

วิธีคำนวณสำหรับการถอดออกโพรบิทเชิงอันดับด้วยวิธีโพลามเมโทโรพลิส

นายไชยชนันท์ ธาราสุข

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชีจุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2554

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)

เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)

are the thesis authors' files submitted through the Graduate School.

A Computational Method for Ordinal Probit Regression Base on  
Polar Metropolis

Mr. Chaiyanun Tharasuke

A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science Program in Statistics  
Department of Statistics  
Faculty of Commerce and Accountancy  
Chulalongkorn University  
Academic Year 2011  
Copyright of Chulalongkorn University

Thesis Title                    A Computational Method for Ordinal Probit Regression  
Base on Polar Metropolis  
By                                 Mr. Chaiyanun Tharasuke  
Field of Study                 Statistics  
Thesis Advisor                Assistant Professor Seksan Kiatsupaibul, Ph.D.

---

Accepted by the Faculty of Commerce and Accountancy, Chulalongkorn  
University in Partial Fulfillment of the Requirements for the Degree of Master of  
Science Program in Statistics

..... Dean of Faculty of Commerce and Accountancy  
(Associate Professor Pasu Decharin, Ph.D.)

#### THESIS COMMITTEE

..... Chairman  
(Associate Professor Kanlaya Vanichbuncha, Ph.D.)

..... Thesis Advisor  
(Assistant Professor Seksan Kiatsupaibul, Ph.D.)

..... Examiner  
(Akarin Phaibulpanich, Ph.D.)

..... External Examiner  
(Professor Anthony Hayter, Ph.D.)

ไชยະนันท์ ธาราสุข: วิธีคำนวณสำหรับการถดถอยโพรบิตเชิงอันดับด้วยวิธีโพลาร์เมโทรโพลิส. (A COMPUTATIONAL METHOD FOR ORDINAL PROBIT REGRESSION BASE ON POLAR METROPOLIS) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร. เสกสรร เกียรติสุไพบูลย์, 56 หน้า.

ในปัญหาการประมาณจุดตัดของตัวแบบโพรบิตเชิงอันดับโดยอาศัยแนวคิดเชิงเบสส์นั้น จะมีลักษณะของปัญหาเป็นการแจกแจงสะสมของการแจกแจงปกติที่ถูกจำกัด ซึ่งการใช้วิธีการสุ่มตัวอย่างแบบกิบส์ในปัญหานี้ทำให้การลู่เข้ามีความช้า เนื่องจากลักษณะของการแจกแจงความน่าจะเป็นมีลักษณะเป็นรูปทรงเรขาคณิตแบบกรวย ในงานวิจัยนี้จึงได้นำเสนอวิธีลูกโซ่มาร์คอฟมอนติคาร์โลแบบใหม่ซึ่งอาศัยการดำเนินการพิกัดเชิงขั้วบนวิธีเมโทรโพลิส ซึ่งประสิทธิภาพของวิธีการวัดจากจำนวนรอบจนกระทั่งทุกจุดตัดลู่เข้า โดยอาศัยการประยุกต์วิธีการดังกล่าวในข้อมูลอันดับความน่าเชื่อถือ

ภาควิชา..... สถิติ..... ลายมือชื่อนิสิต.....  
 สาขาวิชา..... สถิติ..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....  
 ปีการศึกษา..... 2554.....

## 5381780826: MAJOR STATISTICS

KEYWORDS: Ordinal Probit Regression / Cutpoints / Gibbs Sampler / Polar Metropolis (PM)

CHAIYANUN THARASUKE: A COMPUTATIONAL METHOD FOR ORDINAL PROBIT REGRESSION BASE ON POLAR METROPOLIS.

ADVISOR: ASST. PROF. SEKSAN KIATSUPAIBUL, Ph.D., 56 pp.

The problem of estimating the cutpoints of an ordinal probit model in a Bayesian setting is a problem of constrained cumulative normal distribution. The Gibbs sampler method, applied to this problem, demonstrated a slow convergence rate due to the conical geometry of the support of the distribution. This research presents a new MCMC method based on the Metropolis algorithm operating on the polar coordinate of the problem. The efficiency of both algorithms is measured based on the number of iterations necessary until all cutpoints converge. The algorithm is applied to a credit rating data set to demonstrate its efficiency.

Department:..... Statistics ..... Student's Signature.....

Field of Study:..... Statistics ..... Advisor's Signature.....

Academic Year:..... 2011 .....

## ACKNOWLEDGEMENTS

This thesis could not successfully completed without the kindness of my advisor, Assistant Professor Seksan Kiatsupaibul, Ph.D., who gave the best advice, appreciate suggestion, close supervision and be guidance of this thesis since begin until complete. And many thanks for the committees, Associate Professor Kanlaya Vanichbuncha, Ph.D. (Chairman), Akarin Phaibulpanich, Ph.D. (Examiner) and Professor Anthony Hayter, Ph.D. (External Examiner) for all comment and best recommendation.

I would like to special thank for my best friend, Doctor Natchalee Srimaneekarn, D.D.S, who stand beside me in difficult time and for everything, which she always gives me.

Finally, my graduation would not be achieved without best wish from my family.

## CONTENTS

	<b>Page</b>
<b>ABSTRACT IN THAI</b> .....	iv
<b>ABSTRACT IN ENGLISH</b> .....	v
<b>ACKNOWLEDGEMENTS</b> .....	vi
<b>CONTENTS</b> .....	vii
<b>LIST OF TABLES</b> .....	ix
<b>LIST OF FIGURES</b> .....	x
<b>CHAPTER I INTRODUCTION</b> .....	1
1.1 Background and rationale .....	1
1.2 Definition .....	2
1.2.1 Latent variable .....	2
1.2.2 Full conditional probability .....	2
1.3 Objectives .....	2
1.4 Scope of study .....	2
1.5 Significance of the research .....	3
1.6 Methodology .....	3
<b>CHAPTER II LITERATURE REVIEW</b> .....	5
2.1 Ordinal probit model .....	5
2.2 Markov Chain Monte Carlo .....	6
2.2.1 Monte Carlo Integration .....	7
2.2.2 Markov Chains .....	7
2.3 Metropolis-Hastings sampling .....	8
2.4 Gibbs sampler .....	9
2.5 Polar Metropolis .....	10
2.6 Consistent Batch Means .....	11
2.7 Related literature .....	11

	<b>Page</b>
<b>CHAPTER III METHODOLOGY</b> .....	12
3.1 Data of Research.....	13
3.1.1 Credit ratings of companies in Thailand in 2010.....	13
3.1.2 Simulated data using Polar Metropolis.....	13
3.1.3 Simulated data using Gibbs sampler.....	14
3.2 Cutpoints estimation using Polar Metropolis.....	15
3.3 Cutpoints estimation using Gibbs sampler.....	16
3.4 The criterion for the performance and procedures comparison.....	16
<b>CHAPTER IV RESULTS</b> .....	18
4.1 Results of comparison for credit rating data.....	18
4.2 Results of comparison for 2 cutpoints of simulated data.....	21
4.3 Resultsof comparison for 3 cutpoints of simulated data.....	23
4.4 Resultsof comparison for 4 cutpoints of simulated data.....	25
4.5 Resultsof comparison for 5 cutpoints of simulated data.....	27
<b>CHAPTER V CONCLUSION AND DISCUSSION</b> .....	29
5.1 Conclusion.....	29
5.2 Discussion and further research.....	31
<b>REFERENCES</b> .....	33
<b>APPENDIX</b> .....	34
<b>BIOGRAPHY</b> .....	56



## LIST OF TABLES

<b>Table</b>	<b>Page</b>
1 Mean and half-width of the first cutpoint ( $\gamma_1$ ) and the second cutpoint ( $\gamma_2$ ) using Polar Metropolis and the Gibbs sampler in credit rating data ( $n_{iters} = 41,000$ ).....	18
2 Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 2 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 25,000$ ).....	22
3 Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 3 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 25,000$ ).....	24
4 Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 4 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 25,000$ ).....	26
5 Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 5 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 50,000$ ).....	28

## LIST OF FIGURES

<b>Figure</b>	<b>Page</b>
1 The average mean plot of the first cutpoint from Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints regarding The credit ratings of Thai companies in 2010.....	19
2 The average mean plot of the first cutpoint from Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints regarding The credit ratings of Thai companies in 2010.....	20
3 The graph shows number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 2 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.....	21
4 The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints for simulated data	22
5 The graph shows number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 3 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.....	23
6 The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 3 cutpoints for simulated data	24
7 The graph shows number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 4 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.....	25
8 The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 4 cutpoints for simulated data	26

<b>Figure</b>	<b>Page</b>
9 The graph shows number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 5 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.....	27
10 The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 5 cutpoints for simulated data	28
11 The graph plots acceptance rate versus number of cutpoints at 50 data points.....	30
12 The graph shows the proposal density is uniform distribution and The full conditional distribution of cutpoints in the case of 2 cutpoints	31

## CHAPTER I

### INTRODUCTION

#### 1.1 Background and rationale

Ordinal probit model is a widely recognized statistical model for ordinal response data. McCullagh (1980) was among the first who studied this model. He proposed structural models appropriate for measurements on a purely ordinal scale that could be analyzed with different cumulative link functions such as logit and complementary log-log. McCullagh (1980) also provide a parameter estimation method based on the iteratively reweighted least squares.

This model was re-introduced by Albert and Chib (1993) as a Bayesian statistical model. Albert and Chib (1993) employed the Gibbs sampler and data augmentation to estimate cutpoints. The results showed slow convergence. It was noted that high autocorrelation in the MCMC sample for the cutpoints may also seriously affect the convergence of  $\beta$ . Cowles (1995) improved convergence rate of the Gibbs sampler by using the Metropolis-Hasting algorithm in sampling both cutpoints and the latent variables together. However, it proved to be difficult to tune.

Chiarawongse et al. (2011) created an algorithm that was an adaptation of Metropolis-Hasting and operated on polar coordinates (Polar Metropolis or PM) to compute expected return on assets with qualitative input. The posterior density of the returns is a constrained normal distribution. The results showed that Polar Metropolis was more efficient than the Gibbs sampler and Hit-and-run sampler.

A Bayesian ordinal probit model has a similar structure to that of Chiarawongse et al. (2011). The posterior distribution of parameters is also a constrained distribution but concerns different type of density. Therefore, Polar Metropolis algorithm should perform well for a Bayesian ordinal probit model.

The purpose of this paper is to apply Polar Metropolis for Bayesian parameter estimation of the ordinal probit model and to compare efficiency between Polar Metropolis and the Gibbs sampler as measured by the average number of iterations required for the algorithm to converge to the solution.

## 1.2 Definition

### 1.2.1 Latent variable

**Latent variables** (hidden variables) are not directly observed variables (as opposed to observable variables), but are inferred by preference through a statistical model from other variables that are observed.

### 1.2.2 Full conditional probability

Suppose we wish to sample a random vector  $\theta = [\theta_1, \dots, \theta_d]^T$  according to target probability density function ( $\pi'(\theta)$ ) of posterior distribution  $\pi$ . Let  $\pi'(\theta_i|\theta_{-i}) = (\theta_1, \theta_2, \dots, \theta_{i+1}, \theta_{i+1} \dots \theta_d)$  represent the full conditional probability density function of  $i$  - th component,  $\theta_i$ , given the other components  $\theta_1, \theta_2, \dots, \theta_{i+1}, \theta_{i+1} \dots \theta_d$ .

## 1.3 Objective

To apply Polar Metropolis for Bayesian parameters estimation to the ordinal probit model and to compare efficiency between Polar Metropolis and the Gibbs sampler.

## 1.4 Scope of Study

1. Compare efficiency between Polar Metropolis and the Gibbs sampler.
2. Measure performance by the average number of iterations which are required until all cutpoints converge.

3. Conduct analyzes on the ordinal probit model.
4. Analyze a set of real data: a 2010 credit rating of Thai company data set that received by an international credit rating agency.
5. Analyze simulated data of size  $N= 50, 100, 250, 500$  and  $1000$  with 2, 3, 4 and 5 cutpoints. The simulation environment is the following.
  - 5.1 The first independent variable  $(x_1) \sim N(0.2, 0.03)$
  - 5.2 The second independent variable  $(x_2) \sim N(0.2, 0.03)$
  - 5.3  $x_1$  correlate with  $x_2$  by 0.3
  - 5.4 Regression coefficient of  $x_1 (\beta_1)$  equal to -0.73
  - 5.5 Regression coefficient of  $x_1 (\beta_2)$  equal to -9.15
  - 5.6 Cutpoint vectors  $(\tilde{\gamma})$  are at intervals -2.2 to -0.6 for 2, 3 and 4 cutpoints and -3 to 1 for 5 cutpoints
  - 5.7 Each batchis equal to 500 iterations

### 1.5 Significance of the study

This research aims to provide a more efficient computational method for Bayesian parameters estimation of the ordinal probit model.

