International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:10, No:9, 2016

An Estimating Parameter of the Mean in Normal Distribution by Maximum Likelihood, Bayes, and Markov Chain Monte Carlo Methods

Autcha Araveeporn

Abstract**—**This paper is to compare the parameter estimation of the mean in normal distribution by Maximum Likelihood (ML), Bayes, and Markov Chain Monte Carlo (MCMC) methods. The ML estimator is estimated by the average of data, the Bayes method is considered from the prior distribution to estimate Bayes estimator, and MCMC estimator is approximated by Gibbs sampling from posterior distribution. These methods are also to estimate a parameter then the hypothesis testing is used to check a robustness of the estimators. Data are simulated from normal distribution with the true parameter of mean 2, and variance 4, 9, and 16 when the sample sizes is set as 10, 20, 30, and 50. From the results, it can be seen that the estimation of MLE, and MCMC are perceivably different from the true parameter when the sample size is 10 and 20 with variance 16. Furthermore, the Bayes estimator is estimated from the prior distribution when mean is 1, and variance is 12 which showed the significant difference in mean with variance 9 at the sample size 10 and 20.

Keywords**—**Bayes method, Markov Chain Monte Carlo method, Maximum Likelihood method, normal distribution.

I. Introduction

NORMAL distribution is an important distribution in the field of statistics and is often used in tremendous data especially in social science. Most data are presented in terms of continuous probability distribution such as income, weight, and height of a person. The parameters of normal distribution consist of mean and standard deviation which is remarkably useful to explain any characteristics of a population. The mean determines the location of the center of population and standard deviation determines the dispersion from the mean of population.

Several methods of parameter estimation are common such as the moments method, the maximum likelihood method, the minimum chi-square method, the least square method, and the Bayes method. The estimators obtained from these methods have been shown well; e.g. unbiasedness, sufficiency, completeness, and minimum variance unbiased estimator.

In this paper, we interested in the maximum likelihood method because the estimator is shown in a class of minimum variance unbiased estimator [1]. The Bayes method depends on a prior probability distribution to estimate a posterior distribution which is obtained from a Bayes estimator. Moreover, the posterior distribution can be used with Markov Chain Monte Carlo (MCMC) method [2] to estimate MCMC estimator.

A. Araveeporn is with the Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand (e-mail: kaautcha@hotmail.com).

II. Methods of Parameter Estimation

The parameter estimation of the mean in normal distribution consists of the following three methods.

A. Maximum Likelihood (ML) Method

The ML method corresponds to many well-known estimations in statistics because it is easy to understand and calculate the estimators. The basic idea of ML estimation is to treat the likelihood function as a function of parameter, and find the value of parameter that maximizes it.

Suppose that we have the random variables $X_1, \ldots, X_n$ which assumed a normal distribution function depended on unknown parameter mean (μ), and variance (σ^2); however, our goals will approximate the mean. The probability density function of x_i depended on μ, σ^2 is written by

$$f(x_i|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\ \exp - \left\{ \frac{1}{2\sigma^2}(x_i-\mu)^2 \right\}. \quad (1)$$

The likelihood function is denoted as

$$L(\mu) = \prod_{i=1}^{n} f(x_i|\mu,\sigma^2). \quad (2)$$

The ML estimator is solved as:

$$\begin{aligned} L(\mu) &= (2\pi)^{-n/2}\sigma^{-2n}\exp - \left\{ \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \right\} \\ \ln L(\mu) &= -\frac{n}{2}\ln(2\pi) - 2n\ln\sigma\ - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \\ \frac{\partial \ln L(\mu)}{\partial\mu} &= -\frac{2\sum_{i=1}^{n}(x_i-\mu)(-1)}{2\sigma^2}\ =\ 0 \\ \mu &= \frac{\sum_{i=1}^{n}x_i}{n}\ =\ \bar{x}. \end{aligned}$$

