Progress in Nonlinear Dynamics and Chaos
Vol. 7, No. 1 & 2, 2019, 73-83
ISSN: 2321 – 9238 *(online)*
Published on 9 October 2019
www.researchmathsci.org
DOI: http://dx.doi.org/10.22457/pindac.81v7n1a7

Progress in Nonlinear Dynamics and Chaos

Comparison Some Methods to Estimation Parameters Gamma's Distribution by Using Simulation

Kareem K. Oazir

Department of Statistics, University of Sumer, Alrifaee, Iraq
Email: kareem.oazer@yahoo.com

Received 2 September 2019; accepted 9 October 2019

Abstract. A comparison of the estimation methods used and the comparison of the OLS and PES methods was performed with a distribution of one parameter. The simulation was used to compare the OLS method with the gamma distribution under the assumption that Bayes' is known to have a gamma distribution under the assumption that is not known. The estimation process for the parameters of the simple linear regression model is an important topic, despite the fact that it has been written about it through researches and studies that differ according to the methods used in the estimation process, whether this method is traditional or Bayes'.

The purpose of this study is to employ advance information about the parameters to be estimated based on the concept of the Pseudo theory of distribution of a single parameter in the estimation process and for the sizes of different samples. Therefore, a comparison was made between OLS method, the process of derivations to obtain the Pseudo-estimation formulas.

Keywords: Ordinary Least Squares Method, Bayes' Approach in the Estimation, estimation of marks in case O^2 is known, estimating the marks in case of O^2 is unknown

AMS Mathematics Subject Classification (2010): 97N20, 62N09

1. Introduction

In this particular side, we touch upon a brief idea about estimating the marks of simple linear declension, by using the classical methods and of them is (the regular minimized square method) OLS and inclusively the properties of this method, after that we go to dealing with the Bayes' analytic method for this model in estimating the marks which represent the basic target of this research, depending on previous information about the marks represented by Gamma distribution.

2. Ordinary least squares method OLS

It is a style of matching a straight line of an observation sample x. y and includes minimizing the total squares at the points diversions from the line to the least degree if possible , and the meaning of this is that to depending on calculating the values of unknown marks for model and Which make the total of random mistakes of the squares that are in its minimized end.

$$Min\sum(y_i-\hat{y}_i)^2$$

We can obtain the marks values on the following form:

$$\hat{\beta}_1=\frac{n\sum x_iy_i-\sum x_i\sum y_i}{n\sum x_i^2-(\sum x_i)^2}$$

$$\hat{\beta}_0=\bar{y}-b_1\bar{x}$$

To choose the statistical abstractness for the estimations of the declension marks, we need to know about the contrast of each of

$$\mathrm{var}(\hat{\beta}_0)=\overset{2}{\sigma}_u\frac{\sum X_i^2}{n\sum x_i^2} \tag{1}$$

$$\mathrm{var}(\hat{\beta}_1)=\overset{2}{\sigma}_u\frac{1}{\sum x_1^2} \tag{2}$$

3. Estimator's properties of OLS method

The ordinary least squares method estimators are the best unbiased estimators and being not biased

$$E(\hat{b})=b$$

Mean :

$$\text{Biased}=E(\hat{b})-b \tag{3}$$

As for describing the best unbiased linear estimation means it has the smallest contrast, sometimes the researcher may desire to accept some of the bias against a smaller contrast by minimizing the medium of (MSE) square mistake.

$$MSE(\hat{b})=\mathrm{var}\,\hat{b}+(\text{biased}\,\hat{b})^2 \tag{4}$$

4. Bayes' approach in the estimation

Bayes' approach focuses in estimation in its concept generally on employing prior information about unknown marks $\theta = [\theta_1, \theta_2 \ldots \theta_p]$ that are requested to be estimated, considering that these marks are random variables and are not fixed quantities which can be described by a form of contingent distribution, defined by the former function of contingency density (Prior p.d.f) and these information are recognized from the data and previous experiments or from the theory which controls these phenomena .

