Precision}}{\text{Posterior Precision}} \times \text{OLS Estimate}$$

For β_0,

$$\frac{\frac{1}{S^2_{\beta_0}}}{\frac{1}{S^2_{\beta_0}} + \frac{n}{\sigma^2}} \times PM_{\beta_0} + \frac{\frac{n}{\sigma^2}}{\frac{1}{S^2_{\beta_0}} + \frac{n}{\sigma^2}} \times \beta_0 OLS \ \ Estimate$$

For β_1,

$$\frac{\frac{1}{S^2_{\beta_1}}}{\frac{1}{S^2_{\beta_1}} + \frac{SSX_{f_i}}{\sigma^2}} \times PM_{\beta_1} + \frac{\frac{SSX_{f_i}}{\sigma^2}}{\frac{1}{S^2_{\beta_1}} + \frac{SSX_{f_i}}{\sigma^2}} \times \beta_1 OLS \ \ Estimate$$

For β_2,

$$\frac{\frac{1}{S^2_{\beta_2}}}{\frac{1}{S^2_{\beta_2}} + \frac{SSX_{g_i}}{\sigma^2}} \times PM_{\beta_2} + \frac{\frac{SSX_{g_i}}{\sigma^2}}{\frac{1}{S^2_{\beta_2}} + \frac{SSX_{g_i}}{\sigma^2}} \times \beta_2 OLS \ \ Estimate$$

For each of the 99 models, we obtain the Mean Absolute Deviation (MAD) and choose the model with the least MAD. In comparison with the ordinary least square estimation, the best model shall be the model with the least MAD which also is expected to be less than the mean square error (MSE) of the ordinary least square estimate, while the mean absolute deviation of the 50th model is expected to be equivalent to the mean square error of the ordinary least square estimate. The corresponding hyperparameter shall be the optimal hyperparameter.

Table (1) below shows the original data used for the estimation.

Table 1: Original Data

S/N	Area (Sq/m)	Age of house (yrs)	Price ('000)
1	82	10	47500
2	125	12	64550
3	73	15	41250
4	143	17	79750
5	162	13	89550
6	115	6	53350
7	154	5	81200
8	83	7	59800
9	95	3	53200
10	93	8	50100
11	107	9	68400
12	76	10	35300
13	104	11	59300
14	75	9	41100
15	79	8	63500
16	89	22	51400
17	85	13	45900
18	99	16	54700
19	106	19	60600
20	117	7	54500
21	73	18	53100
22	81	5	48000
23	97	10	59000
24	126	8	60300
25	137	16	73000
26	129	9	70000
27	84	14	49600
28	76	16	48300
29	102	13	55300
30	69	18	30000
31	78	7	41500
32	84	21	49000
33	94	7	56300
34	105	9	63500
35	110	16	66600

RESULTS AND DISCUSSION

Table (2) presents the 99 grid points (models) with their corresponding mean absolute deviation. The table reveals that the corresponding mean absolute deviations (MAD) of the 1st model to 46th model and 51st model to 99th model are greater than the mean squared error (MSE) of the ordinary least squared estimates. It is clear from the table, that the corresponding mean absolute deviation (MAD) of the 50th model with value 4404.13, is equivalent to the mean squared error of the ordinary least squared estimates. Also, the table shows that the MAD of the 47th, 48th and 49th models with values 4401.24, 4402.15 and 4403.15 respectively are all less than the mean squared error (MSE) of the ordinary least squared estimates. The 47th model has the least mean absolute deviation (MAD) among the three.

Figure (1) shows the graph of the mean absolute deviation versus the corresponding models. The above results are summarized in the table and figure below.