The $\hat{\mu}_{ML}$ is ML estimator is written as

$$\hat{\mu}_{ML} = \frac{\sum_{i=1}^{n}x_i}{n} = \bar{x}. \quad (3)$$

B. Bayes Method

In Bayesian probability theory, if the posterior distributions are in the same distribution as the prior probability distribution, the prior and posterior will be called conjugate distributions, and the prior is called a conjugate prior for the likelihood function. In this case, the normal distribution is conjugate distribution with respect to a normal likelihood function: If the likelihood function is normal, choosing a normal prior over the mean will ensure that the posterior distribution is also normal. The resulting posterior distribution is

$$\begin{aligned} f(\mu|\sigma^2, x_i) &\propto L(\mu)g(\mu|\sigma^2) \\ &\propto (\sqrt{2\pi\sigma^2})^{-n}\exp - \left\{\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\} \\ &\times (\sqrt{2\pi\sigma_0^2})^{-1}\exp - \left\{\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right\} \end{aligned}$$

$$\propto \exp\left(-\frac{1}{2}\left(\frac{n\sigma_0^2+\sigma^2}{\sigma^2\sigma_0^2}\right)\left\{\mu^2 - 2\mu\left[\frac{n\bar{x}\sigma_0^2+\mu_0\sigma^2}{n\sigma_0^2+\sigma^2}\right]\right\}\right). \tag{4}$$

The parameter μ can write in form of normal distribution as [3]

$$\mu|\sigma^2, x_i \sim \text{Normal}\left(\frac{n\bar{x}\sigma_0^2+\mu_0\sigma^2}{n\sigma_0^2+\sigma^2}, \frac{\sigma^2\sigma_0^2}{n\sigma_0^2+\sigma^2}\right), \tag{5}$$

then the Bayes estimator or $\hat{\mu}_{Bayes}$ can be computed by

$$\hat{\mu}_{Bayes} = \frac{n\bar{x}\sigma_0^2+\mu_0\sigma^2}{n\sigma_0^2+\sigma^2}, \tag{6}$$

$$\hat{\sigma}^2_{Bayes} = \frac{\sigma^2\sigma_0^2}{n\sigma_0^2+\sigma^2}. \tag{7}$$

In this case, the prior is denoted as $\mu_0 = 1$ and $\sigma_0^2 = 12$.

C. Markov Chain Monte Carlo (MCMC) Method

The MCMC method was first introduced by [4] as a method to simulate values from posterior distribution which are developed from the Bayes method. The Gibbs sampling [5], [6] is a popular method which are generated values from the posterior distribution approximated as a MCMC estimator. Therefore, we carry out the WinBUGS program [7] which is a statistical software for Bayesian analysis to estimate MCMC estimator. In order to construct a Gibbs sampling from MCMC, the posterior distribution from Bayes method is used to calculate:

$$\begin{aligned} \mu|\sigma^2, x_i &\sim \text{Normal}\left(\frac{n\bar{x}\sigma_0^2+\mu_0\sigma^2}{n\sigma_0^2+\sigma^2}, \frac{\sigma^2\sigma_0^2}{n\sigma_0^2+\sigma^2}\right) \\ &\sim \text{Normal}(\hat{\mu}_{Bayes}, \hat{\sigma}^2_{Bayes}). \end{aligned} \tag{8}$$

Using this result, the Gibbs sampling algorithm proceeds as:

1) Set prior parameter: μ_0 and σ_0^2.
2) Set $\mu = \mu^{(t-1)}$.
3) Calculate $\hat{\mu}_{Bayes}$ and $\hat{\sigma}^2_{Bayes}$.
4) Generate μ from $Normal(\hat{\mu}_{Bayes}, \hat{\sigma}^2_{Bayes})$.
5) Set $\mu^{(t)} = \mu, t = 1, 2, \ldots, T$.

Finally the MCMC estimator is approximated by

$$\hat{\mu}_{MCMC} = \frac{1}{T}\sum_{t=1}^{T}\mu^{(t)}. \tag{9}$$

III. Simulation Study

The simulation study is to generate the data in terms of normal distribution with true parameters $\mu = 2$ and $\sigma^2 = 4, 9,$ and 16 at the sample sizes n = 10, 20,30, and 50. The data are generated 500 replications in each situation by R program [8]. To investigate the performance of ML, Bayes, and MCMC methods, these estimators are computed by

$$\begin{aligned} \hat{\mu}_{ML} &= \bar{x}, \\ \hat{\mu}_{Bayes} &= \frac{n\bar{x}\sigma_0^2+\mu_0\sigma^2}{n\sigma_0^2+\sigma^2}, \\ \hat{\mu}_{MCMC} &= \frac{1}{T}\sum_{t=1}^{T}\mu^{(t)}. \end{aligned}$$