Whereas recent distribution function of the recent observation samples are under study, in them the value of random variable (y) for these observations is a distributive function that depends on θ and is symbolized with P(y/θ) named the Function) (Likelihood Bayes' estimation for these marks relies on the (Posterior p.d.f) which has been obtained by mixing the primary P.D.F of the marks with the likelihood function of observations , in which Larson defines the posterior P.d.f of the marks (θ)as the conditional function in the marks field (θ) and with existing of the recent sample marks and are expressed mathematically as the following: :

$$P(\theta\,|\,\mathrm{Y})\propto P(\theta).P(y/\theta) \tag{5}$$

At which ∝ is referring that the quantity is proportional, and for the purpose of clarifying Bays approach in estimation very clearly, after obtaining the posterior P.D.F for the vector of (θ) marks, and in which it's specified what is called the (Loss function) and that is symbolized usually by which (Mood) defines it as the real value function fulfilling two conditions:

$$L(\hat{\theta},\theta) \geq 0 \quad \forall \hat{\theta}, \forall \theta$$
$$L(\hat{\theta},\theta) = 0 \quad \forall \hat{\theta} = \theta \tag{6}$$

Therefore, the linear estimation by bays approach depends on finding out the $\hat{\theta}$ value that decreases the expectation of the loss.

It's worth mentioning that there are several kinds of loss functions, the most spread one and most used is the (Quadratic Loss Function) which is named in case of existing a vector with loss of (Weighted Squared Error Loss Function) expressed as the following:

$$L(\hat{\theta},\theta) = (\hat{\theta}-\theta)'C(\hat{\theta}-\theta) \tag{7}$$

Since C is representing the positive matrix of m*m dimension and is an equalized not random matrix, and the expectation for this case is as follows:

$$E\left[L(\hat{\theta},\theta)/y\right] = \int_{\Omega^*} L(\hat{\theta},\theta).P(\theta/y)d\theta \tag{8}$$

Estimating the marks values in the equation (11) by using bays approach and by relying on (squared error loss function weighted), it resembles the arithmetic mean for the (posteriors. P.D.F) f (θ/y) as it is clarified in the following:

$$E\left[L(\hat{\theta},\theta)/y\right] = E\left[(\hat{\theta}-\theta)/y\right]^2 \tag{9}$$

$$= \hat{\theta}^2 \int f(\theta/y)d\theta - 2\hat{\theta}\int \theta f(\theta/y)d\theta + \int \theta^2 f(\theta/y)d\theta \tag{10}$$

By taking the first derivative of the equation (9) in regard of $\hat{\theta}$ and equalizing it to zero for the purpose of obtaining the least loss expectation we have.

$$\frac{\partial^2 E\left[L\left(\hat{\theta},\theta\right)\backslash Y\right]}{\partial^2 \hat{\theta}^2} = 2 > 0 \tag{11}$$

This means that the critical point is eventually a local minimum that's why Bays approach of estimation is expecting the next distribution.

5. Previous probability density functions prior p.d.f

Using Bays approach in estimation requires the existence of first density function, thus marking out the type of posterior density function for the marks is one of the important subjects. As Zellner mentioned that determining the type of functions depends totally on the type of information previously acquired by the research, for it can be a contained information in the data of former samples, which have been provided through a practical approach, for the prior p.d.f that resembles this type of information are called (data based prior p.d.f) . and if there is former information provided through researching or by the result of theoretical considerations which have no relation with any data, whatever the former

sample is, the posterior p.d.f that is resembled in such an information called (NDB) (Non Data-Based Prior P.d.f) And we might have first p.d.f functions representing the provided information from other samples previously, and also from provided information and from theoretical considerations. In which it is possible to divide the posterior p.d.f as follows:
In case of not providing enough posterior information, Zellner has mentioned that Jeffery has suggested two rules to test posterior distribution at the following shape:

<u>First rule</u>

If the marks has value in infinite field from (-∞,∞) the posterior p.d.f of it will be taking an organized distribution (Uniform Distribution)

$$P(\beta)d\beta \propto d\beta \qquad -\infty < \beta < \infty \tag{12}$$

This type is considered one of the improper p.d.f types, by in case of merging it with the likelihood function; we would obtain a proper posterior distribution.