### 1.6 Methodology

1. Review the literature.
2. Use a 2010 credit rating of Thai company data set for Bayesian parameters estimation of an ordinal probit model based on the Gibbs sampler.
3. Use a 2010 credit rating of Thai company data set for Bayesian parameters estimation of an ordinal probit model based on Polar Metropolis (PM).
4. Use simulated data for Bayesian parameters estimation of an ordinal probit model based on the Gibbs sampler.
5. Use simulated data for Bayesian parameters estimation of an ordinal probit model based on Polar Metropolis (PM).

6. Compare efficiency between the Polar Metropolis and the Gibbs sampler as measured by the average number of iterations required for the algorithm to converge to the solution.
7. Draw and discuss conclusions.

## CHAPTER II

### LITERATURE REVIEW

#### 2.1 Ordinal probit model

Response variable of ordinal probit model is ordinal data. Thus, let  $Z = (z_1, z_2, \dots, z_n)'$  is denoted by a  $n \times 1$  vector of  $n$  response variables, which gets one of the  $k$  ordered values  $j = 0, \dots, k$ ,  $X_i$  is vector of independent variables and  $\beta$  is a vector of regression coefficients that there are equal to  $p \times 1$  dimensions. Let  $\gamma_j$  are ordered cutpoints and  $\varepsilon \sim (0,1)$ , the probability of observing  $z_i = j (0 \leq j \leq k)$  conditional on  $\beta$  and  $\gamma$  is given by

$$P(z_i = j | \beta, \gamma) = \Phi(\gamma_j - X_i^T \beta) - \Phi(\gamma_{j-1} - X_i^T \beta)$$

Which defines the link  $\Phi$  is denoted by the normal c.d.f. for  $i = 1, 2, \dots, n$  and  $j = 0, \dots, k$  where  $-\infty = \gamma_0 < \gamma_1 < \gamma_2 < \dots < \gamma_k = \infty$ . Thus, Albert and Chib (1993) incorporated latent variables or hidden variables  $(y^*) = (y_1^*, y_2^*, \dots, y_n^*)'$  to extra parameters for running the Gibbs sampler that depends on a vector of covariates  $x_i$  through this model:

$$y_i^* = X_i^T \beta + \varepsilon_i$$

Such that  $z_i = j$  iff  $\gamma_{j-1} \leq y_i^* \leq \gamma_j$  for  $j = 1, \dots, k-1$  and we assume  $\beta \sim \text{prior}$  and  $\gamma \sim (\text{prior flat})$  indicator function of the ordering of the cutpoints. Hence, the joint posterior distribution for the latent variables  $(y^*)$  and other parameters is given by:

$$\pi(\beta, \gamma, Y^* | Z) \propto \left\{ \prod_{i=1}^n [N(y_i^* | x_i^T \beta, 1) I\{\gamma_{j-1} < y_i^* \leq \gamma_j\} I_{\{\gamma_1 \leq \dots \leq \gamma_{k-1}\}}] \right\}$$

Hence, the full conditional distribution for  $y_i^*$  and  $\beta$  are:

$$P(y_i^* | \beta, \gamma, z_i) \propto (x_i^T \beta, 1) I(\gamma_{j-1} < y_i^* \leq \gamma_j)$$



Thus, the distribution of  $y^*$  is truncated normal distribution, and the distribution of regression coefficients is multivariate normal distribution, given by

$$P(\beta|z, y^*) = ((X^T X)^{-1} X^T Y^*, (X^T X)^{-1})$$

If there are hidden variables  $y^*$ , which are made the full conditional distribution for each  $\gamma_j$  is uniform distribution, given by

$$\begin{aligned} & (\gamma_j | z, y^*, \beta, \{\gamma_i, i \neq j\}) \\ &= [\max(\max\{y_i^*: z_i = j\}, \gamma_{j-1}), \min(\min\{y_i^*: z_i = j + 1\}, \gamma_{j+1})] \end{aligned}$$

## 2.2 Markov Chain Monte Carlo (MCMC)

In the case of the posterior distribution is not conjugate or high dimensional integration for Bayesian parameter estimation, a computational for posterior distribution from the density is very difficult. Hence, Markov Chain Monte Carlo (MCMC) is presented for fixing this problem. This method attempts to simulate straight draws from the distribution of interest which is very difficult, which the MCMC algorithm is assigned for sampling from the posterior distribution that are in one or more dimensions and moves around that posterior distribution. The idea of the MCMC comes from the concept of the second part. The first is the Monte Carlo which use to approximate posterior distribution and the second is the Markov chains, which infer to a new sampling points form their posterior distribution given by the previous sampling points. In each round of the process to create new value of the Markov chain. Details of this technique, Markov chain Monte Carlo, are as follows.

### 2.2.1 Monte Carlo Integration

We assume an arbitrary parameter ( $\theta$ ) has a posterior density ( $\pi(\theta|y^*)$ ). And we are interested in especially  $g(\theta)$  which is the real-value function on state space. Hence, the posterior expectation of  $g(\theta)$  is given by:

$$E(g(\theta)|y^*) = \int g(\theta) \pi(\theta|y^*) d\theta$$

Where the set  $\theta_1, \dots, \theta_N$ , independent and identically distributed, draws from the posterior density. Hence, Monte Carlo Integration is used to approximate posterior expectation from sample mean:

$$\bar{g} = \frac{\sum_{i=1}^N g(\theta_j)}{N}$$

And the Monte Carlo standard error is given by

$$SE_{\bar{g}} = \sqrt{\frac{\sum_{i=1}^N (g(\theta_j) - \bar{g})^2}{(N-1)N}}$$

Thus, if the sample points are appropriate generated from exact posterior distribution, the Monte Carlo Integration is efficiency. (Albert, 2009).

### 2.2.2 Markov Chains

Suppose, we require to simulate random vector  $\theta$  which has p.d.f.  $f$  on space  $\zeta \subset \mathfrak{R}^d$ . Markov Chains are simulated sequence of vector  $\{\theta_n; n = 0, 1, 2, \dots\}$  from transition probability  $P$ , as

$$P(u, v) = \Pr(\theta_{n+1} = v | \theta_n, \theta_{n-1}, \dots, \theta_0) = \Pr(\theta_{n+1} = v | \theta_n = u) \quad u, v \in \theta$$

Where given  $(\zeta, S)$  is measure space and a sequence of random parameters  $\{\theta_n; n = 0, 1, 2, \dots\}$  is called Markov Chains if for every  $A \in S$  it satisfies the Markov property:

$$\Pr(\theta_{n+1} \in A | \theta_n, \theta_{n-1}, \dots, \theta_0) = \Pr(\theta_{n+1} \in A | \theta_n) = P(\theta_n, A)$$

where  $P(\theta_n, A)$  has two properties.

1. For every  $u \in \mathcal{S}$ ,  $P(u, \cdot)$  is probability measure over  $(\mathcal{S}, \mathcal{S})$ .
2. For every  $A \in \mathcal{S}$ ,  $P(\cdot, A)$  is measurable.

The function  $P$  is called the one-step transition probability where given (one-step transition matrix)  $P = [p_{ij}]$ . Thus, the initial states do not change transition probability. Hence, random parameter vector of Markov Chains  $(\theta_1, \dots, \theta_k, \dots)$  by transition probability from state  $\theta_n$  to state  $\theta_{n+1}$  depends only on current state  $(\theta_n)$  not on previous states  $(\theta_1, \dots, \theta_{n-1})$ .

If Markov Chains  $\{\theta_n; n = 0, 1, 2, \dots\}$  have irreducible and aperiodic property (see appendix A.3) where reversible Markov Chains with respect to  $\pi$ . Hence, Markov Chains converge to target distribution  $\pi$  with in total variation or it means  $\theta_n$  can be approximate to  $\pi$ .

### 2.3 Metropolis-Hastings sampling

In Bayesian parameter estimation, we require to generate parameter  $(\theta)$  from probability density function  $(\pi(\theta|y^*))$  of posterior distribution  $(\pi)$ . Hence, Metropolis-Hastings generates sequence of draws from proposal density  $q(\theta_c|\theta_i)$ , and then sample points are filtered through below:

1. Initialize any parameter  $(\theta_0)$  where  $\pi(\theta_0) > 0$ .
2. Given the current state  $\theta_i$ , generate candidate parameter  $(\theta_c)$  from proposal density  $q(\theta_c|\theta_i)$  which is a probability of returning a value of  $\theta_c$  where given a previous value of  $\theta_i$ .

3. To compute the ratio that given by  $R = \frac{\pi(\theta_c)q(\theta_i|\theta_c)}{\pi(\theta_i)q(\theta_c|\theta_i)}$  with acceptance probability  $\alpha = \min(1, R)$ .

4. To compare  $\alpha$  with generate  $U \sim U(0,1)$  and deliver

$$\theta_i = \begin{cases} \theta_c & \text{if } U \leq \alpha \\ \theta_i & \text{otherwise,} \end{cases}$$

5. Let  $i = i+1$  and repeat step 2.

This distribution can converge to target distribution from idea of the MCMC.

## 2.4 Gibbs Sampler

The Gibbs sampler can be viewed as a particular instance of the Metropolis – Hastings algorithm in which random parameter is always accepted ( $\alpha = 1$ ) (every proposal density is accepted). The idea behind the Gibbs sampler is that all random parameter can direct draws from the full conditional distribution (sampled from one by fixed another). Suppose, we interest in parameter  $\theta = (\theta_1, \dots, \theta_d)$  where joint posterior distribution of  $\theta$  which denotes by  $\pi(\theta|y^*)$ , which is sometimes high dimensional integration. The Gibbs sampler's algorithm is as follow:

1. Initialize any parameter  $\theta_0 = [\theta_1, \dots, \theta_d]^T = [\theta_{1,0}, \dots, \theta_{d,0}]^T$  and  $n = 0$

2. To generate  $\theta_{h,n+1}$  from  $\pi'(\theta_h|\theta_{-h,n})$  where

$$\begin{aligned} \theta_{-h,n} &= [\theta_1, \dots, \theta_{h-1}, \theta_{h+1}, \dots, \theta_d]_n^T \\ &= [\theta_{1,n}, \dots, \theta_{h-1,n}, \theta_{h+1,n}, \dots, \theta_{d,n}]^T \end{aligned}$$

3. Let  $\theta_{n+1} = [\theta_{1,n}, \dots, \theta_{h,n+1}, \dots, \theta_{d,n}]^T$  and add  $n$  to  $n+1$ . Repeat step 2.

## 2.5 Polar Metropolis

Polar Metropolis was presented by Chiarawongse et al. (2011) which use to compute expected return on assets with qualitative input, that posterior expectation of its is truncated normal distribution. Polar Metropolis's algorithm is as follow:

1. Given  $\|\tilde{\gamma}^i\|$  is the length of  $\tilde{\gamma}^i$ , the current direction is given by  $\varpi = \frac{\tilde{\gamma}^i}{\|\tilde{\gamma}^i\|}$ .  $L$  be a line segment is defined by  $L = \{li: li = \tilde{\gamma}^i + s\varpi, s \in \mathcal{R}\} \cap R$ , and let  $s$  is the surface of the intersection between the hypersphere with radius  $\|\tilde{\gamma}^i\|$  and the polytope  $R$ .
2. While half-width is more than the precision.
3. Randomly select either an L-move or an S-move by a coin toss.

L move: random sampling  $\tilde{\gamma}$  from the conditional density  $g(\cdot | \tilde{\gamma} \in L)$  and accept it as  $\tilde{\gamma}^{i+1}$  with probability

$$\min\left\{1, \left(\frac{\|\tilde{\gamma}^{i+1}\|}{\|\tilde{\gamma}^i\|}\right)^{N-1}\right\}. \text{ Otherwise, let } \tilde{\gamma}^{i+1} = \tilde{\gamma}^i.$$

S move: random sampling  $\tilde{\gamma}$  from the uniform distribution on  $S$  and accept it as  $\tilde{\gamma}^{i+1}$  with probability  $\min\left\{1, \frac{f(\tilde{\gamma}^{i+1})}{f(\tilde{\gamma}^i)}\right\}$ . Otherwise, let  $\tilde{\gamma}^{i+1} = \tilde{\gamma}^i$ .

## 2.6 Consistent Batch Means

The posterior mean  $E_{\pi}(g(\theta))$  is estimated by using Markov Chain Monte Carlo (MCMC) where  $\theta$  depends on posterior distribution  $\pi$ . Hence, the methods for estimating are as follow:

1. To generate Markov Chain  $\{\theta_i, i=1,2,\dots,n\}$  which have the total iterations  $(n = \text{batchsize} * \text{simround})$ , where *simround* is total group of sample points and *batchsize* is size of batch. Thus, we can estimate sample mean of group  $k - th$  , as

$$\bar{z}_k = \frac{\sum_{i=(k-1)*\text{batchsize}+1}^{k*\text{batchsize}} g(\theta_i)}{\text{batchsize}} \text{ for } k=1,\dots, \text{simround}.$$

2. Let  $\hat{\sigma}_g^2 = \frac{\text{batchsize}}{\text{simround}-1} \sum_{k=1}^{\text{simround}} (\bar{z}_k - \bar{g}_n)^2$ , where  $\bar{g}_n = \frac{1}{n} \sum_{i=1}^n g(\theta_i)$

If the *batchsize* and the number of consistent batch mean (*ncbm*) are allowed to increase as the overall length of the simulation increases by setting  $\text{ncbm} = \lfloor \sqrt{n} \rfloor$  and  $\text{batchsize} = \lfloor n/\text{ncbm} \rfloor$ . Hence, Monte Carlo standard error of  $\bar{g}_n$  is  $\frac{\hat{\sigma}_g}{\sqrt{\text{ncbm}}}$ .