FIGURE 1: Graph of Mean Absolute Deviation versus Models

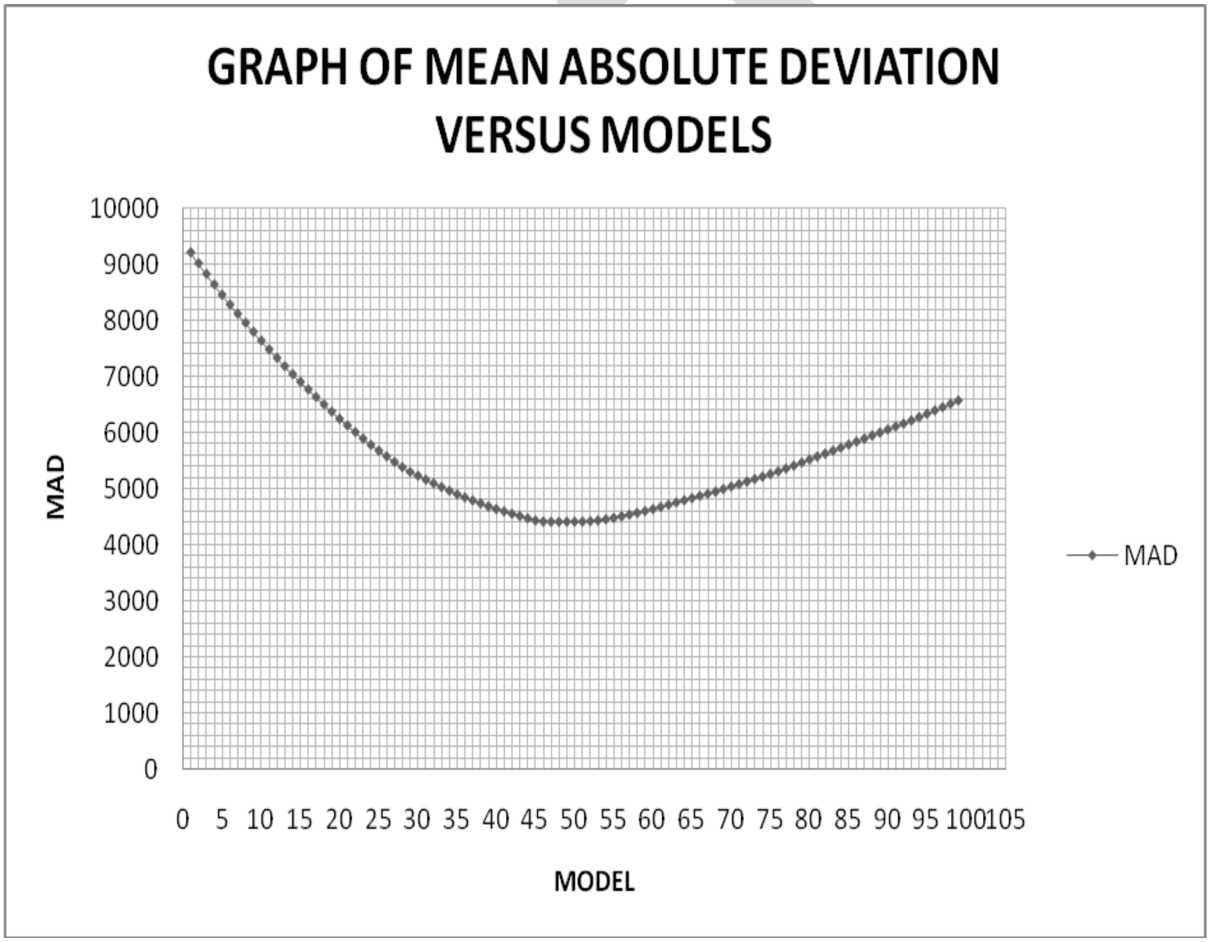

TABLE 2: Mean Absolute Deviations of corresponding models

MODEL	MAD	MODEL	MAD	MODEL	MAD
1st	9199.16	34th	4955.67	67th	4900.74
2nd	9006.13	35th	4891.36	68th	4938.44
3rd	8816.23	36th	4835.40	69th	4983.79
4th	8629.38	37th	4781.63	70th	5029.09
5th	8445.51	38th	4728.51	71st	5073.93
6th	8271.19	39th	4676.03	72nd	5118.32
7th	8106.54	40th	4628.82	73rd	5162.27
8th	7944.44	41st	4586.70	74th	5205.78
9th	7784.82	42nd	4545.09	75th	5251.92
10th	7627.62	43rd	4503.96	76th	5300.04
11th	7472.78	44th	4463.32	77th	5347.68
12th	7320.26	45th	4426.13	78th	5401.58
13th	7171.48	46th	4405.59	79th	5455.93
14th	7031.92	**47th**	**4401.24**	80th	5509.75
15th	6894.38	**48th**	**4402.15**	81st	5563.06
16th	6758.82	**49th**	**4403.15**	82nd	5615.87
17th	6625.20	**50th**	**4404.13**	83rd	5668.17
18th	6493.46	51st	4405.10	84th	5719.97
19th	6363.58	52nd	4411.62	85th	5774.75
20th	6235.50	53rd	4427.04	86th	5829.73
21st	6116.68	54th	4447.19	87th	5884.20
22nd	6000.11	55th	4470.18	88th	5938.16
23rd	5885.11	56th	4498.13	89th	5991.62
24th	5771.67	57th	4529.09	90th	6044.60
25th	5668.66	58th	4559.72	91st	6097.09
26th	5567.54	59th	4592.13	92nd	6149.10
27th	5467.76	60th	4628.66	93rd	6203.22
28th	5375.23	61st	4666.13	94th	6264.23
29th	5289.69	62nd	4706.27	95th	6324.68
30th	5221.17	63rd	4745.98	96th	6384.60
31st	5153.52	64th	4785.28	97th	6443.98
32nd	5086.73	65th	4824.17	98th	6502.84
33rd	5020.79	66th	4862.66	99th	6561.18

