British Journal of Mathematics & Computer Science
15(1): 1-8, 2016, Article no.BJMCS.23145
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org

Comparison of Least Squares Method and Bayesian with Multivariate Normal Prior in Estimating Multiple Regression Parameters

Felix O. Mettle[1], Louis Asiedu[1*], Enoch N. B. Quaye[1] and Abeku A. Asare-Kumi[1]

[1]Department of Statistics, School of Physical and Mathematical Sciences, College of Basic and Applied Sciences, University of Ghana, Ghana.

Authors' contributions

This work was carried out in collaboration between all authors. Author FOM designed the study, wrote the protocol and supervised the work. Authors LA and ENBQ wrote the codes for statistical computations and performed the statistical analysis. Author FOM wrote the first draft of the manuscript. Author AAAK managed the literature searches and edited the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2016/23145
Editor(s):
(1) H. M. Srivastava, Department of Mathematics and Statistics, University of Victoria, Canada.
Reviewers:
(1) Cliff Richard Kikawa, Tshwane University of Technology, South Africa.
(2) Ette Harrison Etuk, Rivers State University of Science and Technology, Nigeria.
(3) Loc Nguyen, Vietnam.
(4) S. Kim, Wayne State University, Detroit, USA.
Complete Peer review History: http://sciencedomain.org/review-history/13526

Original Research Article

Received: 16th November 2015
Accepted: 12th February 2016
Published: 3rd March 2016

Abstract

Based on an assumption of multivariate normal priors for parameters of multivariate regression model, this study outlines an algorithm for application of traditional Bayesian method to estimate regression parameters. From a given set of data, a Jackknife sample of least squares regression coefficient estimates are obtained and used to derive estimates of the mean vector and covariance matrix of the assumed multivariate normal prior distribution of the regression parameters. Driven to determine whether Bayesian methods to multivariate regression parameter estimation present a stable and consistent improvement over classical regression modeling or not, the study results indicate that the Bayesian method and Least Squares Method (LSM) produced almost the same estimates for the regression parameters and coefficient of determination (to 4.dp) with the Bayesian method having smaller standard errors.

**Corresponding author: E-mail: lasiedu@ug.edu.gh;*

Keywords: *Bayesian regression parameters; least square method; jackknife regression estimates; multivariate normal priors.*

2010 Mathematics Subject Classification: 62F15; 62JO5.

1 Introduction

In recent years Bayesian methods have become widespread in many domains including computer vision, signal processing, and information retrieval and genome data analysis. The availability of fast computers allows the required computations to be performed in reasonable time, and thereby makes the benefits of a Bayesian treatment accessible to an ever broadening range of applications [1].

Bayesian inference allows informative priors so that prior knowledge or results of a previous model can be used to inform the current model. Bayesian inference can also avoid problems with model identification by manipulating prior distributions. Classical Statistical inference with any numerical approximation algorithm does not have prior distributions, and can become stuck in regions of flat density, causing problems with model identification.

Bayesian inference considers the data to be fixed, which is true for real life data, and parameters to assume values within a specified range according to a prior distribution. LSM considers the unknown parameters to be fixed, and the data to be random. Estimation is not based only on the data at hand, but together with hypothetical repeated samples of similar data. The Bayesian approach delivers the answer to the right question in the sense that Bayesian inference provides answers conditional on the observed data and not based on the distribution of estimators or test statistics over hypothetical samples not observed [2].

Clearly, Bayesian methods have become widespread in many domains. Studies by [3,4] apply Bayesian method of moments (BMOM) which does not assume likelihood functions and prior density. In their study, [3] demonstrated how the BMOM can be employed to analyze parametric and semiparametric models. Also, [4] carried out Bayesian analysis of regression errors. [5] focused on using Bayesian inference with assumed multivariate normal prior to estimate missing data and their covariance matrix in choice conjoint experiment. [6] provided Bayesian interpretations for White's (errors) heteroskedastic consistent (HC) covariance estimator, and various modifications of it, in linear regression models. For existing literature on Bayesian data analysis, readers can refer to the work by [7] on Bayesian theory for normally distributed random variables.