Thus, the half-width of the interval is given by  $t_{\frac{\alpha}{2}, \text{ncbm}-1} \frac{\hat{\sigma}_g}{\sqrt{\text{ncbm}}}$ .

## 2.7 Related literature

Chiarawongse et al. (2011) presented Polar Metropolis for compute expected return on assets with qualitative input, which the posterior expectation of return is truncated normal distribution.

From this research, they indicated that Polar Metropolis is more efficient than the Gibbs sampler. Hence, we interested to apply Polar Metropolis in credit rating problem which is ordinal probit model with posterior expectation of cutpoints as a product of constrained cumulative normal distribution. Does Polar Metropolis still have efficiency?

## **CHAPTER III**

### **METHODOLOGY**

This chapter presents the methodology of the study. The purpose of this study is applied the Polar Metropolis for Bayesian parameter estimation of ordinal probit model and to compare efficiency between Polar Metropolis and the Gibbs sampler method. The experiments were conducted using the credit ratings of Thai companies in 2010 and obtained by simulation using R-program. The researcher was conducted as follows:

1. Simulate by using Polar Metropolis which complied with the scope of study and calculate half-width by employing consistent batch means.
2. Simulate by using the Gibbs sampler which complied with the scope of study and calculate half-width by employing consistent batch means.
3. Estimate cutpoints by using Polar Metropolis and the Gibbs sampler through a 2010 credit rating of Thai company data set, and calculate half-width by employing consistent batch means.
4. Count the number of iterations from cutpoints estimation where half-width is less than the required accuracy.
5. Repeat until 10 times per case and compute average number of iterations of them.
6. Analyze and conclude the results.

In the process, the details of the operation are as follows

### 3.1 Data of study.

#### 3.1.1 Credit ratings of companies in Thailand in 2010.

A 2010 credit ratings of Thai companies composed of 30 companies, 3 ratings, 2 cutpoints. This data has two independent variables: the ratio between retained earnings and total assets (RETA) and the ratio of earnings before interest and taxes and total assets (EBITA) by a procedure to estimate the parameters of the cutpoints are as follows.

1. let the precision equal to 0.05 and each batch equal to 500 iterations.
2. Estimate cutpoints by using two methods; the Polar Metropolis and the Gibbs sampler methods
3. Compute half-width by employing consistent batch means.
4. Increase one ( $i = i+1$ ) or more batches until the half-width is less than the required accuracy.
5. Repeat step 1 to 4 until 10 times and compute average number of iterations of them.

#### 3.1.2 Simulated data using Polar Metropolis.

This process would generate parameters, the expected value of the cutpoints  $\tilde{\gamma} = (\gamma_1, \dots, \gamma_d)$ , where  $d$  is the last cutpoint of data set, that determine exactly number of cutpoint. This data from the simulation was used, which is based on ordinal probit model and given the number of independent variables, the correlation between two independent variables, the number of cutpoints and the regression coefficients, according to the scope of study. The procedures for cutpoints estimation are as follow:



1. Let the precision equal to 0.1 and each batch equal to 500 iterations.
2. Estimate cutpoints by using Polar Metropolis
3. Compute half-width by employing consistent batch means.
4. Increase one ( $i = i+1$ ) or more batches until the half-width is less than the required accuracy.
5. Repeat step 1 to 4 until 10 times and compute average number of iterations of them.

### **3.1.3 Simulated data using the Gibbs sampler.**

This process would generate parameters, the expected value of the cutpoints  $\tilde{\gamma} = (\gamma_1, \dots, \gamma_d)$ , where  $d$  is the last cutpoint of data set, that determine exactly number of cutpoint. This data from the simulation was used, which is based on ordinal probit model and given the number of independent variables, the correlation between two independent variables, the number of cutpoints and the regression coefficients, according to the scope of study. The procedures for cutpoints estimation are as follow:

1. Let the precision equal to 0.1 and each batch equal to 500 iterations.
2. Estimate cutpoints by using The Gibbs sampler.
3. Compute half-width by employing consistent batch means.
4. Increase one ( $i = i+1$ ) or more batches until the half-width is less than the required accuracy.
5. Repeat step 1 to 4 until 10 times and compute average number of iterations of them.

### 3.2 Cutpoints estimation using Polar Metropolis.

The key to the Polar Metropolis is that using polar coordinate transforms the cone support to a rectangular support by combining the Metropolis-Hasting algorithm and operating on polar coordinates. This method is randomly selected either an L-move or an S-move by a coin toss, L-move: to generate cutpoints in the range of radius and S-move: to generate cutpoints in the range of angle. Let  $n_{round}$  is a number of iterations, an initial point  $\tilde{\gamma}_0 \in R$ . Below are PM's algorithm for sampling cutpoints in ordinal probit model problem.

1. Given initial parameters as  $\tilde{\beta}_0, \tilde{\gamma}_0$  and  $\tilde{\gamma}_0^*$
2. Given  $\|\tilde{\gamma}^i\|$  is the length of  $\tilde{\gamma}^i$ , the current direction is given by  $\varpi = \frac{\tilde{\gamma}^i}{\|\tilde{\gamma}^i\|}$ , L be a line segment is defined by  $L = \{li: li = \tilde{\gamma}^i + s\varpi, s \in \mathcal{R}\} \cap R$ , and let s is the surface of the intersection between the hypersphere with radius  $\|\tilde{\gamma}^i\|$  and the polytope  $R$ .
3. While half-width is more than the precision.
4. Randomly select either an L-move or an S-move by a coin toss.

L move: random Sampling  $\tilde{\gamma}$  from the conditional density  $g(\cdot | \tilde{\gamma} \in L)$  and accept it as  $\tilde{\gamma}^{i+1}$  with probability

$\min\left\{1, \left(\frac{\|\tilde{\gamma}^{i+1}\|}{\|\tilde{\gamma}^i\|}\right)^{N-1} \cdot \left(\frac{\Phi(\tilde{\gamma}^{i+1})}{\Phi(\tilde{\gamma}^i)}\right) \cdot \left(\frac{\phi(\tilde{\gamma}^{i+1})}{\phi(\tilde{\gamma}^i)}\right)\right\}$  and generate new  $\tilde{\gamma}^*$  from their usual full conditional distribution which depend on the new  $\tilde{\gamma}^{i+1}$ . Otherwise, let  $\tilde{\gamma}^{i+1} = \tilde{\gamma}^i$ .

S move: random Sampling  $\gamma$  from the uniform distribution on S and accept it as  $\tilde{\gamma}^{i+1}$  with probability  $\min\left\{1, \left(\frac{\Phi(\tilde{\gamma}^{i+1})}{\Phi(\tilde{\gamma}^i)}\right)\right\}$  and generate new  $\tilde{\gamma}^*$  from

their usual full conditional distribution which depend on the new  $\tilde{\gamma}^{i+1}$ .  
Otherwise, let  $\tilde{\gamma}^{i+1} = \tilde{\gamma}^i$ .

5. Generate the coefficient of regression ( $\tilde{\beta}$ ).
6. End for

### 3.3 Cutpoints estimation using the Gibbs sampler.

1. Let initialize parameters  $\beta_0, \gamma_0$  and  $y_0^*$ , Let  $n = 0$ .

2. Generate  $y_{n+1}^*$  from full conditional probability

$$y^* | \beta, \gamma, z_i \sim (x_i^T \beta, 1) I(\gamma_{j-1} < y_i^* \leq \gamma_j),$$

$$\text{where } \pi'(y_{n+1}^* | y_{-1}^*) \text{ and } y_{-1}^* = (\beta_n, \gamma_n).$$

3. Generate  $\gamma_{n+1}$  from full conditional probability

$$(\gamma_j | z, y^*, \beta, \{\gamma_i, i \neq j\}) = [\max(\max\{y_i^* : z_i = j\}, \gamma_{j-1}), \\ \min(\min\{y_i^* : z_i = j + 1\}, \gamma_{j+1})],$$

$$\text{where } \pi'(\gamma_{n+1} | \gamma_{-h}) \text{ and } \gamma_{-h} = (\beta_{n+1}, y_{n+1}^*).$$

4. Generate  $\beta_{n+1}$  from full conditional probability

$$\beta | z, y^* \sim ((X^T X)^{-1} X^T Y^*, (X^T X)^{-1}),$$

$$\text{where } \pi'(\beta_{n+1} | \beta_{-k}) \text{ and } \beta_{-k} = (\gamma_{n+1}, y_{n+1}^*).$$

5. Let  $n = n+1$  and repeat step 2 to 5.

### 3.4 The criterion for the performance and procedures comparison.

The performance comparison criterion between two methods is the number of iterations required until all cutpoints converge, the smaller number of iterations the more efficient of method. Thus, the convergence is determined, while the half-width is less than the required accuracy. After received  $n$  sample points of cutpoints  $\tilde{\gamma} = (\gamma_1, \dots, \gamma_d)$ . The procedure for calculating expected value of cutpoints using the Polar Metropolis and the Gibbs sampler is as follows:

1. Burn-in 10 percent of sample points and take 90 percent for calculating the number of consistent batch mean ( $ncbm$ ) from round down of the square root of the remaining data ( $\lfloor \sqrt{(0.9) * n} \rfloor$ ).
2. Use consistent batch mean to estimate Monte Carlo standard error  $\frac{\hat{\sigma}_f}{\sqrt{ncbm}}$ .
3. Compute half-width from  $t_{0.975, (ncbm-1)} \frac{\hat{\sigma}_f}{\sqrt{ncbm}}$ .

## CHAPTER IV

### RESULTS

The results of Bayesian parameter estimation for the ordinal probit model show that in credit ratings of companies in Thailand in 2010, Polar Metropolis is more efficient than the Gibbs sampler. In addition, in simulated data, Polar Metropolis is still more efficient than the Gibbs sampler for 2, 3 and 4 cutpoints. However, in the case of 5 cutpoints, the Gibbs sampler is more efficient than Polar Metropolis. The results of this study can be summarized as being that Polar Metropolis is better than the Gibbs sampler if the data characteristic is a credit ratings problem and there is low number of cutpoints. This chapter presents comparisons of half-width by Polar Metropolis and the Gibbs sampler where the number of iterations ( $n_{iters}$ ) is the same. It shows only part of the credit ratings problem and the data obtained from simulation is fixed at 50 data points for 2, 3, 4 and 5 cutpoints. Therefore, the comparisons of their performance are chosen only against the first cutpoint.

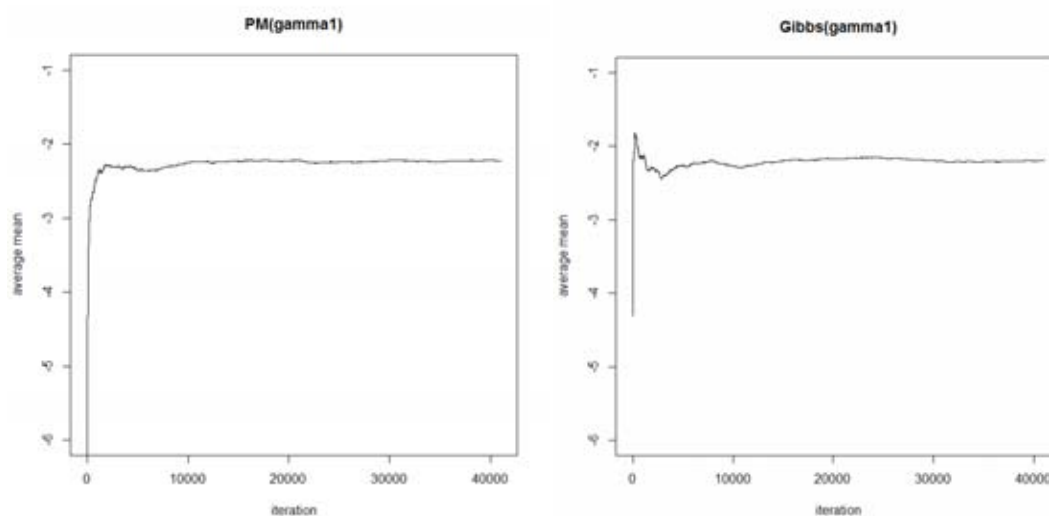
#### 4.1 Results of comparison of credit rating data.

**Table 1.** Mean and half-width of the first cutpoint ( $\gamma_1$ ) and the second cutpoint ( $\gamma_2$ ) using Polar Metropolis and the Gibbs sampler for credit rating data ( $n_{iters} = 41,000$ ).

Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
( $\gamma_1$ )	-2.2266	0.0497	-2.1968	0.0601
( $\gamma_2$ )	-0.6298	0.0428	-0.594	0.0491

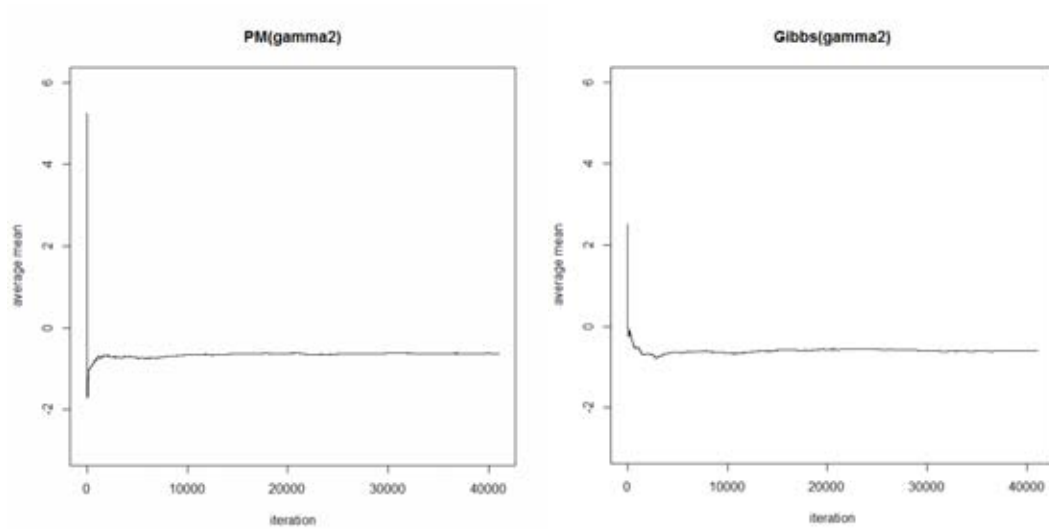
Table 1 presents results of cutpoints estimation. The half-width of the first cutpoint using Polar Metropolis (0.0497) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0601), and half-width of the second cutpoint

by using Polar Metropolis (0.0428) is less than half-width of the first cutpoint by using Gibbs sampler (0.0491). The results show that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of credit rating problems.



**Figure 1.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints regarding the credit ratings of Thai companies in 2010.

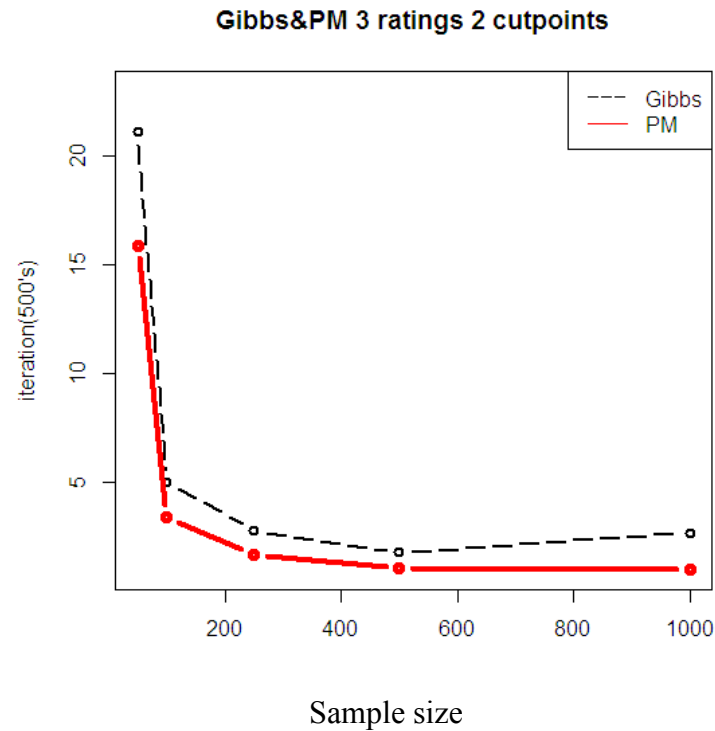
Figure 1 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 41,000 iterations are fixed. The first cutpoints using Polar Metropolis show convergence, but those of the Gibbs sampler do not yet.



**Figure 2.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) with 2 cutpoints for credit ratings of Thai companies in 2010.

Figure 2 presents the average mean plots of Polar Metropolis and the Gibbs sampler where 41,000 iterations are fixed. The second cutpoints using Polar Metropolis and the Gibbs sampler show convergence.

#### 4.2 Results of comparison for 2 cutpoints of simulated data.



**Figure3.** The graph shows the number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 2 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.

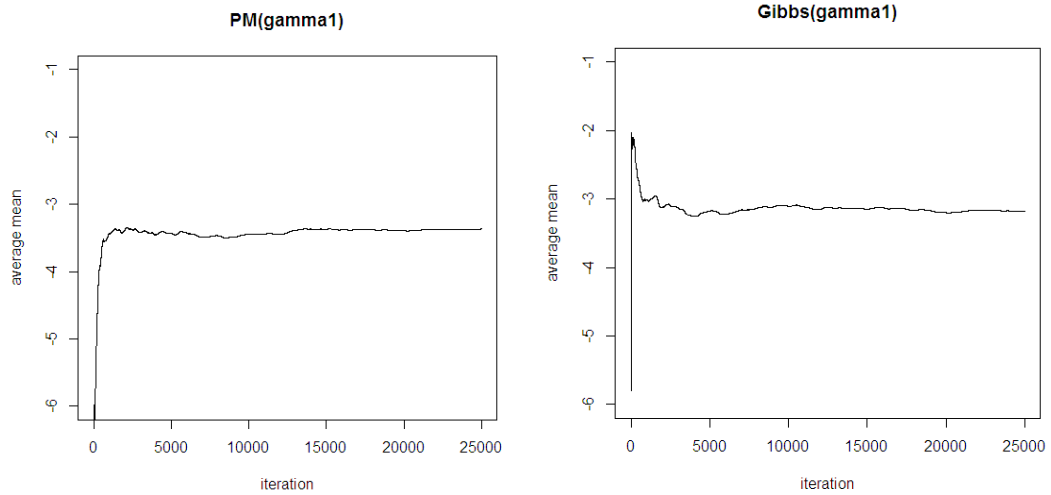
The graph shows that Polar Metropolis is more efficient than the Gibbs sampler regarding cutpoints estimation of 2 cutpoints of simulated data because number of iterations using Polar Metropolis are fewer than number of iterations using the Gibbs sampler at every data point.



**Table 2.** Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 2 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 25,000$ ).

Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
$(\gamma_1)$	-3.3680	0.0755	-3.1819	0.0841

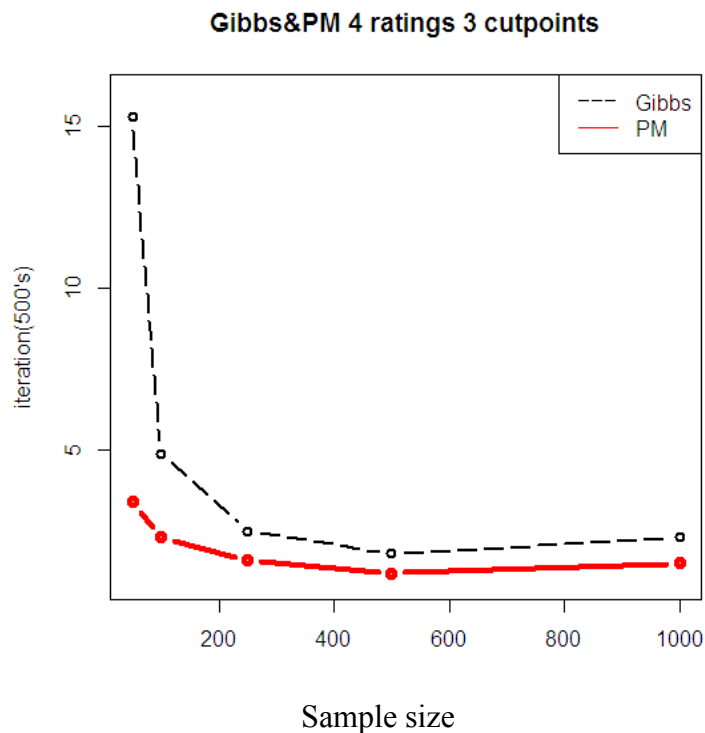
Table 2 presents results from cutpoints estimation. Results show that the half-width of the first cutpoint using Polar Metropolis (0.0755) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0841). The results can be summarized as showing that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of 2 cutpoints from simulated data.



**Figure 4.** The average meanplot of the first cutpointof Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints for simulated data.

Figure 4 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 25,000 iterations are fixed. The first cutpoints using Polar Metropolis show convergence, but those of the Gibbs sampler do not yet.

### 4.3 Results of comparison for 3 cutpoints of simulated data.



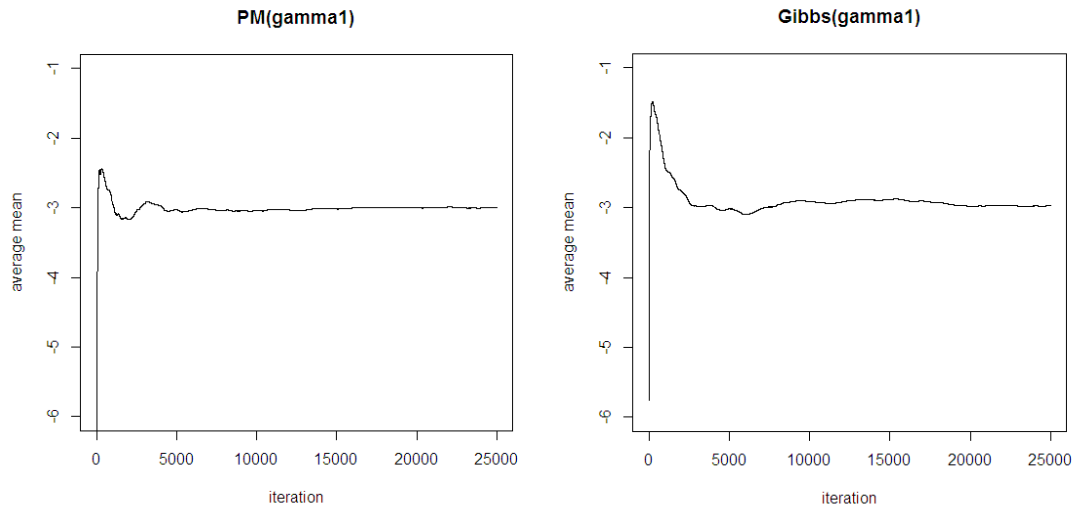
**Figure 5.** The graph shows the number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 3 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.

The graph shows that Polar Metropolis is more efficient than the Gibbs sampler regarding cutpoints estimation of 3 cutpoints of simulated data because number of iterations using Polar Metropolis are fewer than number of iterations using the Gibbs sampler at every data point.

**Table 3.** Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 3 cutpoints and fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 25,000$ ).

Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
$(\gamma_1)$	-2.9927	0.0526	-2.9727	0.0692

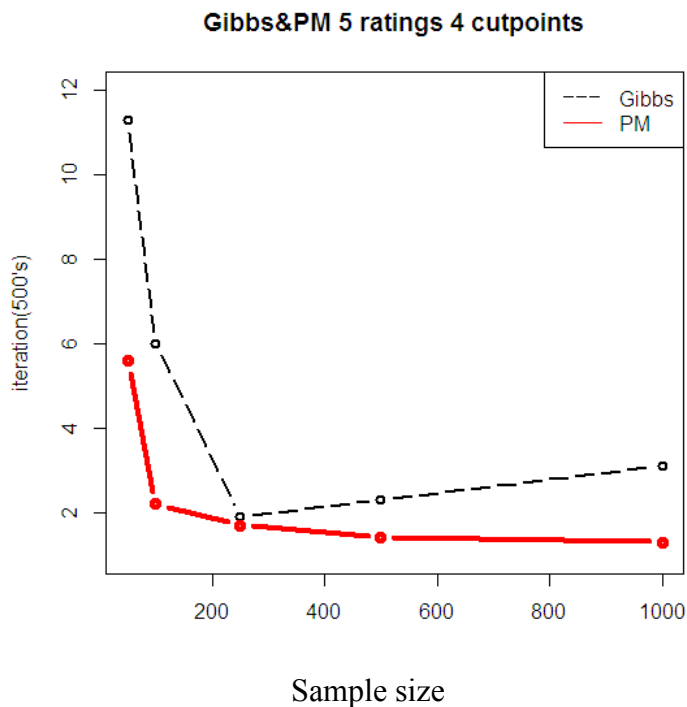
Table 3 presents results from cutpoints estimation. Results show that the half-width of the first cutpoint using Polar Metropolis (0.0526) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0692). The results can be summarized as showing that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of 3 cutpoints from simulated data.



