การเปรียบเทียบการลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โลสำหรับการอนุมานเชิงเบส์เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง

นายพงศ์ศักดิ์ สถิตรุ่งพรชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553

A COMPARISON OF MARKOV CHAIN MONTE CARLO CONVERGENCE FOR BAYESIAN INFERENCE WITH TRUNCATED NORMAL POSTERIOR DISTRIBUTION

Mr.Pongsak Sathitrungpornchai

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 2010

หัวข้อวิทยานิพนธ์	การเปรียบเทียบการลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โลสำหรับการอนุมานเชิงเบส์ เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง
โดย	นายพงศ์ศักดิ์ สถิตรุ่งพรชัย
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.เสกสรร เกียรติสุไพบูลย์

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.... คณบดีคณะพาณิชยศาสตร์และการบัญชี

(รองศาสตราจารย์ ดร.อรรณพ ตันละมัย)

คณะกรรมการสอบวิทยานิพนธ์

.... ประธานกรรมการ

(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)

.... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์ ดร.เสกสรร เกียรติสุไพบูลย์)

.... กรรมการ

(รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา)

.... กรรมการภายนอกมหาวิทยาลัย

(อาจารย์ ดร.อัศรินทร์ ไพบูลย์พานิช)

พงศ์ศักดิ์ สถิตรุ่งพรชัย : การเปรียบเทียบการลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โล สำหรับการอนุมานเชิงเบส์ เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง. (A COMPARISON OF MARKOV CHAIN MONTE CARLO CONVERGENCE FOR BAYESIAN INFERENCE WITH TRUNCATED NORMAL POSTERIOR DISTRIBUTION) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร.เสกสรร เกียรติสุไพบูลย์, 53 หน้า.

การอนุมานเชิงเบส์เมื่อการแจกแจงก่อนเป็นแบบปกติ และ พารามิเตอร์ที่เป็นค่าคาดหวังมีการเรียงอันดับอย่างสมบูรณ์ จะมีการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง การอนุมานสามารถทำได้โดยใช้วิธีในกลุ่มลูกโซ่มาร์คอฟมอนติคาร์โล การวิจัยนี้จึงทำการเปรียบเทียบประสิทธิภาพของเทคนิคในกลุ่มลูกโซ่มาร์คอฟมอนติคาร์โล 2 วิธี คือ การสุ่มตัวอย่างแบบฮิตแอนด์รัน และ การสุ่มตัวอย่างแบบกิบส์ โดยนำเทคนิคทั้งสองมาแก้ปัญหาการอนุมานเชิงเบส์ที่มีการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง ประสิทธิภาพนิยามโดยค่าครึ่งช่วงความเชื่อมั่นที่คำนวณจากวิธีค่าเฉลี่ยกลุ่ม และค่า MPSRF ของบรูกซ์-เกลแมน

จากการศึกษาพบว่า ในกรณีที่จำนวนมิติของพารามิเตอร์ที่เป็นค่าคาดหวังมีค่าต่ำ การสุ่มตัวอย่างแบบฮิตแอนด์รัน จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ แต่ในกรณีที่จำนวนมิติของพารามิเตอร์ที่เป็นค่าคาดหวังมีค่าสูง การสุ่มตัวอย่างแบบกิบส์ จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

ภาควิชา ....สถิติ..........	ลายมือชื่อนิสิต.....พงศ์ศักดิ์ สถิตรุ่งพรชัย.....
สาขาวิชา ....สถิติ..........	ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก..........
ปีการศึกษา ....2553..........	

5181862426 : MAJOR STATISTICS
KEYWORDS : CONVERGENCE DIAGNOSTIC / GIBBS SAMPLER / HIT-AND-RUN SAMPLER / MARKOV CHAIN MONTE CARLO

PONGSAK SATHITRUNGPORNCHAI : A COMPARISON OF MARKOV CHAIN MONTE CARLO CONVERGENCE FOR BAYESIAN INFERENCE WITH TRUNCATED NORMAL POSTERIOR DISTRIBUTION.THESIS ADVISOR : ASST. PROF. SEKSAN KIATSUPAIBUL, Ph.D., 53 pp.