The least squares estimation procedure has been used in problems that arise in many scientific investigations involving the study of observations whose theoretical mean values are known functions of parameters which are to be estimated ([8,9,10], etc). Suppose $\theta_1, \theta_2, \dots, \theta_k$ and let the available data consist of n observations $y_1, y_2, \dots, y_n$ with the expected values $E(y_i) = h_i(\theta_1, \theta_2, \dots, \theta_k), i = 1, 2, \dots, n.$ The least squares estimates are the values of $\theta_1, \theta_2, \dots, \theta_k$ that minimize $= \sum_{i=1}^{n}[y_i - E(y_i)]^2$; which are often obtained by solving the equations $\frac{\partial Q}{\partial \theta_i} = 0,\ i = 1, 2, \dots, k.$

This paper seeks to compare the results from a classical LSM approach to that of a Bayesian approach.

2 Model Specification

The multiple linear regression model of a response variable Y and k predictor variables $X_1, X_2, \cdots, X_k$ for a sample size of n is given by

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i, \qquad i = 1, 2, \cdots, n \tag{1}$$

$$= \boldsymbol{\beta}' \boldsymbol{X}_i + \varepsilon_i \tag{2}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k)'$; $\boldsymbol{X}_i = (1, X_{1i}, X_{2i}, \cdots, X_{ki})'$ and $\varepsilon_i \sim N(0, \sigma^2)$,

The multiple regression assumes that the errors are independent and distributed according to the normal distribution with zero mean and a constant variance denoted by σ^2. As a consequence of this, coupled with the assumption of fixed $X_1, X_2, \cdots, X_k$ and $\beta_0, \beta_1, \beta_2, \cdots, \beta_k$; the Y_is are also independent with each having a normal distribution with mean and variance given respectively $\boldsymbol{\beta}'\boldsymbol{X}_i$ and σ^2 $(i = 1, 2, \cdots, n)$. The least squares estimator for $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k)'$ is given by $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$; $\boldsymbol{X}$ is an $n \times k$ matrix with ith row $\boldsymbol{X}_i$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)'$ [8].

However if the components of $\boldsymbol{\beta}$ can assume values within a given range based on a prior distribution instead of fixed parameters as in LSM of estimating the regression parameters, then the conditional density function for each of the $Y_i's$ is given by

$$f(y_i|\boldsymbol{\beta}, \boldsymbol{X}_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y_i - \boldsymbol{\beta}'\boldsymbol{X}_i)^2} ;|y_i| \geq 0 \tag{3}$$

The conditional joint density function of $Y_i, i = 1, 2, \cdots, n$ is given by;

$$f(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{X}_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y_i - \boldsymbol{\beta}'\boldsymbol{X}_i)^2} \quad , \boldsymbol{y} = (y_1, y_2, \ldots, y_n) \tag{4}$$

Now suppose the random vector $\boldsymbol{\beta}$ has a multivariate normal distribution with mean vector $\boldsymbol{\mu} = (\mu_0, \mu_1, \mu_2, \cdots, \mu_k)'$ and covariance matrix $\boldsymbol{\Sigma}$; that is $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k)' \sim N_{k+1}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; then the joint density function of $\boldsymbol{y}$ and the coefficients $\boldsymbol{\beta}$ is given by;

$$\begin{aligned} f(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{X}_i) &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \boldsymbol{\beta}'\boldsymbol{X}_i)^2} \times (2\pi)^{-\left(\frac{k+1}{2}\right)} |\boldsymbol{\Sigma}|^{-1/2} e^{-1/2(\boldsymbol{\beta}-\mu)'\Sigma^{-1}(\boldsymbol{\beta}-\mu)} \\ &= K e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \boldsymbol{\beta}'\boldsymbol{X}_i)^2 - \frac{1}{2}\{\sum_{i=1}^{n} a_{ii}(\beta_i - \mu_i)^2 + 2\sum_{i \neq j}(\beta_i - \mu_i)(\beta_j - \mu_j)a_{ij}\}}; \end{aligned} \tag{5}$$

where, $K = (2\pi\sigma^2)^{-\frac{n}{2}} (2\pi)^{-\left(\frac{k+1}{2}\right)} |\Sigma|^{-1/2}$

Now let,

$$\begin{aligned} V &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \boldsymbol{\beta}'\boldsymbol{X}_i)^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{k}\beta_j X_{ji})^2 \\ &= \frac{1}{\sigma^2}\sum_{i=1}^{n} y_i^2 + \frac{n}{\sigma^2}\beta_0^2 + \frac{1}{\sigma^2}\sum_{j=1}^{k}\sum_{i=1}^{n}\beta_j^2 X_{ji}^2 - \frac{2}{\sigma^2}\left[\beta_0\sum_{i=1}^{n} y_i + \sum_{j=1}^{k}\sum_{i=1}^{n}\beta_j X_{ji} y_i\right] \\ &\quad + \frac{2}{\sigma^2}\sum_{j=0}^{k-1}\sum_{s=j+1}^{k}\sum_{i=1}^{n}\beta_j\,\beta_s X_{ji} X_{si} \end{aligned} \tag{6}$$