CONCLUSION

The present study constitutes an important contribution to the field of econometrics. In an obvious way, this work is important because of its direct ties to substantive research and the process of simplifying the selection of hyperparameters of prior distribution in empirical Bayesian inference. This is a valuable focus for research, given the increasing availability of alternative estimation methods within software packages. By using the mean absolute deviation, the MAD of the 47^{th}, 48^{th} and 49^{th} models are all less than the mean squared error (MSE) of the ordinary least squared estimates. Therefore, the 47^{th}, 48^{th} and 49^{th} models are better than the OLS. Since the 47^{th} model has the least MAD and gives the best model, we choose it as the corresponding optimal hyperparameter.

REFERENCE

Abu-Dayyeh, W. and Al-Sawi, E. (2007). Modified inference about the mean of the exponential distribution using moving extreme ranked set sampling. Statistical Papers, 50(2): 249-259.

Al-Hadhrami, S.A. (2007). Estimation of the population mean using ranked set sampling with auxiliary variable. PhD Thesis, UKM, Malaysia.

Al-Hadhrami, S.A. and Al-Omari, A.I. (2010). Bayesian inference on the variance of Normal distribution using moving extreme ranked set sampling.

Al-Hadhrami, S.A., Al Omari, A.E. and Al-Saleh, M.F. (2009). Estimation of standard deviation of normal distribution using moving extreme ranked set sampling. World Academy of Science, Engineering and Technology. 37, 988-993.

Al-Saleh, M.F. and Al-Ananbeh, A.M. (2007). Estimation of the means of the bivariate normal using moving extreme ranked set sampling with concomitant variable. Statistical Papers, 48(2): 179-195.

Al-Saleh, M.F. and Al-Hadhrami, S. (2003a).Parametric estimation for the location parameter for symmetric distributions using moving extremes ranked set sampling with application to trees data. Environmetrics, 14(7) 651-664.

Al-Saleh, M.F. and Al-Hadhrami, S. (2003b). Estimation of the mean of exponential distribution using moving extreme ranked set sampling. Statistical Papers, 44, 367-387.

Al-Saleh, M.F. and Samawi, H. (2010).On estimating the odds using moving extreme ranked set sampling. Statistical Methodology, 7 (2), 133- 140.

Chen, M.-H., Shao, Q.-M., and Ibrahim, J. (2000), Monte Carlo Methods in Bayesian Computation, New York: Springer Verlag.

Congdon, P. (2001), Bayesian Statistical Modelling, Hoboken, NJ: Wiley.

Congdon, P. (2003), Applied Bayesian Models, Hoboken, NJ: Wiley.

Congdon, P. (2005), Bayesian Models for Categorical Data, Hoboken, NJ: Wiley.

Finney, D. J. (1947), The Estimation from Individual Records of the Relationship between Dose and Quantal Response," Biometrika, 34, 320-334.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions.Econometrica, 77(3):953–973.

Gelfand, A. E., and Smith, A. F. M. (1990), Sampling-based Approaches to Calculating Marginal Densities," Journal of American Statistical Association, 85, 398-409.

Jouko Lampinen and Aki Vehtari (2001). Bayesian Approach for Neural Networks – Review

and Case Studies, journal article on Neural Networks, 14(3):7-24
Kottas, A. and Gelfand, A.E. (2001). Bayesian semiparametric median regression modeling.
Lindley, D. V. (1957), A Statistical Paradox," Biometrika, 44, 187-192.
Madigan, D. http://stat.rutgers.edu/ madigan/bayes02/. A Web Site for Course on Bayesian Statistics.
Martz, H., and Waller, R. (1985), Bayesian Reliability Analysis, New York: Wiley.
Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953), Equation of State Calculations by Fast Computing Machines," The Journal of Chemical Physics, 21, 1087-1092.
Michel Lubrano and Abdoul Aziz (2012). Bayesian Unconnditional Quantile Regression, An analysis of recent expansions in wage structure and earnings inequality in the U. S., Journal of Applied Econometrics, 20(4):445-465
Robert, C. (2001). The Bayesian Choice: From Decision-Theoretic Motivations to Computational Implementation, Second Edition, New York: Springer Verlag.
Robert, C. and Casella, G. (2004), Monte Carlo Statistical Methods, Second Edition, New York: Springer Verlag.
Tim Conley, Chris Hansen, Rob McCulloch and Peter E. Rossi, (2006). A Non-Parametric Bayesian Approach to the Instrumental Variable Problem, Graduate School of Business, University of Chicago.