In a Bayesian inference, when the prior distribution is a normal distribution and the expectation parameter is completely ordered, the posterior distribution becomes a truncated normal distribution in high dimension. The inference can be performed by employing methods in the class of Markov chain Monte Carlo. The objective of this study is to compare the performances between the two Markov chain Monte Carlo techniques, the hit-and-run sampler and Gibbs sampler, by applying the two techniques to the Bayesian inference problem with a truncated normal posterior distribution. The performance are defined as the half width of the confidence interval formed by the batch means approach and the Brooks–Gelman MPSRF.

From our study, we find that the hit-and-run sampler and the Gibbs sampler perform differently depending on the dimension of the expectation parameter. More specifically, when the dimension of the expectation parameter is low, the hit-and-run sampler works better. On the other hand, the Gibbs sampler performs more effectively when the dimension of the expectation is high.

Department : Statistics	Student's Signature Pongsak Sathitrungpornchai
Field of Study : Statistics	Advisor's Signature [signature]
Academic Year : 2010	

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความกรุณาและเอาใจใส่อย่างดียิ่งของ ผู้ช่วยศาสตราจารย์ ดร. เสกสรร เกียรติสุไพบูลย์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งได้ให้คำปรึกษา คำแนะนำ และช่วยเหลือแก้ไขข้อบกพร่อง จนกระทั่งวิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์ ผู้วิจัยจึงขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.ธีระพร วีระถาวร รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา และ อาจารย์ ดร.อัคริน ทร์ ไพบูลย์พานิช ในฐานะประธานกรรมการและกรรมการสอบวิทยานิพนธ์ ที่กรุณาตรวจสอบและให้คำแนะนำอันเป็นประโยชน์ในการแก้ไขวิทยานิพนธ์ให้สมบูรณ์ยิ่งขึ้น และขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ได้ให้ความรู้แก่ผู้วิจัยจนกระทั่งสำเร็จการศึกษา

ผู้วิจัยขอกราบขอบพระคุณ บิดา มารดา และครอบครัว ที่ช่วยสนับสนุนในทุกๆ ด้านให้ผู้วิจัยได้มีโอกาสทางการศึกษาและคอยให้กำลังใจเสมอมาจนสำเร็จการศึกษา สุดท้ายนี้ขอขอบคุณเพื่อนๆ ทุกท่าน ที่คอยให้คำแนะนำ และให้กำลังใจในการทำวิทยานิพนธ์ตลอดมา

สารบัญ

บทที่ หน้า

บทที่ หน้า

สารบัญตาราง

สารบัญภาพ

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการตัดสินใจลงทุนในตลาดหลักทรัพย์ ผู้ลงทุนจะรับและประมวลผลข้อมูลจากหลายแหล่ง เพื่อสร้างตัวแบบค่าคาดหวังของผลตอบแทนในอนาคตของหุ้นแต่ละตัว ซึ่งค่าคาดหวังของผลตอบแทนนี้จะถูกใช้เป็นปัจจัยในการตัดสินใจลงทุน ในกรณีที่ข้อมูลทางเศรษฐศาสตร์เพียงชี้ชัดว่าหุ้นในกลุ่มอุตสาหกรรมหนึ่งดีกว่าหุ้นในอีกกลุ่มอุตสาหกรรมหนึ่ง เราอาจใช้ตัวแบบอสมการเชิงเส้นเชิงเบส์ (Bayesian Linear Inequalities Model) เสนอโดย Chiarawongse และคณะ (2009) ในการแก้ปัญหานี้ ตัวแบบดังกล่าวเป็นดังนี้