<u>Second rule</u>

If the marks in the field (0,∞) ، such as the normative diversion (σ)، for Jeffery has suggested to take the organized logarithm distribution, therefore the posterior p.d.f. of θ Be as follows:

$$f(\theta)d\theta \; \alpha \; d\theta \qquad\qquad -\infty < \theta < \infty \tag{13}$$

So in this formula (1-13) is being correspondent to the first rule of Jeffery

$$P(\sigma)d\sigma \propto \frac{1}{\sigma}d\sigma \tag{14}$$

Whereas the formula (14) is represented by the improper posterior p.d.f, this type of functions is called (Non informative prior p.d.f)

1. In the case of availability of enough posterior information about the marks and which is represented by restrictions about these marks, therefore we can use the same rules which Jeffery mentioned earlier, and which have been mentioned in the previous point (1), but in a restrictive marks field, that the field to be from (-∞) into (+∞)so we would have gained an organized distribution and restrictive about the marks, when this type of functions is being called as (Informative Prior p.d.f).
2. In case of using a recognized posterior p.d.f which through merging it with likelihood functions for the observation, it results a next p.d.f that is the same posterior p.d.f but with different features, therefore it would be called (Natural conjugate prior p.d.f).
3. There is a generalization in this research, and in case of existing more than one marks, for which one of the properties of this type of posterior p.d.f is being an invariant, thus Jeffery suggested to generalize what he had obtained in (1) which the function of posterior p.d.f is as follows:

$$P(\theta) \propto \left|\ln f\theta\right|^{\frac{1}{2}} \tag{15}$$

$$Inf_{\theta} = -E\left[\frac{\partial^2 Log\, P(\mathrm{Y}/\theta)}{\partial\theta_i \partial\theta_j}\right] \tag{16}$$

So the formula (16) is called (Fisher information Matrix) and the expectation has taken the percentage of y.

On the light of what has been said before, when estimating the marks of the declension model (2), and when there is an availability of posterior information, and by providing a likelihood distribution of the marks$_1 \hat{\beta}$ which is a gamma distribution with one marks (Gamma distribution).

We can estimate the model marks according to two assumptions and like the following:

6. Estimation of marks in case O² is known

Relying on function (16), the next posterior p.d.f can be found for the marks (β_0)via processing the integration calculus for the function (16) with regard to β_1 and as follows :

$$f\,(\beta_{\circ})=\sigma^{-(n+1)}\,(2\pi)^{-\frac{n}{2}}\,\lambda^2\int_{\circ}^{\infty}\beta_1\,e^{-\lambda\beta_1}\,\overline{e}^{\,-\frac{Q_1+Q_2}{2\sigma^{2}}}\;d\beta_1$$

$$=(2\pi)^{-\frac{n}{2}}\,\sigma^{-(n+1)}\,\lambda^2\,e^{-\frac{Q_1}{2\sigma^2}}\int_{\circ}^{\infty}\beta_1\,e^{-\frac{Q_2}{2\sigma^2}}\,e^{-\lambda\beta_1}\;d\beta_1$$

$$A=(2\pi)^{-\frac{n}{2}}\,\sigma^{-(n+1)}\;\lambda^2\,e^{-\frac{Q_1}{2\sigma^2}}$$

$$f\,(\beta_o)=A\int_{o}^{\infty}\beta_1\,e^{-\lambda B_1}\,e^{-\frac{Q_2}{2\sigma^2}}\;d\beta_1 \tag{17}$$

We can simplify Q_2 in the formula (17) in the following:

$$f\,(\beta_o)=A\int_{o}^{\infty}\beta_1\,e^{-\lambda\beta_1}\,e^{\frac{\sum x_i^2}{2\sigma^2}\beta_1^2+\frac{\beta_1}{2}\left(\hat{\beta}_1\sum x_1^2-\left(\beta_o-\hat{\beta}_o\right)\sum x_i\right)}d\beta_1 \tag{18}$$

And with adding and subtracting the final boundary of the square:

$$f\left(\beta_o\right)=Ae^{+\frac{1}{2\sigma^2\sum x_i^2}\left[\hat{\beta}_1\sum x_i^2-\left(\beta_o-\hat{\beta}_o\right)\sum x_i-\lambda\sigma^2\right]^2} \tag{19}$$

7. Estimating the marks in case of O² is unknown

Depending on the function, the next boundary likelihood function can be found (Zellner) [40] for β_o, β_1 (through integrating the function in respect to the marks (σ) as follows:

$$f\left(\beta_o,\beta_1\right)\propto\lambda^2\,\beta_1\;e^{-\lambda\beta_1}\int_{o}^{\infty}\sigma^{-(n+1)}\;e^{-\frac{Q}{2\sigma^2}}\;d\sigma^2 \tag{20}$$