Also, let $W = (\boldsymbol{\beta} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu})$

$$\begin{aligned} &= \sum_{j=0}^{k} a_{jj}\left(\beta_j - \mu_j\right)^2 + 2\sum_{j=0}^{k-1}\sum_{s=j+1}^{k} a_{js}\left(\beta_j - \mu_j\right)(\beta_s - \mu_s),\ \Sigma^{-1} = \left(a_{ij}\right) \\ &= \sum_{j=0}^{k} a_{jj}\left(\beta_j^2 - 2\mu_j\beta_j + \mu_j^2\right) + 2\sum_{j=0}^{k-1}\sum_{s=j+1}^{k} a_{js}\left(\beta_j\beta_s - \mu_s\beta_j - \mu_j\beta_s + \mu_j\mu_s\right) \\ &= \sum_{j=0}^{k} a_{jj}\beta_j^2 - 2\sum_{j=0}^{k} a_{jj}\mu_j\beta_j + \sum_{j=0}^{k} a_{jj}\mu_j^2 + 2\sum_{j=0}^{k-1}\sum_{s=j+1}^{k} a_{js}\beta_j\beta_s - 2\sum_{j=0}^{k-1}\sum_{s=j+1}^{k} a_{js}\mu_s\beta_j - \\ &\quad 2\sum_{j=0}^{k-1}\sum_{s=j+}^{k} a_{js}\mu_j\beta_s + 2\sum_{j=0}^{k-1}\sum_{s=j+1}^{k} a_{js}\mu_j\mu_s \end{aligned} \tag{7}$$

So that,

$$\begin{aligned} V + W &= \left(\frac{n}{\sigma^2} + a_{00}\right)\beta_0^2 + \sum_{j=1}^{k}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n} X_{ji}^2 + a_{jj}\right]\beta_j^2 \\ &\quad + 2\sum_{j=0}^{k-1}\sum_{s=j+1}^{k}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n} X_{ji}X_{si} + a_{js}\right]\beta_j\beta_s - 2\left[\frac{1}{\sigma^2}\sum_{i=1}^{n} y_i + v_j\right]\beta_0 \\ &\quad - 2\sum_{j=1}^{k}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n} X_{ji}y_i + v_j\right]\beta_j + R \end{aligned} \tag{8}$$

where $v_j = \sum_{i=1}^{k} a_{ji}\,\mu_i,\ j = 0, 1, 2, \cdots, k,\quad a_{js} = a_{sj}$ for $j \neq s$ and R is a constant term independent of $\beta_j, (j = 0,1,2, \cdots, k)$.

Clearly,

$V + W = Q(\boldsymbol{\beta})$ is a quadratic form of the matrix $\boldsymbol{\Sigma}_{\beta}^{-1}$ in $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k)'$. Therefore the posterior distribution of $\boldsymbol{\beta}$ is of the form

$$f(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{X}) = Ke^{-\frac{1}{2}Q(\boldsymbol{\beta})} \tag{9}$$

Hence it follows the multivariate normal distribution with mean vector given by

$$\boldsymbol{\mu}_{\beta} = -\frac{1}{2}\Sigma_{\boldsymbol{\beta}}\boldsymbol{C} \tag{10}$$

where, $\boldsymbol{\Sigma}_{\beta}$ is a $(k+1) \times (k+1)$ matrix with an inverse $\boldsymbol{\Sigma}_{\beta}^{-1} = (m_{ij})$ whose elements are given as;

$$\begin{aligned} m_{00} &= \frac{n}{\sigma^2} + a_{00},\ j = 1, 2\ 3, \cdots, k \\ m_{j0} &= \frac{1}{\sigma^2}\textstyle\sum_{i=1}^{n} X_{ji} + a_{j0},\ j = 1, 2\ 3, \cdots, k \\ m_{0j} &= \frac{1}{\sigma^2}\textstyle\sum_{i=1}^{n} X_{ji} + a_{0j},\ j = 1, 2\ 3, \cdots, k \\ m_{ij} &= \frac{1}{\sigma^2}\textstyle\sum_{l=1}^{n} X_{il}X_{jl} + a_{ij,}\ \ i \neq j \\ m_{ii} &= \frac{1}{\sigma^2}\textstyle\sum_{i=1}^{n} X_{il}^2 + a_{ii},\ l = 1, 2\ 3, \cdots, k \end{aligned} \tag{11}$$