ถ้าให้ $\underset{\sim}{\alpha}$ เป็นค่าคาดหวังของผลตอบแทน ซึ่งเป็นเวกเตอร์สุ่มของหุ้น k ตัว ให้การแจกแจงก่อน (prior distribution) ของ $\underset{\sim}{\alpha}$ เป็นการแจกแจงแบบปกติหลายตัวแปร ซึ่งมีค่าเฉลี่ย $\underset{\sim}{\Pi}$ และเมทริกซ์ความแปรปรวนร่วม Σ โดยที่ Σ เป็นเมทริกซ์ที่เป็นบวกแน่นอน (positive definite matrix)

$$\underset{\sim}{\alpha} \sim N_k\left(\underset{\sim}{\Pi}, \Sigma\right)$$

และให้ทัศนะของค่าคาดหวังของผลตอบแทน แสดงอยู่ในรูปของระบบอสมการเชิงเส้น

$$A\underset{\sim}{\alpha} \le \underset{\sim}{b}$$

โดยที่ A เป็นเมทริกซ์ขนาด $M \times k$ และ $\underset{\sim}{b}$ เป็นเวกเตอร์ขนาด M

พิจารณาค่าคาดหวังของผลตอบแทนของหุ้น k ตัว $\underset{\sim}{\alpha} = \left(\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(k)}\right)$ การเรียงอันดับอย่างสมบูรณ์สามารถแสดงได้ด้วยอสมการเชิงเส้น k-1 อสมการ

$$\alpha^{(i_j)} \ge \alpha^{(i_{j+1})}$$

เมื่อ $j = 1,2,\ldots,k-1$ โดยที่ $(i_1, i_2, \ldots, i_k)$ เป็นการเรียงอันดับของ $(1,2,\ldots,k)$

ให้ $S = \left\{\underset{\sim}{\alpha} \in \Re^k \mid A\underset{\sim}{\alpha} \le \underset{\sim}{b}\right\}$ และ I_S เป็นตัวแปรบ่งชี้ (indicator variable) บน S ตัวแบบอสมการเชิงเส้นเชิงเบส์ นิยามโดย การแจกแจงภายหลัง (posterior distribution) ของค่าคาดหวังของผลตอบแทน ดังนี้

$$f\left(\underset{\sim}{\alpha} \mid S\right) \propto I_S \exp\left(-\frac{1}{2}\left[\left(\underset{\sim}{\alpha} - \underset{\sim}{\Pi}\right)^T \Sigma^{-1} \left(\underset{\sim}{\alpha} - \underset{\sim}{\Pi}\right)\right]\right)$$

ซึ่งการแจกแจงภายหลังนี้ เป็นการแจกแจงแบบปกติที่ถูกตัดหาง (truncated normal distribution)

การประมาณค่าคาดหวังของผลตอบแทนในตัวแบบนี้ เป็นดังนี้

$$\underset{\sim}{\hat{\alpha}} = E\left[\underset{\sim}{\alpha} \mid S\right] = \frac{1}{K}\int_{\underset{\sim}{x}\in S} \underset{\sim}{x} \exp\left(-\frac{1}{2}\left(\underset{\sim}{x}-\underset{\sim}{\Pi}\right)^T \Sigma^{-1}\left(\underset{\sim}{x}-\underset{\sim}{\Pi}\right)\right) d\underset{\sim}{x}$$

เมื่อ $K = \int_{\underset{\sim}{x}\in S} \exp\left(-\frac{1}{2}\left(\underset{\sim}{x}-\underset{\sim}{\Pi}\right)^T \Sigma^{-1}\left(\underset{\sim}{x}-\underset{\sim}{\Pi}\right)\right) d\underset{\sim}{x}$