Next, we obtain these estimators from 3 methods. The hypothesis testing is used to test the mean of estimator in normal distribution which is different from the true parameters. In this case, the hypotheses are

$$H_0 : \mu_{\hat{\mu}} = \mu \quad \text{and} \quad H_1 : \mu_{\hat{\mu}} \neq \mu$$

. The t statistic is computed as:

$$t = \frac{\bar{\hat{\mu}} - \mu}{s_{\hat{\mu}}/\sqrt{m}},$$

where $s_{\hat{\mu}} = \sqrt{\frac{\sum_{j=1}^{m}(\hat{\mu}_j - \bar{\hat{\mu}})^2}{m-1}}$, $df = m - 1$, m is a number of replications. For the level of significance at $\alpha = 0.05$, we will reject H_0 if $|t| > t_{\alpha/2, m-1}$.

IV. Results

The parameter estimations of normal distribution by ML, Bayes, and MCMC methods are given in Tables I-III. The first and the second columns of these tables present the sample sizes and the true parameters from simulated data. A mean, a standard deviation, a lower and an upper bound of 95% confidence interval are shown in the next four columns. The last two columns list the t statistics and p-values for hypothesis testing. The p-values of the ML and MCMC from the tables indicate that the means of the estimated parameters are different from the true parameters with μ = 2 , and σ^2 = 16 at the sample sizes n = 10 and 20. For Table II, the Bayes estimator shows the significant difference in the mean with μ = 2 , and σ^2 = 9 at the sample sizes n = 10 and 20.

Figs. 1-3 show the histograms of the estimated parameter with ML method which followed a normal distribution. The Bayes method shows the histogram in normal distribution at at $\sigma^2 = 4, 9,$ and 16 on Figs. 4-6. For MCMC method, the histograms follow a normal distribution at $\sigma^2 = 4, 9,$ and 16 on Figs. 7-9.

TABLE I
THE MEAN, STANDARD DEVIATION (SD), LOWER CONFIDENCE INTERVAL (LCI), UPPER CONFIDENCE INTERVAL (UCI), T STATISTICS (T), AND P-VALUES BY ML METHOD

n	$\mu = 2$	mean	S.D.	t	p-values
	$\sigma^2 = 4$	1.9799	0.6474	-0.6933	0.4844
n=10	$\sigma^2 = 9$	1.9830	0.9136	-0.4140	0.6790
	$\sigma^2 = 16$	2.1312	1.2436	2.3589	0.0187*
	$\sigma^2 = 4$	2.0141	0.4648	0.6806	0.4964
n = 20	$\sigma^2 = 9$	1.9678	0.6889	-1.0440	0.2970
	$\sigma^2 = 16$	2.1036	0.9016	2.5760	0.0104*
	$\sigma^2 = 4$	2.0189	0.3527	1.2013	0.2302
n = 30	$\sigma^2 = 9$	1.9979	0.5237	-0.0891	0.9290
	$\sigma^2 = 16$	1.9860	0.7354	0.4237	0.6719
	$\sigma^2 = 4$	1.9990	0.2846	-0.0673	0.9463
n = 50	$\sigma^2 = 9$	2.0112	0.3991	0.6305	0.5287
	$\sigma^2 = 16$	2.0250	0.6074	0.9215	0.3572

* indicated significance level at 5 %

TABLE II
THE MEAN, SD, LCI, UCI, T, AND P-VALUES BY BAYES METHOD

n	$\mu = 2$	mean	S.D.	t	p-values
	$\sigma^2 = 4$	1.9489	0.6259	-1.8233	0.0688
n=10	$\sigma^2 = 9$	1.9145	0.8531	-2.2396	0.0255*
	$\sigma^2 = 16$	2.000	1.1042	0.0018	0.9985
	$\sigma^2 = 4$	1.9976	0.4571	-0.1140	0.9093
n = 20	$\sigma^2 = 9$	1.9328	0.6634	-2.2642	0.0239*
	$\sigma^2 = 16$	2.0363	0.8458	0.9619	0.3365
	$\sigma^2 = 4$	2.0076	0.3488	0.4913	0.6234
n = 30	$\sigma^2 = 9$	1.9738	0.5116	-1.1414	0.2542
	$\sigma^2 = 16$	1.9444	0.7049	-1.7617	0.0787
	$\sigma^2 = 4$	1.9924	0.2827	-0.5935	0.5531
n = 50	$\sigma^2 = 9$	1.9964	0.3934	-0.2008	0.8409
	$\sigma^2 = 16$	1.9986	0.5921	-0.0507	0.9596