and $\boldsymbol{C}$ is a column vector of order $(k+1)$ with l^{th} element given as;

$$\begin{aligned} C_0 &= -2\left[\frac{1}{\sigma^2}\sum\nolimits_{i=1}^{n} y_i + \sum\nolimits_{j=0}^{k} a_{0j}\mu_j\right] \\ C_l &= -2\left[\frac{1}{\sigma^2}\textstyle\sum_{i=1}^{n} y_i X_{li} + \sum_{j=1}^{k} a_{lj}\mu_j\right],\ l = 1, 2\ 3, \cdots, k \end{aligned} \tag{12}$$

2.1 Estimation of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$

To estimate the parameters of the prior distribution of the regression parameters, we used jackknife samples as follows.

Let $\hat{\beta}_l = (\hat{\beta}_{0l}, \hat{\beta}_{1l}, \hat{\beta}_{2l}, \cdots, \hat{\beta}_{kl})$; $l = 1, 2\ 3, \cdots, n$ be the l^{th} jackknife estimate of the regression parameters from a given dataset which consist of a response variable Y and predictor variables $X_1, X_2, \cdots, X_k$. Then the estimate of the mean vector $\boldsymbol{\mu}$ of the random vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k)'$ is given as $\widehat{\boldsymbol{\mu}} = (\hat{\mu}_0,\ \hat{\mu}_1\ ,\ \hat{\mu}_2\ ,\ \cdots, \hat{\mu}_k)'$.

where

$$\hat{\mu}_j = \frac{1}{n}\textstyle\sum_{i=1}^{n} \beta_{ji}\ , j = 0,1, 2\ 3, \cdots, k \tag{13}$$

and an estimate of the covariance matrix of $\boldsymbol{\beta}$ given by

$$\widehat{\boldsymbol{\Sigma}}_{\beta} = \frac{1}{n-1}\textstyle\sum_{j=1}^{n}\big(\widehat{\boldsymbol{\beta}}_j - \widehat{\boldsymbol{\mu}}_j\big)\big(\widehat{\boldsymbol{\beta}}_j - \widehat{\boldsymbol{\mu}}_j\big)' = (\hat{a}_{ij}). \tag{14}$$

The estimate of the standard error of the i^{th} coefficient based on the Bayesian estimate is the square root of the i^{th} diagonal element of $\widehat{\boldsymbol{\Sigma}}_{\beta}$. That is

$$se(\hat{\beta}_i) = \sqrt{\hat{a}_{ii}}, \qquad i = 0, 1, 2, \dots, k \tag{15}$$

3 Results

To compare the LSM and the Bayesian method, we used a set of data based on a sample of size 63 from [2]. The data include one response variable, Sales Price, and four predictor variables namely, Square Feet, Rooms, Bedrooms, Age (See Appendix for data). The multiple regression model in equation (1) was fitted to the data to obtain the parameter vector $\boldsymbol{\beta}$ estimates through the Least Squares method.

Based on equations (13) and (14), jackknife estimates of the mean vector and covariance matrix of the random vector $\boldsymbol{\beta}$ were computed from the same data as follows respectively.

$\hat{\boldsymbol{\mu}} =$ (11.3763, 0.0539, -13.3180, -0.3815, 5.8404) and

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 4.0974 & -0.0009 & -0.5721 & -0.0022 & -0.1996 \\ -0.0009 & 0.0000 & -0.0001 & 0.0000 & -0.0004 \\ -0.5721 & -0.0001 & 0.4854 & 0.0002 & -0.1024 \\ -0.0022 & 0.0000 & 0.0002 & 0.0002 & -0.0014 \\ -0.1996 & -0.0004 & -0.1024 & -0.0014 & 0.1796 \end{bmatrix}$$