การอนุมานสถิติเชิงเบส์เพื่อประมาณค่าคาดหวังของผลตอบแทนในตัวแบบดังกล่าวซึ่งเป็นค่าคาดหวังภายใต้การแจกแจงภายหลัง สามารถใช้เทคนิคในกลุ่มลูกโซ่มาร์คอฟมอนติคาร์โล (Markov Chain Monte Carlo หรือ MCMC) ซึ่ง Chiarawongse และคณะ (2009) เลือกใช้วิธี การสุ่มตัวอย่างแบบฮิตแอนด์รัน (Hit-and-Run sampler) ซึ่งเป็นวิธีหนึ่งในกลุ่มเทคนิคดังกล่าว ในการประมาณค่าคาดหวังของผลตอบแทน อย่างไรก็ตามเทคนิคในกลุ่มลูกโซ่มาร์คอฟมอนติคาร์โล มีหลายวิธี และ วิธีที่เป็นที่นิยมอย่างแพร่หลายวิธีหนึ่งได้แก่ การสุ่มตัวอย่างแบบกิบส์ (Gibbs sampler) ซึ่งสามารถใช้แก้ปัญหาดังกล่าวได้เช่นเดียวกัน ดังนั้นเพื่อให้การแก้ปัญหาดังกล่าวเป็นไปอย่างมีประสิทธิภาพ งานวิจัยชิ้นนี้จึงมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการอนุมานสถิติของตัวแบบอสมการเชิงเส้นเชิงเบส์ ระหว่างวิธีการสุ่มตัวอย่างแบบฮิตแอนด์รันและการสุ่มตัวอย่างแบบกิบส์

1.2 วัตถุประสงค์ของการวิจัย

เพื่อทำการเปรียบเทียบความเร็วในการลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โลสำหรับการอนุมานเชิงเบส์ เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง ระหว่างการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์

1.3 ขอบเขตของการวิจัย

1. เปรียบเทียบเฉพาะการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์

2. เกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพ ใช้2 เกณฑ์ คือ เปรียบเทียบด้วยค่าครึ่งช่วงความเชื่อมั่นที่คำนวณจากวิธีค่าเฉลี่ยกลุ่ม (batch means) และเปรียบเทียบด้วยค่า MPSRF ของ บรูกซ์-เกลแมน (Brooks–Gelman MPSRF)

3. จำนวนตัวแปร(มิติ)ของ $\underset{\sim}{\alpha}$ ที่ศึกษาเท่ากับ 2 ,3 ,4 ,5 ,7 และ 10 ตัวแปร

4. สำหรับการแจกแจงก่อน (prior distribution) ของ $\underset{\sim}{\alpha}$ ในตัวแบบอสมการเชิงเส้นเชิงเบส์ ซึ่งมีการแจกแจงแบบปกติหลายตัวแปร กำหนดให้ค่าเฉลี่ย $\underset{\sim}{\Pi}$ เท่ากับ $\underset{\sim}{0}$ และ เมทริกซ์ความแปรปรวนร่วม

$$\Sigma = \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \rho & \cdots & \rho & 1 & \rho \\ \rho & \cdots & \rho & \rho & 1 \end{bmatrix}$$ โดยที่ ρ มีค่า 0 ,0.5 และ 0.9

1.4 คำจำกัดความที่ใช้ในการวิจัย

การลู่เข้า (Convergence) หมายถึง ค่าเฉลี่ยของตัวอย่างมีแนวโน้มที่จะลู่เข้าสู่ค่าเฉลี่ยของประชากร เมื่อขนาดตัวอย่างเพิ่มขึ้น นั่นคือ

ให้ $\{X_n; n = 0,1,2,...\}$ เป็นกระบวนการเฟ้นสุ่มที่เป็นลูกโซ่มาร์คอฟเมื่อ X_n เป็นเวกเตอร์ของตัวแปรสุ่ม ซึ่ง $X_n \in S$ โดยที่ S เป็นปริภูมิสถานะ (state space) โดยเงื่อนไขทั่วไปบางเงื่อนไข(some regularity conditions) การแจกแจงของ X_n จะลู่เข้าสู่การแจกแจงคงตัว (stationary distribution) π และเป็นอิสระกับจุดเริ่มต้น x_0 ดังนี้

$$P(X_n \in A \mid X_0 = x_0) \to \pi(A) \quad \text{สำหรับทุก ๆ } A \in S$$

โดย Markov Chain Strong Law of Large Numbers (MCSLLN) จะได้ว่า

$$\frac{1}{N}\sum_{i=0}^{N-1} g(X_i) \to \pi g \quad \text{เมื่อ } \pi g = \int_S g(x)\pi(dx) \text{ สำหรับ } N \text{ ที่มีค่ามาก ๆ}$$