After proceeding with mathematical operations, we can give the final picture of the next function for the marks (β_o, β_1)

$$f(\beta_o,\beta_1)\propto \lambda^2\,\beta_1\,e^{-\lambda\beta_1}\,Q^{-\frac{n}{2}}\frac{\Gamma(\frac{n}{2}-1)}{2^{-\frac{1}{2}(n-2)}} \qquad (21)$$

Depending on function (21) the next boundary of the likelihood function can be found pursuant to the marks (β_0) by making an integration of the function for the marks β_1 and as follows :

$$Q^{-\frac{n}{2}}=\left(\frac{VQ}{V}\right)^{-\frac{n}{2}} \qquad (22)$$

$$Q^{-\frac{n}{2}}=\exp\left[-\frac{n}{2}\log\left(\frac{VQ}{V}\right)\right] \qquad (23)$$

8. Comparisons

We will cover in this side the procedure of comparing between the used estimation methods and a procedure of comparison for each of the estimating approaches (OLS) and Bays Approach, on availability of distribution about one of the marks (Gamma Distribution), that has one marks which has used the stimulation to compare between OLS estimation method , and Bays approach . Under the availability of gamma distribution by assuming that σ^2 is known in Bays approach, with providing of gamma distribution under the assumption that σ^2 is unknown.

1-4 description for the special stimulation experiment of the research

In this stimulation experiment, random data with sample sizes will be used such as (15, 30, 68, and 90,140) an observation for the dependent variable (yi) in dependence of the hypothetical declension model as follows:

$$y_i=\beta_0+\beta_1 X_i+U_i \qquad (24)$$

The hypothetical values of the marks do represent the estimators of OLS method, for the data from the practical reality, they will later be defined, as for the marks distribution of)λ;α(, the researcher is suggesting five cases.

stat / paramtar	1	2	3	4	5
α	2	2	2	2	2
λ	1	2	2.75	3.75	4.25

Table 1: The hypothetical values of distribution marks that are employed in estimation

9. Analysing the stimulation results

For the repeated size of (1000), of an experiment for the hypothesized declension model, and for each case of the explained cases in table (1) A comparison has been made between the used methods in the estimation operation, which is represented by the minimal squared method (OLS), and by Bays Approach for the use of estimation, on availability of Gamma Distribution (BCK), in case of assuming that is known and unknown.

We can subsequently define the hypnotized marks of the models. The comparison has been made by calculating the proportional efficiency, that represents the average of square mistakes, of the marks estimation by Bays approach, used to estimate the availability of Gamma distribution under the assumption that o2 is known to the average of minimal squared estimators with their square mistakes. If the efficiency is equal to the real one, this proves that the two styles have the same efficiency, or if it was less than a real one then this would refer to that (BCK) is much efficient than (OLS) and vice versa .

As follows the explanations of the used symbols in the tables:

OLS: Ordinary least squares Method

IBCK: Bays Approach of estimation by providing Gamma distribution under the assumption that σ^2 is known.

2BCK: Bays Approach of estimation by providing Gamma distribution under the assumption that σ^2 is unknown.

Case one:

The following table (2) shows the average of square mistakes (Mse) for the estimators relying on the first case when $\alpha=2,\lambda=1$ and as follows

State	N / estimators		15	30	40	68	90	140
$\sigma^2$is known	$\hat{\beta}_0$	OLS	103.35	88.10	81.76	75.04	73.09	71.00
		1BCK	103.16	88.14	81.02	74.41	76.62	63.40
	$\hat{\beta}_1$	OLS	0.86	0.83	0.77	0.75	0.74	0.75
		1BCK	0.86	0.83	0.76	0.74	0.72	0.70
$\sigma^2$is unknown	$\hat{\beta}_0$	OLS	103.66	91.97	75.67	74.78	73.12	70.20
		2BCK	101.47	92.06	70.75	73.98	67.67	62.80
	$\hat{\beta}_1$	OLS	0.85	0.86	0.75	0.78	0.69	0.75
		2BCK	0.87	0.86	0.75	0.74	0.67	0.71

Table 2: The average of square mistakes for the marks within the first case

From the table (2) we noticed the following:
a-In case of assuming that (o2) is known.
b- For the estimator ($\hat{\beta}$) and for the sample sizes (15, 40, 68,140), the approach of IBCK showed that the average of square mistakes is less than OLS method.
c- For the estimator ($_1\hat{\beta}$) and for the sample sizes of (40-68—90-140), the IBCK showed that the average of square mistakes is less than OLS method.
1-In case of assuming that **(σ^2)** is unknown

a- For the estimator ($\hat{\beta}$) and for the sample sizes (15.40.68.90.140), the approach of 2BCK showed that the average of square mistakes is less than OLS method.
b- For the estimator ($_1\hat{\beta}$) and for the sample sizes of (40-68—90-140), the 2BCK showed that the average of square mistakes is less than OLS method.