The inverse of $\hat{\boldsymbol{\Sigma}}$ is given by,

$$\hat{\boldsymbol{\Sigma}}^{-1} = (a_{ij}) = \begin{bmatrix} 16.4841 & 22011.8900 & 45.5908 & 641.2054 & 94.8122 \\ 22011.8940 & 30094728.4200 & 61474.6287 & 860629.5473 & 128436.9868 \\ 45.5908 & 61474.6300 & 128.9556 & 1786.3398 & 265.1956 \\ 641.2054 & 860629.5500 & 1786.3398 & 31627.9544 & 3755.6499 \\ 94.8122 & 128436.9900 & 265.1956 & 3755.6499 & 556.9152 \end{bmatrix}$$

Now based on the set of equations (11) an estimate of the posterior distribution of $\boldsymbol{\beta}$ is given by,

$$\hat{\boldsymbol{\Sigma}}_\beta = \begin{bmatrix} 3.8700 & -0.0008 & -0.5503 & -0.0023 & -0.1959 \\ -0.0008 & 0.0000 & -0.0002 & 0.0000 & -0.0004 \\ -0.5503 & -0.0002 & 0.4782 & 0.0002 & -0.1003 \\ -0.0023 & 0.0000 & 0.0002 & 0.0002 & -0.0014 \\ -0.1959 & -0.0004 & -0.1003 & -0.0014 & 0.1747 \end{bmatrix}$$

Also from the set of equation (12), we have an estimate of the vector $\boldsymbol{C}$ given as;

$$\hat{\boldsymbol{C}} = \begin{bmatrix} 1087.8880 \\ 1495893.4600 \\ 3019.4140 \\ 40087.1390 \\ 6369.1730 \end{bmatrix}$$

From equation (10), a Bayesian estimate of the parameter vector $\boldsymbol{\beta}$ is given as

$$\hat{\boldsymbol{\mu}}_\beta = \begin{bmatrix} 11.3775 \\ 0.0539 \\ -13.3181 \\ -\ 0.3814 \\ 5.8407 \end{bmatrix}$$

The estimates of the standard errors of the coefficients were also computed based on equation (15) and compared with those of the least squares standard errors as shown in Table 1.

3.1 Comparison of the LSM and Bayesian method

Table 1 shows the coefficient estimates and the corresponding standard errors for the Least Squares model and the Bayesian Model.

Table 1. Estimates of the coefficient of the two models

Variable name	Least squares		Bayesian	
	Coefficients	Standard error	Coefficients	Standard error
Intercept	11.3755	12.2752	11.3775	1.9672
Square. feet	0.0538	0.0085	0.0539	0.0015
Bedrooms	-13.3119	6.2096	-13.3181	0.6916
Age	-0.3813	0.1182	- 0.3814	0.0126
Rooms	5.8476	2.6646	5.8407	0.4180

The following table (Table 2) consists of variance results of the Least Squares model.

Table 2. Analysis of variance of the least squares model

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	4	55155.70	13788.92429	34.351239	1.0778E-14
Residual	58	23281.77	401.4098083		
Total	62	78437.47			

The analysis of variance for the multiple regression model gives an F statistics value of 34.3512 with a corresponding P-value of $1.0778e^{-14}$ which is significant at 5% significance level. The coefficient of determination (R^2) value from the LSM is 0.7032 indicating 70.32% of the variability in the response data is explained by the predictor variables.

The sum of squares due to error (SSE) of the Bayesian model is computed as follows,

$$\sum_{i=1}^{63}\left(Y_i - \hat{Y}_i\right)^2 = 23281.7800$$

Now our total sum of squares is $SST = \sum_{i=1}^{63}\left(Y_i - \overline{Y}\right)^2 = 78437.4660.$

The coefficient of determination of the Bayesian model is given by

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{23281.7800}{78437.4660}$$
$$= 0.7032$$

Clearly the coefficients of determination (R^2 values) for the two models are almost the same.

4 Discussion and Conclusion

From Table 1, it can be seen that estimated coefficients $\widehat{\boldsymbol{\beta}}$ are almost the same for the Least Squares model and the Bayesian model. In estimating the coefficient of determination $\boldsymbol{R^2}$, of the two fitted models, they both reported almost the same values.

This study reveals that, though the Least Squares method is just sufficient for estimating the coefficients of the regression parameters, the Bayesian estimates recorded comparatively very small standard errors; making the Bayesian method more robust. The use of additional information provided by the assumption of multivariate normal prior distribution of the $\boldsymbol{\beta}'\boldsymbol{s}$ accounted for the smaller standard errors of the Bayesian estimates.