* indicated significance level at 5 %

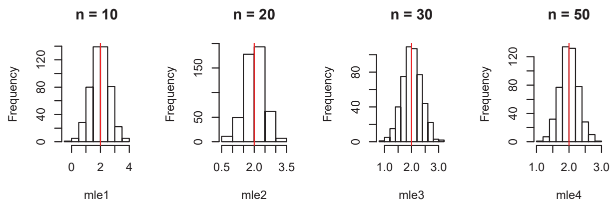

Fig. 1 Histograms of estimated parameters μ with ML method when $\mu = 2$ and $\sigma^2 = 4$

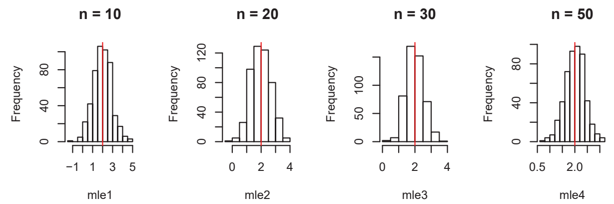

Fig. 2 Histograms of estimated parameters μ with ML method when $\mu = 2$ and $\sigma^2 = 9$

V. CONCLUSION

The mean of estimated parameter from ML and MCMC methods are not different from the true parameters in most cases except the large variance and small sample sizes. However, Bayes method are not different from the true parameters in most cases except the moderate variance and

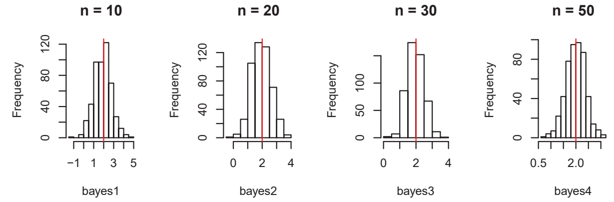

Fig. 3 Histograms of estimated parameters μ with ML method when $\mu = 2$ and $\sigma^2 = 16$

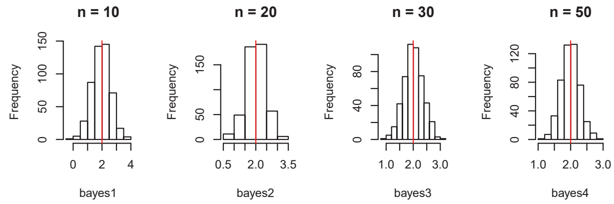

Fig. 4 Histograms of estimated parameters μ with Bayes method when $\mu = 2$ and $\sigma^2 = 4$

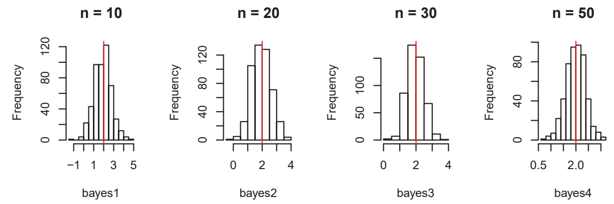

Fig. 5 Histograms of estimated parameters μ with Bayes method when $\mu = 2$ and $\sigma^2 = 9$

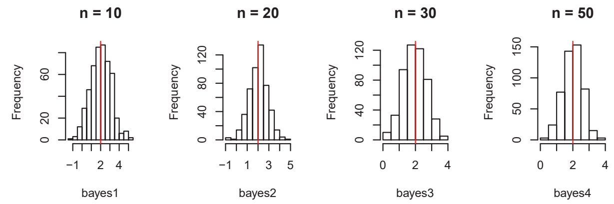

Fig. 6 Histograms of estimated parameters μ with Bayes method when $\mu = 2$ and $\sigma^2 = 16$