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อเป็นแนวทางในการตัดสินใจว่าควรใช้การสุ่มตัวอย่างแบบฮิตแอนด์รันหรือแบบกิบส์ในการจำลองข้อมูล ของการอนุมานเชิงเบส์ เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง

2. เพื่อทราบความเร็วในการลู่เข้า เมื่อหยุดการจำลองด้วยวิธีต่าง ๆ

1.6 วิธีดำเนินการวิจัย

1. ศึกษาค้นคว้าเอกสารและข้อมูลที่เกี่ยวข้องกับการวิจัย
2. จำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันให้เป็นไปตามขอบเขตที่กำหนด
3. จำลองข้อมูลด้วยการสุ่มตัวอย่างแบบกิบส์ให้เป็นไปตามขอบเขตที่กำหนด
4. คำนวณค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม
5. คำนวณค่า MPSRF ของ บรูกซ์-เกลแมน
6. วิเคราะห์ และสรุปผลการวิจัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดและทฤษฎี

2.1.1 การสุ่มตัวอย่างแบบฮิตแอนด์รัน (Hit-and-Run Sampler)

ขั้นตอนวิธีของการสุ่มตัวอย่างแบบฮิตแอนด์รัน ถูกนำเสนอโดย Smith R.L. โดยเป็นการสร้างลำดับของจุดตัวอย่างซึ่งลู่เข้าสู่การแจกแจงแบบสม่ำเสมอ (uniform distribution) บนบริเวณเปิดที่มีขอบเขตจำกัด $\Re^k$ ใดๆ ซึ่งต่อมาได้ถูกทำให้สามารถสร้างลำดับของจุดตัวอย่างที่มีการแจกแจงที่ต่อเนื่องของหลายตัวแปรใดๆ ได้ (Bélisle C.J.P., Romeijn H.E. and Smith R.L.(1993) และ Romeijn H.E. and Smith R.L.(1994)) โดยการสุ่มตัวอย่างแบบฮิตแอนด์รันนี้จัดว่าเป็นหนึ่งในวิธีที่เร็วที่สุด ในการสร้างจุดตัวอย่างในหลายมิติบนบริเวณรูปร่างนูน (Lovász L.(1998))

สำหรับการแจกแจงเป้าหมาย (target distribution) ทั่วไป การสุ่มตัวอย่างแบบฮิตแอนด์รันจะเป็นดังนี้

ให้ $S \subset \Re^k$ และให้การแจกแจงเป้าหมายมีฟังก์ชันความน่าจะเป็น π บน S

สร้าง $\underset{\sim}{X}_n$ ด้วยขั้นตอนวิธี ดังนี้

กำหนด $n = 0$ และ $\underset{\sim}{X}_n = \underset{\sim}{x}_0 \in S$

1.) ที่ $\underset{\sim}{x} = \underset{\sim}{x}_n \in S$ สุ่มเลือกทิศทาง $\underset{\sim}{d}$ บนพื้นผิวทรงกลมรัศมี 1 หน่วย ลากเส้นตรง L_n ผ่าน $\underset{\sim}{x}$ และตัดกับ S

$$L_n = S \cap \left\{ \underset{\sim}{l} : \underset{\sim}{l} = \underset{\sim}{x} + \lambda \underset{\sim}{d}, \lambda \in \Re \right\}$$

2.) สุ่มเลือกจุด $\underset{\sim}{x}_{n+1}$ บนเส้นตรง L_n ด้วยการแจกแจงแบบ π โดยมีเงื่อนไขบน L_n

3.) เพิ่ม n เป็น $n+1$ และกลับไป 1.)