Future studies may consider using Bootstrap estimates for the parameters of the prior distribution of $\boldsymbol{\beta}$ and consider a smaller data set to see whether same results will occur. It is anticipated that same finding will result in these situations. The main contribution of this work is to provide a means of using traditional Bayesian methods to estimate parameters of multiple regression coefficients under the assumption of multivariate normal prior distribution as against the existing simulation methods.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Bishop CM, Tipping ME. Bayesian regression and classification, advances in learning theory: Methods, models and applications. NATO Science Series III: Computer and Systems Sciences. 2003;190.

[2] Rossi PG, Allenby RM. Bayesian statistics and marketing. New York: John Wiley & Sons; 2005.

[3] Zellner A, Tobias J, Ryu HK. Bayesian method of moments (BMOM) analysis of parametric and semiparametric regression models. South African Statistical Journal. 1999;33:41-70.

[4] Zellner A. Bayesian method of moments (BMOM) analysis of mean and regression models. Springer New York. 1996;61-72.

[5] Zeithammer R, Lenk P. Bayesian estimation of multivariate-normal models when dimensions are absent. Quantitative Marketing and Economics. 2006;4(3):241-265.

[6] Poirier DJ. Bayesian interpretations of heteroskedastic consistent covariance estimators using the informed Bayesian bootstrap. Econometric Reviews. 2011;30(4):457-468.

[7] German A, Carlin BJ, Stern SH, Rubin BD. Bayesian data analysis. Florida: Chapman and Hall/CRC; 2003.

[8] Bowerman BL, O'Connell TR. Applied statistics. Von Hoffman Press Inc; 1997.

[9] Seber GAF, Wild CJ. Nonlinear regression. Wiley, New York; 2003.

[10] Van de Geer S. Least squares estimators. In Everitt BS, Howell DC, (eds.). Encyclopedia of statistics in behavioral science. Wiley and Sons. 2005;2:1041–1045.

Appendix A: Data used for the analysis

No	Sales price	Square feet	Bed room	Age	Rooms	No	Sales price	Square feet	Bed room	Age	Rooms
1	53.5	1008	2	35	5	**33**	63	1053	2	24	5
2	49	1290	3	36	6	**34**	60	1728	3	26	6
3	50.5	860	2	36	8	**35**	34	416	1	42	3
4	49.9	912	3	41	5	**36**	52	1040	2	9	5
5	52	1204	3	40	6	**37**	75	1496	3	30	6
6	55	1204	3	10	5	**38**	93	1936	4	39	8
7	80.5	1764	4	64	8	**39**	60	1904	4	32	7
8	86	1600	3	19	7	**40**	73	1080	2	24	5
9	69	1255	3	16	5	**41**	71	1786	4	24	8
10	149	3600	5	17	10	**42**	83	1503	3	14	6
11	46	864	3	37	5	**43**	90	1736	3	16	7
12	38	720	2	41	4	**44**	83	1695	3	12	6
13	49.5	1008	3	35	6	**45**	115	2186	4	12	8
14	105	1950	3	52	8	**46**	50	888	2	34	5
15	152.5	2086	3	12	7	**47**	55.2	1120	3	29	6
16	85	2011	4	76	9	**48**	61	1400	3	33	5
17	60	1465	3	102	6	**49**	147	2165	3	2	7
18	58.5	1232	2	69	5	**50**	210	2353	4	15	8
19	101	1736	3	67	7	**51**	60	1536	3	36	6
20	29.4	1296	3	11	6	**52**	100	1972	3	37	8
21	125	1996	3	9	7	**53**	44.5	1120	3	27	5
22	87.9	1874	2	14	5	**54**	55	1664	3	79	7
23	80	1580	3	11	5	**55**	53.4	925	3	20	5
24	94	1920	3	14	5	**56**	65	1288	3	2	5
25	74	1430	3	16	9	**57**	73	1400	3	2	5
26	69	1486	3	27	6	**58**	40	1376	3	103	6
27	63	1008	2	35	5	**59**	141	2038	4	62	12
28	67.5	1282	3	20	5	**60**	68	1572	3	29	6
29	35	1134	2	74	5	**61**	139	1545	3	9	6
30	142.5	2400	4	15	9	**62**	140	1993	3	4	6
31	92.2	1701	3	15	5	**63**	55	1130	2	21	5
32	56	1020	3	16	6						

Data from Bowerman and O'Connell, (1997)

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/13526