**Figure 6.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 3 cutpoints for simulated data.

Figure 6 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 25,000 iterations are fixed. The first cutpoints using Polar Metropolis show convergence, but those of the Gibbs sampler do not yet.

#### 4.4 Results of comparison for 4 cutpoints of simulated data.



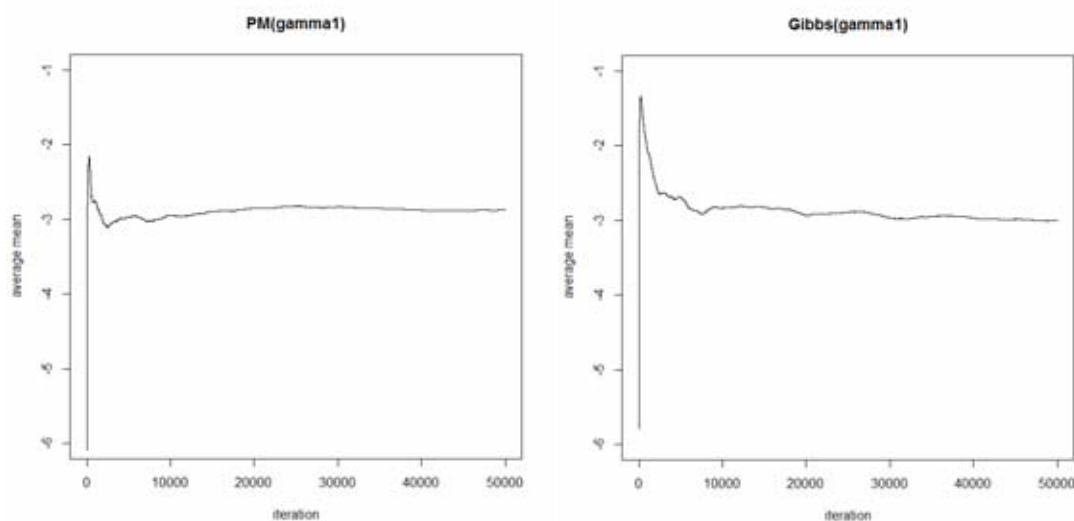
**Figure 7.** The graph shows the number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 4 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.

The graph shows that Polar Metropolis is more efficient than the Gibbs sampler regarding cutpoints estimation of 4 cutpoints of simulated data because number of iterations using Polar Metropolis are fewer than number of iterations using the Gibbs sampler at every data point.

**Table 4.** Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 4 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 25,000$ ).

Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
$(\gamma_1)$	-2.8735	0.0362	-3.0008	0.0597

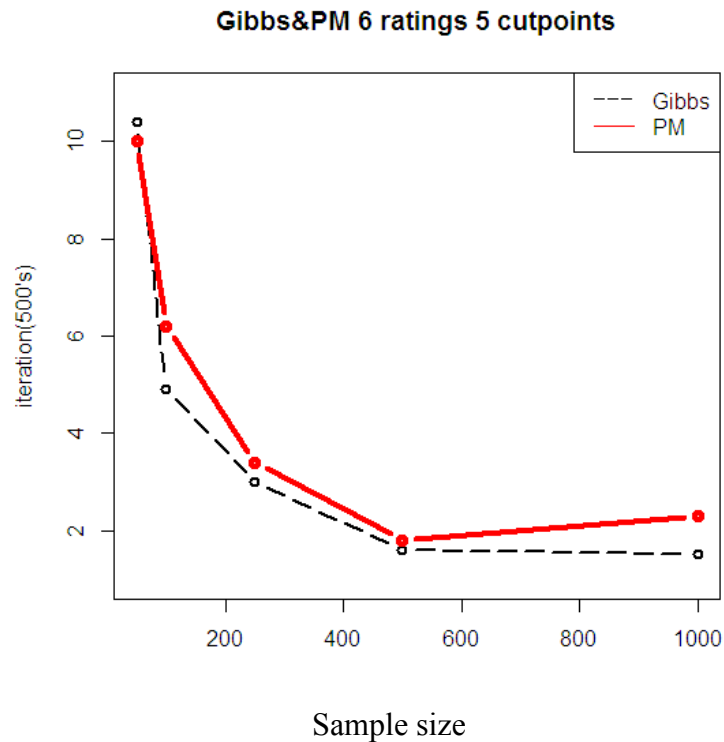
Table 4 presents results from cutpoints estimation. Results show that the half-width of the first cutpoint using Polar Metropolis (0.0362) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0597). The results can be summarized as showing that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of 4 cutpoints from simulated data.



**Figure 8.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 3 cutpoints for simulated data.

Figure 8 presents the average mean mean plot of Polar Metropolis and the Gibbs sampler where 50,000 iterations are fixed. The first cutpoints using Polar Metropolis show convergence, but those of the Gibbs sampler do not yet.

#### 4.5 Results of comparison for 5 cutpoints of simulated data.



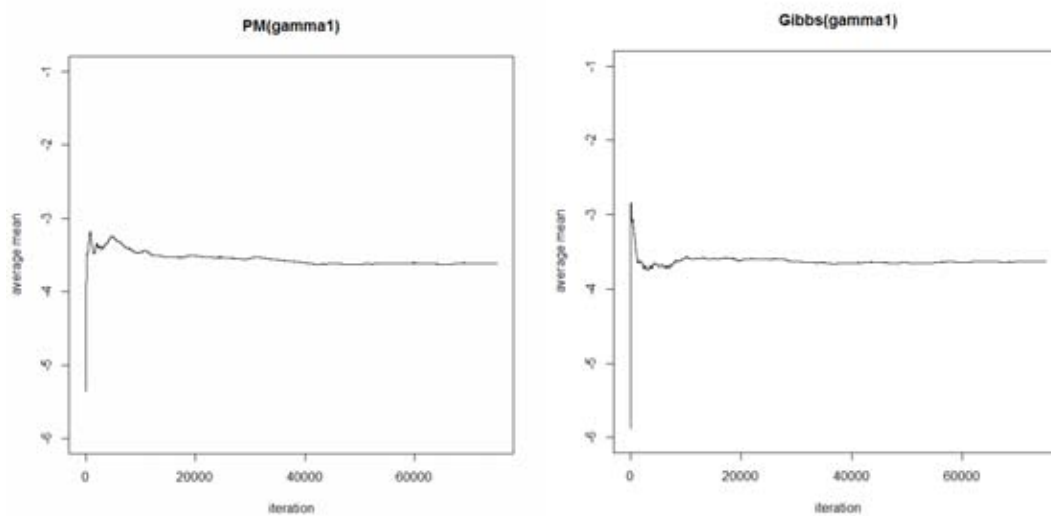
**Figure 9.** The graph shows the number of iterations until all cutpoints converge versus number of observations by comparison between Polar Metropolis (line plot) and the Gibbs sampler (dot plot) for 4 cutpoints of simulated data where the precision is equal to 0.1 and number of observation are 50, 100, 250, 500 and 1,000 data points.

The graph shows that Polar Metropolis is more efficient than the Gibbs sampler regarding cutpoints estimation of 4 cutpoints of simulated data because number of iterations using Polar Metropolis are fewer than number of iterations using the Gibbs sampler at every data point.

**Table 5.** Mean and half-width of the first cutpoints ( $\gamma_1$ ) where given 5 cutpoints and data is fixed 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 50,000$ ).

Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
$(\gamma_1)$	-3.5349	0.0561	-3.6052	0.0637

Table 5 presents results from cutpoints estimation. Results show that the half-width of the first cutpoint using Polar Metropolis (0.0561) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0637). The results can be summarized as showing that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of 5 cutpoints from simulated data.



**Figure 10.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 5 cutpoints for simulated data.

Figure 10 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 75,000 iterations are fixed. The first cutpoints using Polar Metropolis and Gibbs sampler show convergence.

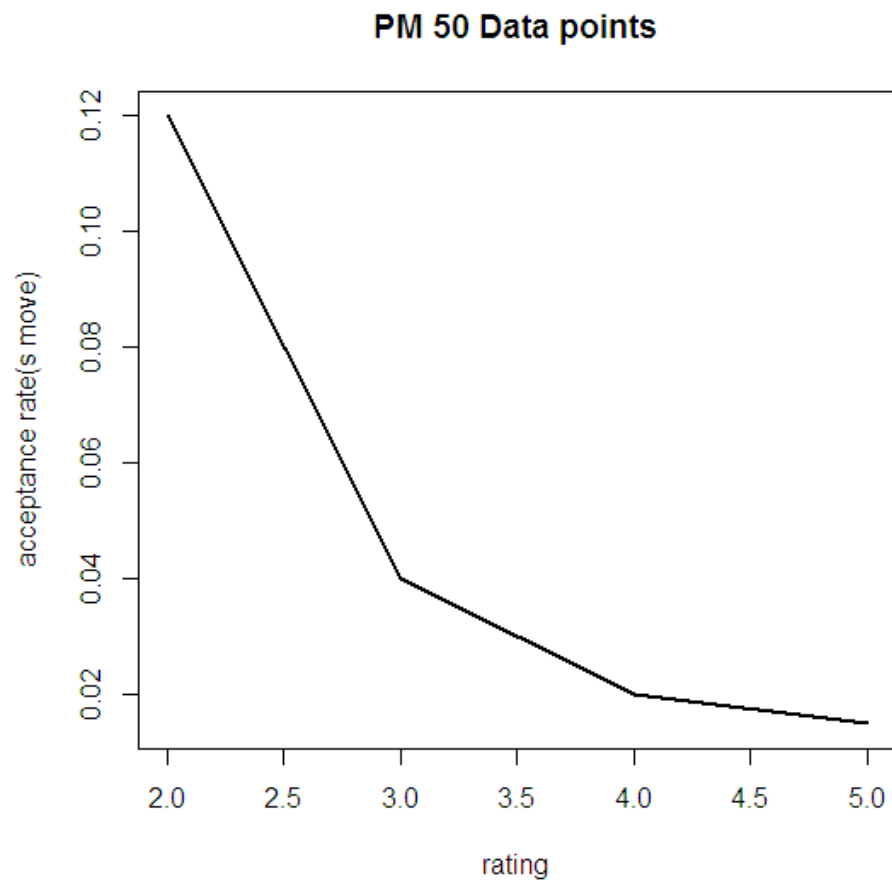
## CHAPTER V

### CONCLUSION AND DISCUSSION

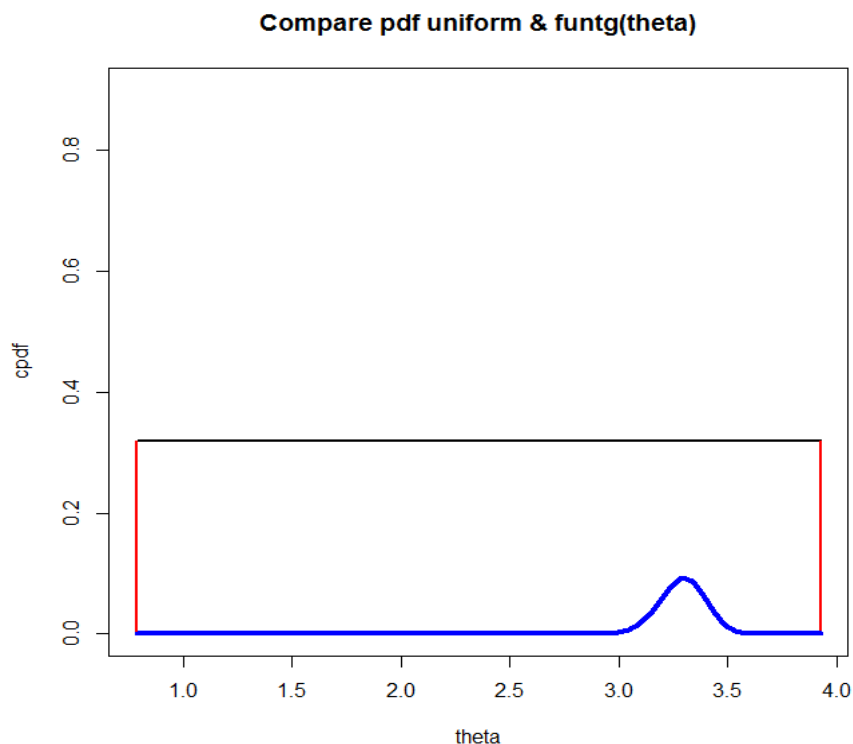
#### 5.1 Conclusion

The problem of estimating the cutpoints of an ordinal probit model in a Bayesian setting is an integration problem of constrained cumulative normal distribution. Although the most popular MCMC method is the Gibbs sampler method, applied to this problem, it has a slow convergence rate due to the conical geometry of the support of the distribution. Therefore, we introduced a new MCMC method for Bayesian parameter estimation of an ordinal probit model which involves using Polar Metropolis. In simulated data sets, it is clear that Polar Metropolis is more efficient than the Gibbs sampler in low dimension problems with 2, 3, and 4 cutpoints. But, from 5 cutpoints, the low acceptance rate of the Metropolis-Hastings algorithm in S-move causes the Polar Metropolis to slow down, which makes the Gibbs sampler more efficient than Polar Metropolis. Results show that the acceptance rate of S-move would drop where there are too many cutpoints (see Fig 11) because unsuitable distribution is used for the proposal density. We provided an example with 2 cutpoints, using uniform distribution at interval of  $[\frac{\pi}{4}, \frac{5\pi}{4}]$  for proposal density to generate cutpoints from their posterior distribution as a product of constrained cumulative normal distribution (see Fig 12).