2.1.2 การสุ่มตัวอย่างแบบกิบส์ (Gibbs Sampler)

การสุ่มตัวอย่างแบบกิบส์ เป็นขั้นตอนวิธีในการสร้างลำดับของจุดตัวอย่างจากฟังก์ชันการแจกแจงร่วม ของตัวแปรสุ่มตั้งแต่ 2 ตัวขึ้นไป ขั้นตอนวิธีของการสุ่มตัวอย่างแบบกิบส์ถูกอธิบายโดย Stuart and Donald Geman ในปีค.ศ.1984

สำหรับการแจกแจงเป้าหมายทั่วไป การสุ่มตัวอย่างแบบกิบส์จะเป็นดังนี้

ให้ S เป็นปริภูมิสถานะของ $\underset{\sim}{X} = \left(X^{(1)},\dots,\ X^{(k)}\right)$

ให้ π เป็นการแจกแจงเป้าหมาย ซึ่งมีฟังก์ชันความน่าจะเป็น $\pi'\left(x^{(1)},\dots,\ x^{(k)}\right)$

ให้ $\underset{\sim}{X}^{(-i)} = \left(X^{(1)},\dots,X^{(i-1)},X^{(i+1)},\dots,X^{(k)}\right)$ เป็นเวกเตอร์ $\underset{\sim}{X}$ ซึ่งไม่มี $X^{(i)}$

สมมติว่า เราสามารถสุ่มตัวอย่างจาก full conditional distribution $\pi'\left(x^{(i)} \mid \underset{\sim}{x}^{(-i)}\right)$

ขั้นตอนวิธีของการสุ่มตัวอย่างแบบกิบส์มีดังนี้

กำหนด $n = 0$ และเริ่มต้นที่ $\underset{\sim}{X}_0 = \left(X_0^{(1)},\dots,\ X_0^{(k)}\right)$

1.) สุ่มตัวอย่าง j แจกแจงแบบสม่ำเสมอ บน $\{1,\dots,\ k\}$

2.) สุ่มตัวอย่าง $X_{n+1}^{(j)}$ จาก $\pi'\left(x_n^{(j)} \mid \underset{\sim}{x}_n^{(-j)}\right)$

3.) ให้ $\underset{\sim}{X}_{n+1} = \left(X_n^{(1)},\dots,\ X_{n+1}^{(j)},\dots,\ X_n^{(k)}\right)$ เพิ่ม n เป็น $n+1$ และกลับไป 1.)

2.1.3 วิธีค่าเฉลี่ยกลุ่ม (Batch Means)

วิธีค่าเฉลี่ยกลุ่ม เป็นวิธีการหนึ่งที่ใช้ประมาณค่าคาดหวัง โดยใช้การคำนวณ Monte Carlo standard error (MCSE) สมมติว่าเราสนใจประมาณค่าคาดหวัง $E_\pi\left(g\left(X\right)\right)$ เมื่อ X มีการแจกแจง π

ขั้นตอนของวิธีค่าเฉลี่ยกลุ่ม มีดังนี้

1.) จำลองลูกโซ่มาร์คอฟ $\{X_i, i = 1,2,..., n\}$ จำนวน $n = ab$ รอบ โดยที่ a และ b เป็นจำนวนเต็ม a เป็นจำนวนกลุ่ม และแต่ละกลุ่มมีจำนวน b

2.) หาค่าเฉลี่ยของตัวอย่าง ในกลุ่มที่ j

$$\overline{y}_j = \frac{\sum_{i=(j-1)b+1}^{jb} g(x_i)}{b}$$

3.) ให้ $\hat{\sigma}_g^2 = \frac{b}{a-1}\sum_{j=1}^{a}\left(\overline{Y}_j - \overline{g}_n\right)^2$ โดยที่ $\overline{g}_n = \frac{\sum_{j=1}^{a}\overline{Y}_j}{a}$