As to the relative efficiency for the methods, in regard of OLS method within the first case, it has been clarified by table (3).

Estimators \ N		15	30	40	68	90	140
$\hat{\beta}_0$	1BCK	0.99	1.00	0.99	0.99	1.04	0.89
	2BCK	0.97	1.00	0.93	0.98	0.92	0.89
$\hat{\beta}_1$	1BCK	1.00	1.00	0.99	0.98	0.97	0.94
	2BCK	1.00	1.03	0.97	0.92	0.98	0.93

Table 3: The proportional efficiency for the estimation methods in regard of OLS method within the first case

From table (3) we can notice that following:

1- The estimator ($\hat{\beta}$) as well as the sample sizes of (30,140), 2BCK approach has showed itself to be more qualified than IBCK whereas IBCK has showed itself to be more efficient at the sample sizes of (15, 40, 68, and 90).
2- For the estimator ($_1\hat{\beta}$) 2BCK has showed more efficiency of estimation than the IBCK Method at the size of two samples (68-50).

As for the average of square mistakes (Mse) for the estimators) $_1\hat{\beta}_0$, $\hat{\beta}$ (within the second case, the following table (4) illustrates it and as follows:

State	n / estimators		15	30	40	68	90	140
$\sigma^2$is Known	$\hat{\beta}_{\square}$	OLS	103.90	93.05	83.60	75.04	72.919	71.50
		1BCK	103.89	88.90	81.69	71.41	69.30	65.01
	$\hat{\beta}_{\square}$	OLS	0.82	0.86	0.77	0.75	0.72	0.75
		1BCK	0.83	0.86	0.77	0.74	0.68	0.80
$\sigma^2$is unknown	$\hat{\beta}_{\square}$	OLS	103.66	91.97	80.34	72.87	71.93	73.17
		2BCK	102.47	91.99	82.60	65.52	70.30	65.29
	$\hat{\beta}_{\square}$	OLS	0.88	0.85	0.77	0.75	0.75	0.71
		2BCK	0.87	0.88	0.76	0.75	0.70	0.76

Table 4: Average values of square mistakes for the estimators within the second case

From table (4) we can notice the following:
1- In case of assuming that **(σ^2)** is known.
2- For the estimator ($\hat{\beta}$) and for the sample sizes (15, 30, 40, 68,140), the approach of IBCK showed that the average of square mistakes is less than OLS method.
3- For the estimator ($\hat{\beta}$) and for the sample sizes of (40-68—90), the IBCK showed that the average of square mistakes is less than OLS method.

D- In case of assuming that **(σ^2)** is unknown
1-For the estimator ($\hat{\beta}$) and for the sample sizes (15...68.90.140), the approach of 2BCK showed that the average of square mistakes is less than OLS method.
2-For the estimator ($\hat{\beta}$) and for the sample sizes of (15, 40-68—90-), the 2BCK showed that the average of square mistakes is less than OLS method.
Table (5) shows the proportional efficiency of the methods regarding to OLS method as the following:

n / estimators		15	30	40	68	90	140
$\hat{\beta}_0$	1BCK	0.99	0.95	0.97	0.95	0.95	0.90
	2BCK	0.98	1.03	1.02	0.89	0.97	0.89
$\hat{\beta}_1$	1BCK	1.01	1.00	0.99	0.89	0.93	1.05
	2BCK	0.99	1.03	0.98	0.99	0.94	1.05

Table 5: The proportional efficiency of the estimation methods, in respect of OLS method within the second case

From table (5) we can notice the following:

1- The estimator ($\widehat{\beta}$) and the sized samples of (15, 68,140) show how 2BCK is much efficient than IBCK method.

2- In the estimator ($\widehat{\beta_1}$) 2BCK method showed a more efficient marks of estimation than IBCK method, at sizes of (15, 40).