TABLE III
THE MEAN, SD, LCI, UCI, T, AND P-VALUES BY MCMC METHOD

n	$\mu = 2$	mean	S.D.	t	p-values
	$\sigma^2 = 4$	1.9698	0.6473	-1.0409	0.2986
n=10	$\sigma^2 = 9$	1.7040	0.9136	-0.7243	0.4692
	$\sigma^2 = 16$	2.1156	1.2436	2.0796	0.0380*
	$\sigma^2 = 4$	2.0084	0.4648	0.4596	0.6849
n = 20	$\sigma^2 = 9$	1.9601	0.6888	-1.2940	0.1963
	$\sigma^2 = 16$	2.0938	0.9017	2.3270	0.0203*
	$\sigma^2 = 4$	2.0149	0.3527	0.9454	0.3449
n = 30	$\sigma^2 = 9$	1.9923	0.5237	-0.3268	0.7439
	$\sigma^2 = 16$	1.9788	0.7355	-0.6434	0.5202
	$\sigma^2 = 4$	1.9958	0.2846	-0.3243	0.7458
n = 50	$\sigma^2 = 9$	2.0066	0.3991	0.3699	0.7116
	$\sigma^2 = 16$	2.0189	0.6074	0.6981	0.4854

* indicated significance level at 5 %

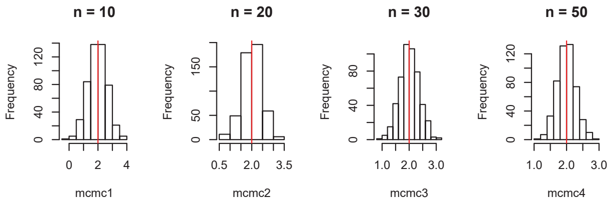

Fig. 7 Histograms of estimated parameters μ with MCMC method when $\mu = 2$ and $\sigma^2 = 4$

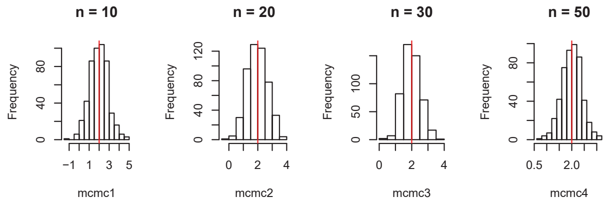

Fig. 8 Histograms of estimated parameters μ with MCMC method when $\mu = 2$ and $\sigma^2 = 9$

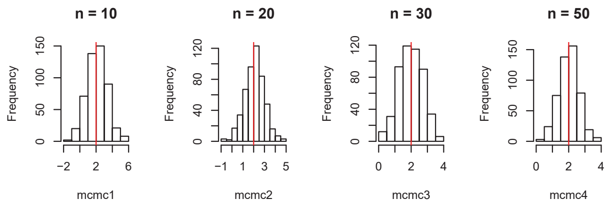

Fig. 9 Histograms of estimated parameters μ with MCMC method when $\mu = 2$ and $\sigma^2 = 16$

small sample sizes, but the Bayes method is depended on the parameters of prior distribution so the output may change in this case. If we did not identify the prior distribution, the ML method will work with good performance for estimating parameter of normal distribution.

References

[1] Rohatgi, V. K. and Saleh, E. (2001) *An Introduction to Probability and Statistics*. John Wiely & Sons, New York.

[2] Gilk, W., Richardson, S. and Spiegelhalter, D. (1996) *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

[3] Carlin, B. P. and Louis, T. A. (2009) *Bayesian Methods for Data analysis*. Florida: CRC Press Taylor & Francis Group

[4] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. and Teller, E. (1953) *Equations of State Calculations by Fast Computing Machine*. Journal of Chemical Physics, **21**, 1087-1092.

[5] Geman, S. and Geman, D. (1984) *Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration of Images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, **6** 721-741.

[6] Gelfand, A., Hills, S., Racine-Poon, A. and Smith, A. (1990) *Illustration of Bayesian inference in normal data models using Gibbs sampling*. Journal of the American Statistical Association, **85**, 972-985.

[7] Ntzoufran, I. (2009) *Bayesian Modeling Using WinBUGS*. John Wiely & Sons, New Jersey.

[8] R Development Core Team. (2004) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.