**Figure 11.** The graph plots acceptance rate versus number of cutpoints at 50 data points.



**Figure12.** The graph shows the proposal density is uniform distribution and the full conditional distribution of cutpoints in the case of 2 cutpoints.

## 5.2 Discussion and further research

As the results of study show a comparison of cutpoints estimation for an ordinal probit model for credit rating using Polar Metropolis and the Gibbs sampler, in low dimensions, 2, 3, and 4 cutpoints, Polar Metropolis is more efficient than the Gibbs sampler. Thus, the results correspond to the research of Chiarawongse et al. (2011) in that Polar Metropolis is more efficient than the Gibbs sampler. Unlike the 2, 3 and 4 cutpoints, in the case of 5 cutpoints, the Gibbs sampler is more efficient than Polar Metropolis because of the problem of the low acceptance rate of cutpoints generated in S-move. Therefore, further research is needed to fine tune Polar Metropolis in its S-move. If successful, we expect the performance of Polar Metropolis to improve.

We perform the experiments with different initial parameters settings. We observed that the initial parameters have some impacts on the convergence rates. Performing the experiments on different sets of simulated data, we observe variation in the convergence rates, which indicates the sensitivity of the algorithm to the data set. However, the correlation between independent variables does not show any direct impact on the experiment results.

The data in this research is simulated only for application to the credit rating problem. Further research, should try other ranking problems such as medical data.

In this research, the confidence interval is the criterion for performance comparison. Thus, in the case of higher dimensions, the MCMC algorithm may stop prematurely. There, further research to confirm the results, may try stopping other criterion.

## REFERENCES

- Albert, J. Bayesian Computation with R. New York: Springer, 2009.
- Albert, J. H. and Chib, S. Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 88 (1993): 669-679.
- Anant Chiarawongse, Seksan Kiatsupaibul, Sunti Tirapat and Benjamin Van Roy. Portfolio selection with qualitative input. Journal of Banking & Finance 36, 2 (2012): 489-496.
- Cowles, M. K.. Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. Journal of Statistics and Computing 6 (1996): 101-111.
- James M. Flegal, Murali Haran and Galin L. Jones..Markov Chain Monte Carlo:Can We Trust the Third Significant Figure?. Statistical Science 23, 2 (2008): 250-260.
- Jeliazkov, I., Graves, J. and Kutzbach, M..Fitting and comparison of models for multivariate ordinal outcomes.Advances in Econometrics: Bayesian Econometrics 23 (2008): 115-156.
- McCullagh, P. Regression models for ordinal data. Journal of the Royal Statistical Society: Series B. 42, 2 (1980): 109-142.

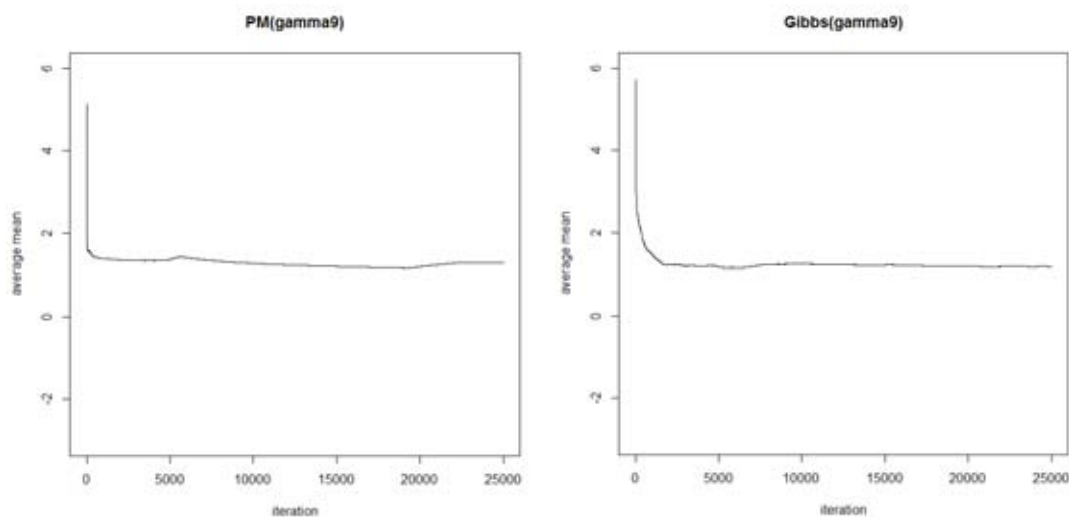
# **APPENDIX**

### A1. Results of comparison of 9 cutpoints of simulated data.

**Table A1.** Mean and half-width of the ninth cutpoints ( $\gamma_9$ ) where given 5 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 25,000$ ).

Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
$(\gamma_9)$	1.2891	0.0603	1.1811	0.0334

Table A1 presents results from cutpoints estimation. Results show that the half-width of the ninth cutpoint using Polar Metropolis (0.0603) is more than the half-width of the ninth cutpoint using the Gibbs sampler (0.0334). The results can be summarized as showing that Polar Metropolis is less efficient than the Gibbs sampler for cutpoints estimation of 5 cutpoints from simulated data.



**Figure A1.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 9 cutpoints for simulated data.

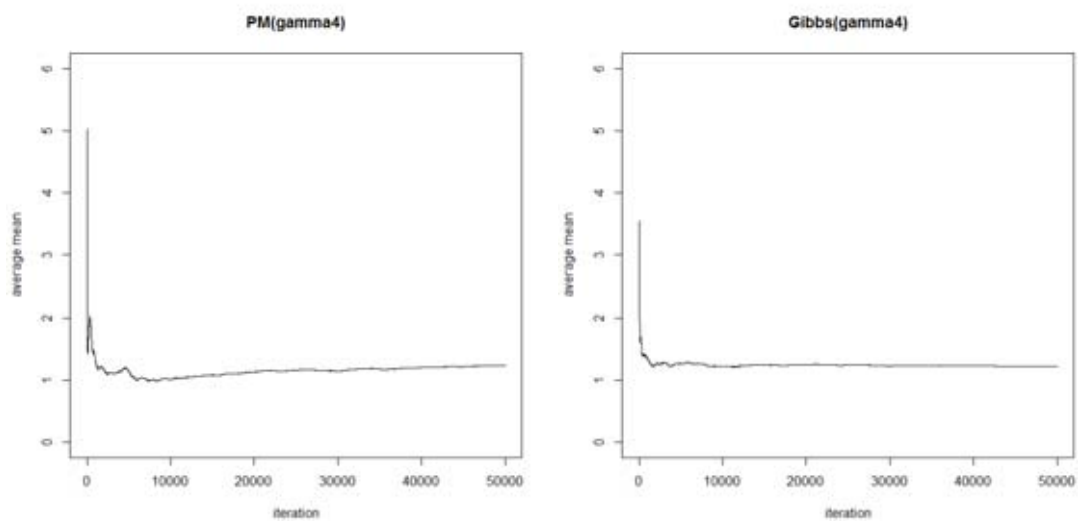
Figure A1 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 25,000 iterations are fixed. The ninth cutpoints using the Gibbs sampler show convergence, but those of Polar Metropolis do not yet.

## A2. Results of comparison for credit rating data (4 cutpoints).

**Table A2.** Mean and half-width of the fourth cutpoints ( $\gamma_4$ ) where given 4 cutpoints and data is fixed at 50 points from simulated data using Polar Metropolis and the Gibbs sampler ( $n_{iters} = 50,000$ ).

Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
$(\gamma_4)$	1.2378	0.0527	1.2252	0.0303

Table A2 presents results from cutpoints estimation. Results show that the half-width of the fourth cutpoint using Polar Metropolis (0.0527) is more than half-width of the fourth cutpoint using the Gibbs sampler (0.0303) The results can summarized as showing that Polar Metropolis is efficient than the Gibbs sampler for cutpoints estimation of 4 cutpoints from credit rating data.



**FigureA2.** The average mean plot of the fourth cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 4cutpoints for credit rating data.

Figure A2 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 50,000 iterations are fixed. The fourth cutpoint using the Gibbs sampler show convergence, but those of Polar Metropolis do not yet.

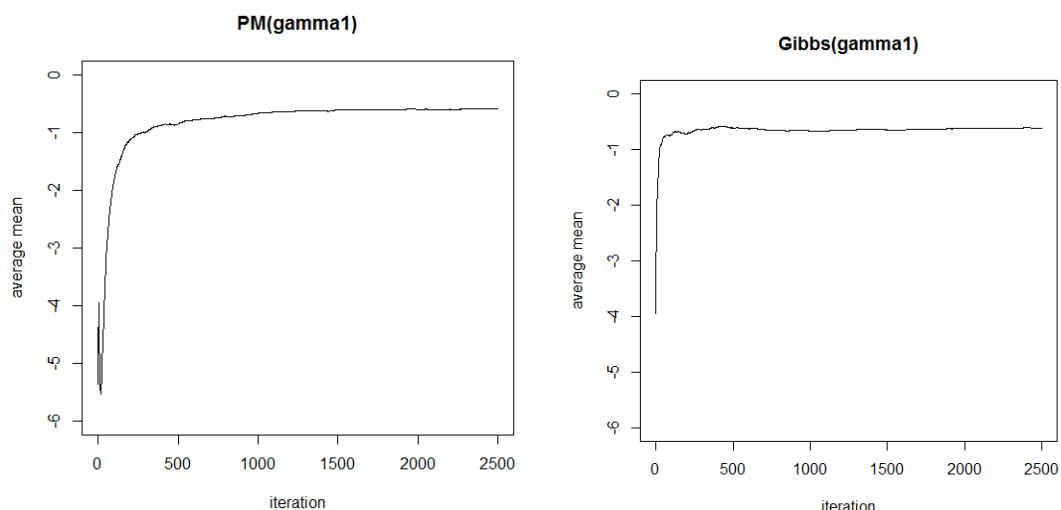


### A3. Results of credit rating for 2 cutpoints where given correlation = -0.37.

**Table A3.** Mean and half-width of the first cutpoint ( $\gamma_1$ ) and the second cutpoint ( $\gamma_2$ ) using Polar Metropolis and the Gibbs sampler for credit rating data ( $n_{iters} = 2,000$ ).

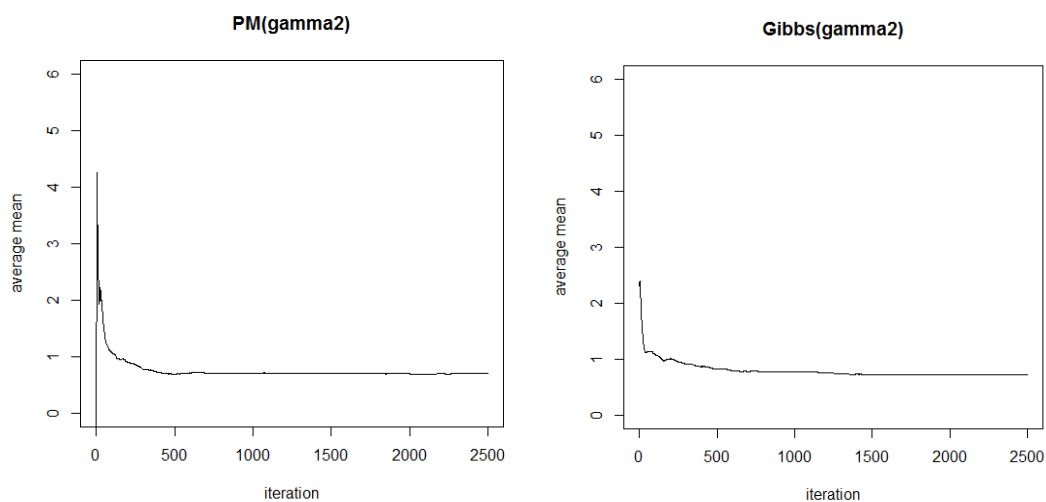
Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
$(\gamma_1)$	-0.5774	0.0477	-0.6147	0.0531
$(\gamma_2)$	0.6973	0.0435	0.8030	0.0467

Table A3. presents results of cutpoints estimation. The half-width of the first cutpoint using Polar Metropolis (0.0477) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0531), and half-width of the second cutpoint by using Polar Metropolis (0.0435) is less than half-width of the first cutpoint by using Gibbs sampler (0.0467). The results show that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of credit rating problems.



**Figure A3.1.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints regarding the credit ratings of Thai companies in 2010.