Table (6) shows the values of square mistakes average for the estimators (1$\widehat{\beta_0}$, $\widehat{\beta}$) within the third case as the following:

State	N estimators		15	30	40	68	90	140
σ^2 is known	$\hat{\beta}_0$	OLS	108.65	97.68	79.79	72.11	73.73	71.50
		1BCK	105.87	99.64	80.26	75.79	68.06	65.01
	$\hat{\beta}_1$	OLS	1.09	0.89	0.76	0.75	0.75	0.76
		1BCK	0.98	0.90	0.75	0.75	0.74	0.76
$\sigma^2$is unknown	$\hat{\beta}_0$	OLS	104.82	91.00	78.35	77.71	72.14	73.17
		2BCK	99.54	86.91	71.01	70.06	69.11	65.29
	$\hat{\beta}_1$	OLS	0.96	0.85	0.76	0.76	0.75	0.75
		2BCK	0.96	0.85	0.77	0.76	0.76	0.67

Table 6: Square mistakes average for the estimators within the third case.

From Table (6) we note the following:

A- Assuming (σ^2) is knowing

1 In case of presuming that o2 is known

1. For the estimator ($\widehat{\beta_0}$) and for sample sizes (15, 90,140), IBCK method showed a square mistakes average less than OLS method.
2. For the estimator ($\widehat{\beta_1}$), IBCK showed square mistakes are less than OLS at sizes of (15, 40, 68, and 90).
3. In case of presuming that o2 is unknown
4. Concerning the estimator ($\widehat{\beta_0}$) ,2BCK method showed a square mistakes average less than OLS method at all sizes of (15,30,40,68,90,140).

As for the table (7) shows the relative efficiency for the methods in respect of OLS method as the following:

estimators \ n		15	30	40	68	90	140
$\hat{\beta}_0$	1BCK	0.97	1.02	1.01	1.05	0.93	0.90
	2BCK	0.94	0.95	0.90	0.90	0.95	0.89
$\hat{\beta}_1$	1BCK	0.92	1.04	0.98	0.99	0.98	1.01
	2BCK	0.99	1.00	1.00	1.00996	1.01	0.88

Table 7: Relative efficiency for the estimation methods in respect of OLS method within the third case

From the table (7) we can notice the following:

1- The estimator ($\widehat{\beta_0}$), and sample sizes (15, 25, 50, 75,150), 2BCK method showed that it was more efficient than 1BCK method.

2- The estimator ($\widehat{\beta_1}$), and for sample sizes of (15, 50, 75,100), 1BCK showed an estimation marks efficiency much efficient than 2BCK

3- The square mistakes average for the estimators ($\widehat{\beta_0}$-$\widehat{\beta_1}$) is explained in table (8) within the fourth case.

1-4 results and guidelines

1. Relying on the five cases, we can deduce that the marks value of (α) does not equal to the marks value of λ that means ($\alpha \neq \lambda$).
2. Depending on the second and third case, we can deduce that the two ways have proved their efficiency to estimate the marks ($\widehat{\beta_1}$) for which their efficient equals to number one the real, or is more larger than the one at the sample size of (30), for which sample sizes have different effects on the efficient.
3. Recommendation: to use the Bays approach in the estimation process by using Gamma distribution for a number marks K.

To use Bays approach in estimation and specially when there is an availability of former information, being described as a likelihood distribution function, former and a natural associated, and using a former likelihood density function by depending on the subsequent samples, and under the assumption that the mark (σ) is unknown from the perspective of Bays school, and by comparing it with the used ways in this study.

REFERENCES

1. A. Tchourbanov, *Prior Distribution*, July 2, pKI 357, UNOmaha, USA (2002).
2. A.Gelman, Prior distributions for variance parameters in models (2005). http://www.stat.columbia.edu/~gelman/research/published/tau9.pdf
3. B. W. Lindgren, *Statistical Theory*, 3ed, Macmillan Publishing Co, Inc, New -York.
4. J.O.Berger, Statistical Decision Theory and Bayesian Analysis, Second Edition, Springer Verlag (1985).
5. G.E.P. Box and G.C.Tiao, *Bayesian in Ferrous in Statistical Analysis*, Addison–Wesley Publishing Company, Inc. (1973).
6. B.P.Carlin and T.A.Louis, *Bayesian Methods for Data Analysis*. 3rd ed. Boca Raton, FL: Chapman & Hall/CRC Press, (2008).