Figure A3.1. presents the average mean plot of Polar Metropolis and the Gibbs sampler where 2,000 iterations are fixed. The first cutpoints using Polar Metropolis and the Gibbs sampler do not converge.



**Figure A3.2.** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) with 2 cutpoints for credit ratings of Thai companies in 2010.

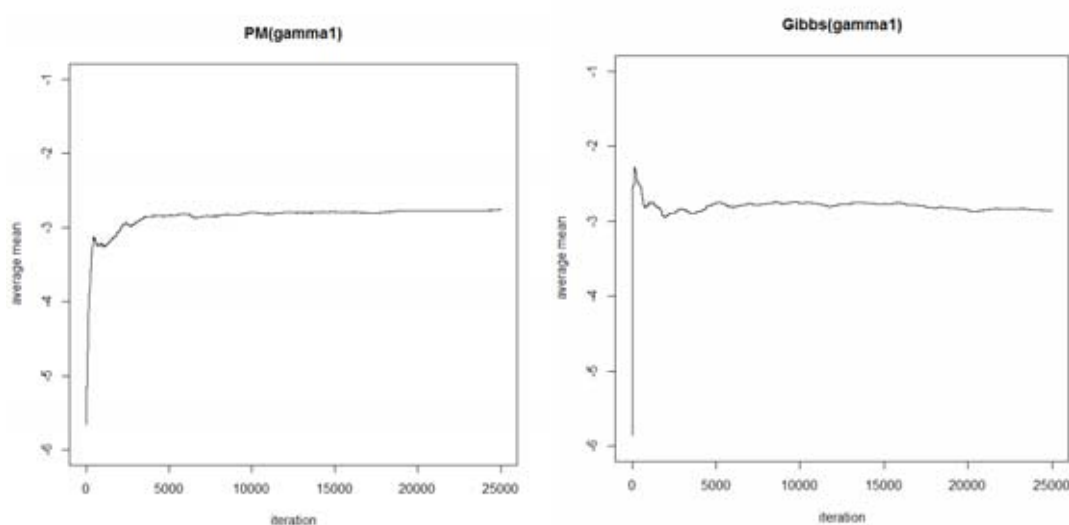
Figure A3.2. presents the average mean plots of Polar Metropolis and the Gibbs sampler where 2,000 iterations are fixed. The second cutpoints using Polar Metropolis show convergence, but those the Gibbs sampler do not yet.

#### A4. Results of simulated data for 2 cutpoints where given correlation = 0.

**Table A4.** Mean and half-width of the first cutpoint ( $\gamma_1$ ) and the second cutpoint ( $\gamma_2$ ) using Polar Metropolis and the Gibbs sampler for credit rating data ( $n_{iters} = 25,000$ ).

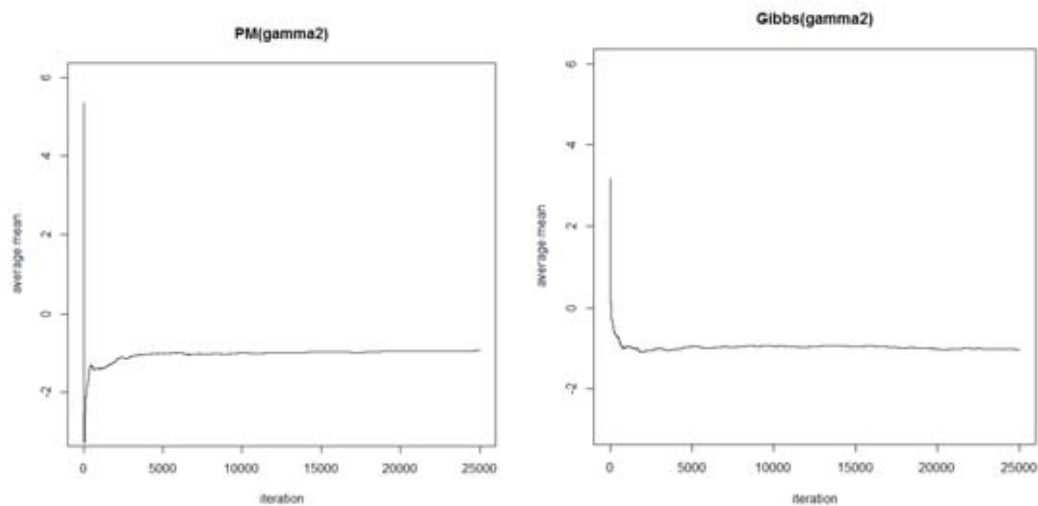
Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
( $\gamma_1$ )	-2.757	0.0565	-2.8624	0.0832
( $\gamma_2$ )	-0.9341	0.0524	-1.0209	0.0673

Table A4. presents results of cutpoints estimation. The half-width of the first cutpoint using Polar Metropolis (0.0565) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0832), and half-width of the second cutpoint by using Polar Metropolis (0.0524) is less than half-width of the first cutpoint by using Gibbs sampler (0.0673). The results show that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of credit rating problems.



**FigureA4.1** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints for simulated data.

Figure A4.1 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 25,000 iterations are fixed. The first cutpoints using Polar Metropolis show convergence, but those of the Gibbs sampler do not yet.



**FigureA4.2** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints for simulated data.

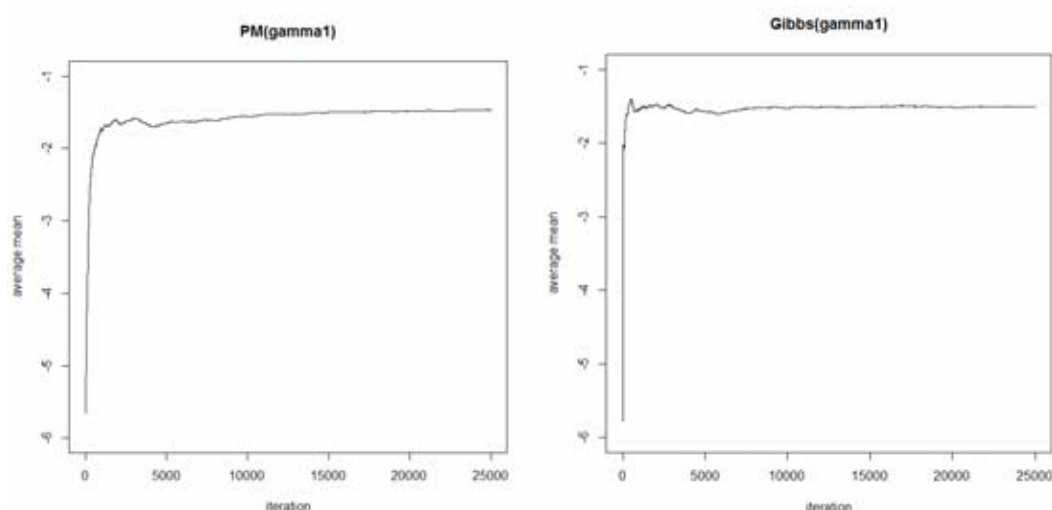
Figure A4.2 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 25,000 iterations are fixed. The first cutpoints using Polar Metropolis and the Gibbs sampler show convergence.

### A5. Results of simulated data for 2 cutpoints where given correlation = 0.9.

**Table A5.** Mean and half-width of the first cutpoint ( $\gamma_1$ ) and the second cutpoint ( $\gamma_2$ ) using Polar Metropolis and the Gibbs sampler for credit rating data ( $n_{iters} = 25,000$ ).

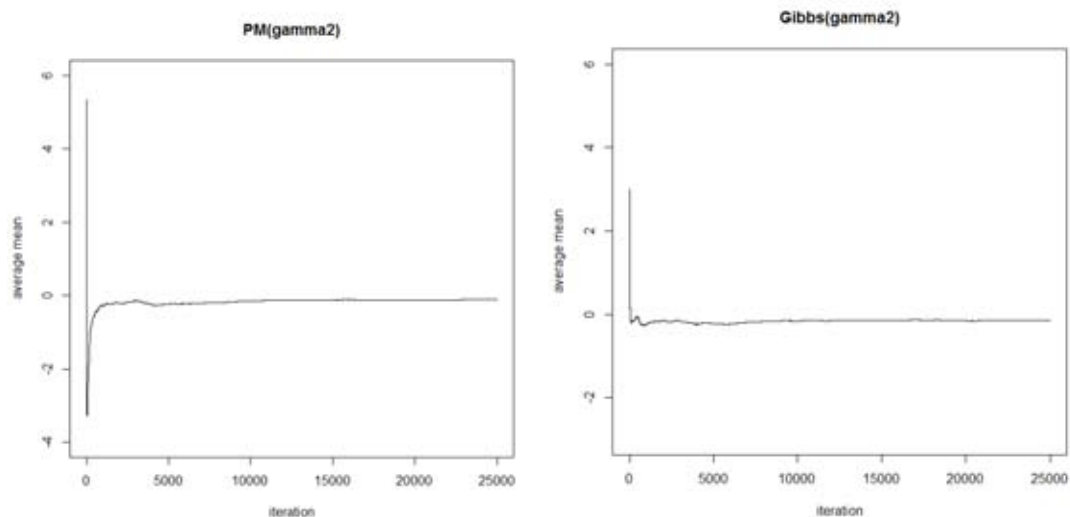
Cut points	Polar Metropolis		Gibbs Sampler	
	Mean	Std.Err	Mean	Std.Err
( $\gamma_1$ )	-1.4693	0.0497	-1.5035	0.0532
( $\gamma_2$ )	-0.1050	0.0454	-0.1449	0.0486

Table A5. presents results of cutpoints estimation. The half-width of the first cutpoint using Polar Metropolis (0.0497) is less than the half-width of the first cutpoint using the Gibbs sampler (0.0532), and half-width of the second cutpoint by using Polar Metropolis (0.0454) is less than half-width of the first cutpoint by using Gibbs sampler (0.0486). The results show that Polar Metropolis is more efficient than the Gibbs sampler for cutpoints estimation of credit rating problems.



**FigureA5.1** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints for simulated data.

Figure A5.1 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 25,000 iterations are fixed. The first cutpoints using Polar Metropolis and the Gibbs sampler do not converge.



**FigureA5.2** The average mean plot of the first cutpoint of Polar Metropolis (right) and the Gibbs sampler (left) in the case of 2 cutpoints for simulated data.

Figure A5.2 presents the average mean plot of Polar Metropolis and the Gibbs sampler where 25,000 iterations are fixed. The first cutpoints using Polar Metropolis and the Gibbs sampler show convergence.

## A6. About the Markov Chains

### A6.1. Stationary distribution

Let  $\pi$  is probability distribution over  $(\zeta, S)$ .  $\pi$  is said to be stationary distribution of probability transition matrix  $P$ , if  $\pi = \pi P$ .

### A6.2. Irreducible Markov chains

Let  $P^n$  is n-step transition probability and  $\Psi$  is finite measure space over  $(\zeta, S)$ . Markov Chain is said to be irreducible, if for every  $A \in S$  with  $\Psi(A) > 0$ , there exists n such that  $P^n(x, A) > 0$  for all  $x \in \zeta$ .

### A6.3. Aperiodic of Markov chains

$P$  has a period if there exist integer  $d \geq 2$  and sequence of subset  $d$  of  $\zeta \{E_0, E_1, \dots, E_{d-1}\}$  such as  $i = 0, 1, \dots, d-1$  for all  $\theta \in \zeta$ .

$$P^n(x, E_j) = 1 \text{ where } j = (i + 1) \bmod d$$

Otherwise  $P$  is said to be aperiodic.

### A6.4. Reversibility

Transition probability  $P$  is reversibility with respect to  $\pi$  where  $\pi$  be a probability distribution over measure space  $(\zeta, S)$ . If for every  $A, B \in S$

$$\int_A P(\theta, B) \pi(d\theta) = \int_B P(\theta, A) \pi(d\theta).$$

where countable space  $\zeta$  has all possible value  $K$ , transition probability  $P$  is given by the matrix  $\tilde{P} = [p_{ij}]$  with size  $K \times K$ . Hence, a stationary Markov Chain is said to be reversible if  $\pi_i p_{ij} = \pi_j p_{ji}$ .

### A6.5. Markov Chain Convergence in total variation

Theorem: If  $P$  is irreducible over  $\Psi$ , where  $\Psi$  is measure space on  $(\zeta, S)$ , and  $\pi$  is stationary distribution. If  $P$  is aperiodic then for almost every  $x$ ,

$$\|P^n(x, \cdot) - \pi\| \Rightarrow 0$$

Where  $\|\cdot\|$  is total variational distance, as

$$\lim_{n \rightarrow \infty} \Pr(\theta_n \in A | \theta_0 = x) = \pi(A), \quad \forall A \in S$$

### A6.6. Ergodicity

Markov Chain is said to be ergodic if finite state Markov Chains that are irreducible and aperiodic have the property that,  $\exists n \geq 0$ ,  $\mathbf{P}^n$  has all entries positive.

If the Markov Chains Monte Carlo (MCMC) algorithms can be constructed, the ergodic theorem can examine most of settings, the average as  $\frac{1}{N} \sum_{n=1}^N f(\theta_n)$  converges to  $E[f(\cdot)]$ .



## A7. Code for R-Program to compute Bayesian parameters.

### A7.1. Bayesian parameters estimation using Polar Metropolis.

```

dataset<-read.csv("MCMC.csv",header=TRUE);
x<-as.matrix(dataset[,2:3]);
y<-as.matrix(dataset[,1]);
library(MASS);
library(mvtnorm);
#Defined variable
simround<-500;
precision<-0.05;
betavec<-as.matrix(c(-1,-6));
levels(y)<-c("1","2","3");
ratingvec<-levels(y);
meanystar<-as.matrix(c(NA,nrow=NA));
gammavec<-seq(from=-6, to=6, length=length(ratingvec)-1);
cgammavec<-c(NA,length=length(ratingvec)-1);
stdc<-0.625;
dim<-length(gammavec);
pivec<-c(NA,length=length(ratingvec));
xtb<-rep(NA, length(dataset[,1]),nrow=length(dataset[,1]),ncol=1);
ystar<-rep(NA, length(dataset[,1]));
ystar<-as.matrix(dataset[,2:3])%*%betavec;
cgammamat<-c();
countl<-0;
countll<-0;
counts<-0;
countss<-0;
betamat<-c();
halfwidth<-1;
batchnum<-0;
gibbsloop<-function(){
while(halfwidth> precision){

```

```

for (i in 1:simround){
#calculate xtb
for(j in 1:length(ratingvec)){
  meanystar<-
as.matrix(dataset[dataset[,1]==ratingvec[j],2:3],ncol=2)%*%betavec;
  xtb[dataset[,1]==ratingvec[j]]<-meanystar;
}

#simulate candidate gamma
if(runif(1)<=0.5){
#L Move
countl<-countl+1;
rgam<-sqrt(sum(gammavec^2));
rcgam<-qnorm(runif(1)*(1-pnorm(0,rgam,stdc))+pnorm(0,rgam,stdc),rgam,stdc);
cgammavec<-rcgam*(gammavec/rgam);

#calculate r (ry/rx)
r<-(rcgam/rgam)^(dim-1);

#calculate proposal density (qy/qx)
qydqx<-(dnorm(rgam,rcgam,stdc)/
(1- pnorm(0,rcgam,stdc)))/(dnorm(rcgam,rgam,stdc)/(1-pnorm(0,rgam,stdc)));

# Calculatepir (piy/pix)
for (j in 1:length(ratingvec)){
  gammaminusxtb<-
c(NA,length=length(xtb[dataset[,1]==ratingvec[j]]));
  if (j==1){
    gammaminusxtb<-pnorm((cgammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1)/pnorm((gammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1);
    pivec[j]<-prod(gammaminusxtb);
  }
}
}

```

```

}else{
  if (j==length(ratingvec)){
    gammaminusxtb<-(1-pnorm((cgammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1))/(1-pnorm((gammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1));
    pivec[j]<-prod(gammaminusxtb);
  }else{
    gammaminusxtb<-(pnorm((cgammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1)-pnorm((cgammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1))/(pnorm((gammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1)-pnorm((gammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1));
    pivec[j]<-prod(gammaminusxtb);
  }
}
}

piydpix<-prod(pivec);

#calculate alpha
alphal<-piydpix*r*qydx;
minr<-min(alphal,1);
if(runif(1)<=minr){
  countll<-countll+1;
  gammavec<-cgammavec;

  #Simulate y*
  for (j in 1:length(ratingvec)){
    if(j == 1){
      probvec<-
runif(length(xtb[dataset[,1]==ratingvec[j]]))*pnorm(gammavec
[j],xtb[dataset[,1]==ratingvec[j]],1);
    }else{

```

```

    if(j == length(ratingvec)){
      probvec<-runif(length(xtb[dataset[,1]==ratingvec[j]]))*
      (1-pnorm(gammavec[j-1],xtb[dataset[,1]==ratingvec[j]],1))+
      pnorm(gammavec[j-1],xtb[dataset[,1]==ratingvec[j]],1);
    }else{
      probvec<-
      runif(length(xtb[dataset[,1]==ratingvec[j]]))*(pnorm(gammave
      c[j],xtb[dataset[,1]==ratingvec[j]],1)-pnorm(gammavec[j-
      1],xtb[dataset[,1]==ratingvec[j]],1)) + pnorm(gammavec[j-
      1],xtb[dataset[,1]==ratingvec[j]],1);
    }
  }
  ystar[dataset[,1]==ratingvec[j]]<-
  qnorm(probvec,xtb[dataset[,1]==ratingvec[j]],1);
}
}

}else{
  counts<-counts+1;
  #S Move
  #Uniform Surface Hypersphere
  svec<-rnorm(length(gammavec),0,1);
  svec<-svec/sqrt(sum(svec^2));
  cgammavec<-sort(svec);
  rgam<-sqrt(sum(gammavec^2));
  cgammavec<-rgam*cgammavec;

  # Calculatepir (piy/pix)
  for (j in 1:length(ratingvec)){
    gammaminusxtb<-
    c(NA,length=length(xtb[dataset[,1]==ratingvec[j]]));
    if (j==1){

```

```

      gammaminusxtb<-pnorm((cgammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1)/pnorm((gammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1);
      pivec[j]<-prod(gammaminusxtb);
    }else{
      if (j==length(ratingvec)){
        gammaminusxtb<-(1-pnorm((cgammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1))/(1-pnorm((gammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1));
        pivec[j]<-prod(gammaminusxtb);
      }else{
        gammaminusxtb<-(pnorm((cgammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1)-pnorm((cgammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1))/(pnorm((gammavec[j]-
xtb[dataset[,1]==ratingvec[j]]),0,1)-pnorm((gammavec[j-1]-
xtb[dataset[,1]==ratingvec[j]]),0,1));
        pivec[j]<-prod(gammaminusxtb);
      }
    }
  }
  piydpix<-prod(pivec);

  #calculate alpha
  alphas<-piydpix
  minr<-min(alphas,1)

  if(runif(1)<=minr){
    countss<-countss+1;
    gammavec<-cgammavec;

    #Simulate y*
    for (j in 1:length(ratingvec)){

```

```

    if(j == 1){
      probvec<-
      runif(length(xtb[dataset[,1]==ratingvec[j]]))*pnorm(gammavec[j],xtb[dataset[
,1]==ratingvec[j]],1);
    }else{
      if(j == length(ratingvec)){
        probvec<-runif(length(xtb[dataset[,1]==ratingvec[j]]))*
        (1-pnorm(gammavec[j-1],xtb[dataset[,1]==ratingvec[j]],1))+
        pnorm(gammavec[j-1],xtb[dataset[,1]==ratingvec[j]],1);
      }else{
        probvec<-runif(length(xtb[dataset[,1]==ratingvec[j]]))*
        (pnorm(gammavec[j],xtb[dataset[,1]==ratingvec[j]],1)-
        pnorm(gammavec[j-1],xtb[dataset[,1]==ratingvec[j]],1)) +
        pnorm(gammavec[j-1],xtb[dataset[,1]==ratingvec[j]],1);
      }
    }
  }
  ystar[dataset[,1]==ratingvec[j]]<-qnorm(probvec,xtb[dataset[,1]==ratingvec[j]],1);
}
}
}
}
cgammat<-rbind(cgammat,matrix(gammavec,nrow=1));

#Simulate beta
betavec<-
t(rmvnorm(1,(ginv(t(x)%*%x)%*%t(x)%*%ystar),as.matrix(ginv(t(x)%*%x))));

}

#Compute halfwidth
batchnum<-batchnum+1;
newhalfwidthvec<-rep(NA,length=(length(ratingvec)-1));
#Number of batches for inference according to Consistent batch mean
ncbm<-floor(sqrt(0.9*batchnum*simround));

```

```

for(i in 1:(length(ratingvec)-1)){
  #Number of inference batch and the batch size is equal
  infvec<-cgammat[-(1:(0.1*batchnum*simround)),i];
  infvec<-infvec[1:(ncbm*ncbm)];
  infmat<-matrix(infvec,ncol=ncbm);
  batchmeanvec<-colMeans(infmat);
  newhalfwidthvec[i]<-qt(0.975,ncbm-1)*
  sd(batchmeanvec)/sqrt(length(batchmeanvec));
}

halfwidth<-max(newhalfwidthvec);
}

result<-c(batchnum,countll/countl,countss/counts);
}

batchvec<-rep(NA,length=10)
aclrate<-rep(NA,length=10)
acsrates<-rep(NA,length=10)
  for(i in 1:10){
    set.seed(100+i);
    outvec<-gibbsloop();
    batchvec[i]<-outvec[1];
    aclrate[i]<-outvec[2];
    acsrates[i]<-outvec[3];
    print(i)
  }

batchmean<-mean(batchvec);
aclmean<-mean(aclrate);
acsmean<-mean(acsrates);

```

### A7.2. Bayesian parameters estimation using the Gibbs sampler.

```

dataset<-read.csv("MCMC.csv",header=TRUE);
x<-as.matrix(dataset[,2:3]);
y<-as.matrix(dataset[,1]);
library(MASS);
library(mvtnorm);
#Defined variable
simround<-500;
precision<-0.05;
betavec<-as.matrix(c(-1,-6));
levels(y)<-c("1","2","3");
ratingvec<-levels(y);
meanystar<-as.matrix(c(NA,nrow=NA));
gammavec<-seq(from=-6, to=6, length=length(ratingvec)-1);
gammavecnew<-c(NA,length(ratingvec)-1);
betamat<-c();
gammamat<-c();
ystar<-rep(NA, length(dataset[,1]),nrow=30,ncol=1);
halfwidth<-1;
batchnum<-0;
gibbsloop<-function(){
while(halfwidth> precision){
for (i in 1:simround){
  for (j in 1:length(ratingvec)){
    #Simulate y*
meanystar<-as.matrix(dataset[dataset[,1]==ratingvec[j],2:3],ncol=2)%*%betavec;
    if(j == 1){
      probvec<-runif(length(meanystar))*pnorm(gammavec[j],meanystar,1);
    }else{
      if(j == length(ratingvec)){
        probvec<-runif(length(meanystar))*(1-pnorm(gammavec[j]-
1],meanystar,1))+ pnorm(gammavec[j-1],meanystar,1);

```



```

    }else{
      probvec<-
      runif(length(meanystar))*(pnorm(gammavec[j],meanystar,1)-
      pnorm(gammavec[j-1],meanystar,1)) + pnorm(gammavec[j-
      1],meanystar,1);
    }
  }
  ystar[dataset[,1]==ratingvec[j]]<-qnorm(probvec, meanystar,1);
}

#Simulate gamma
for ( j in 1:(length(ratingvec)-1) ){
  lb<-ifelse(j==1,max(ystar[dataset[,1]==ratingvec[j]]),
  max(max(ystar[dataset[,1]==ratingvec[j]],gammavec[j-1])));
  ub<-ifelse(j==(length(ratingvec)-1),min(ystar[dataset[,1]==ratingvec[j+1]]),
  min(min(ystar[dataset[,1]==ratingvec[j+1]],gammavec[j+1])));
  gammavec[j]<-runif(1)*(ub-lb)+lb;
}
gammamat<-rbind(gammamat,matrix(gammavec,nrow=1));

#simulate betavector
betavec<-
t(rmvnorm(1,(ginv(t(x)%*%x)%*%t(x)%*%ystar),as.matrix(ginv(t(x)%*%x))));
}

#Compute halfwidth
batchnum<-batchnum+1;
newhalfwidthvec<-rep(NA,length=(length(ratingvec)-1));
#Number of batches for inference according to Consistent batch mean
ncbm<-floor(sqrt(0.9*batchnum*simround));
for(i in 1:(length(ratingvec)-1)){
  #Number of inference batch and the batch size is equal

```

```
infvec<-gammamat[-(1:(0.1*batchnum*simround)),i];
infvec<-infvec[1:(ncbm*ncbm)];
infmat<-matrix(infvec,ncol=ncbm);
batchmeanvec<-colMeans(infmat);
newhalfwidthvec[i]<-qt(0.975,ncbm-
1)*sd(batchmeanvec)/sqrt(length(batchmeanvec));
}
halfwidth<-max(newhalfwidthvec);
}
batchnum;
}
batchvec<-rep(NA,length=10)
for(i in 1:10){
    set.seed(100+i);
    batchvec[i]<-gibbsloop();
    print(i)
}
batchmean<-mean(batchvec);
```

## **BIOGRAPHY**

Chaiyanun Tharasuke was born in Bangkok, Thailand on 7 February 1984. He completed a Bachelor of Science, majoring in Mathematics in 2006 from Department of Mathematics and Applied Mathematics, Faculty of Science, Thammasat University. After three years of work he enrolled in Master of Science in Statistics program at Chulalongkorn University in 2010.