จะได้ว่า ค่าประมาณของ Monte Carlo standard error คือ $\frac{\hat{\sigma}_g}{\sqrt{n}}$

4.) ทำการจำลองจนกระทั่ง $t_{\frac{\alpha}{2},(a-1)}\frac{\hat{\sigma}_g}{\sqrt{n}}$ มีค่าน้อยกว่าความคลาดเคลื่อนที่กำหนด

2.1.4 ค่า MPSRF ของ บรูกซ์-เกลแมน (Brooks–Gelman MPSRF)

การวิเคราะห์ของเกลแมนกับรูบิน (The Gelman-Rubin Diagnostic หรือ GRD) ถูกนำเสนอโดย Gelman and Rubin (1992) เป็นหนึ่งในวิธีที่ได้รับความนิยมที่สุดในการวิเคราะห์การลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โล ได้ถูกปรับปรุงใหม่และนำเสนอเพิ่มเติมโดย Brooks and Gelman (1998)

ขั้นตอนของวิธีของเกลแมนกับรูบิน มีดังนี้

1.) ทำการจำลองลูกโซ่มาร์คอฟที่อิสระกันจำนวน m ลูกโซ่(m independent parallel Markov Chains) $m \geq 2$ ซึ่งมีการแจกแจง π แต่ละลูกโซ่มีความยาว $2l$ ดังนั้นต้องทำการจำลองทั้งหมด $2lm$ รอบ Gelman and Rubin (1992) ได้แนะนำว่า การจำลอง l ค่าแรก ให้ตัดทิ้งไป และทำการอนุมานบนการจำลอง l ค่าสุดท้าย สำหรับลูกโซ่ที่ j แทนด้วย $\{X_{1j}, X_{2j},..., X_{lj}\}$ โดยที่ $j = 1,2,...,m$

2.) ให้ $Y_{ij} = g(X_{ij})$ ทำการคำนวณค่า

ความแปรปรวนระหว่างลูกโซ่ (between-chain variance)

$$\frac{B}{l}=\frac{1}{m-1}\sum_{j=1}^{m}\left(\overline{Y}_{.j}-\overline{Y}_{..}\right)^2$$

ค่าเฉลี่ยของความแปรปรวนภายในลูกโซ่ (the average of the m within-chain variance)

$$W=\frac{1}{m}\sum_{j=1}^{m}s_j^2$$

โดยที่ $\overline{Y}_{.j}=\frac{1}{l}\sum_{i=1}^{l}Y_{ij}$, $\overline{Y}_{..}=\frac{1}{m}\sum_{j=1}^{m}\overline{Y}_{.j}$ และ $s_j^2=\frac{1}{(l-1)}\sum_{i=1}^{l}\left(Y_{ij}-\overline{Y}_{.j}\right)^2$

$\overline{Y}_{..}$ เป็นผลการประมาณแบบจุดของ $E_\pi g$

3.) ประมาณค่าความแปรปรวนเป้าหมาย (target variance) โดยการเฉลี่ยถ่วงน้ำหนัก(weighted average) ค่า W และ B ได้เป็นดังนี้

$$\hat{\sigma}_+^2=\frac{l-1}{l}W+\frac{1}{l}B$$

4.) ให้ $\hat{V}=\hat{\sigma}_+^2+\frac{B}{ml}=\frac{l-1}{l}W+\frac{(m+1)B}{ml}$, $d\approx\frac{2\hat{V}^2}{\widehat{\mathrm{var}}(\hat{V})}$

เมื่อ $$\widehat{\mathrm{var}}(\hat{V})=\left(\frac{l-1}{l}\right)^2\frac{1}{m}\widehat{\mathrm{var}}\left(s_j^2\right)+\left(\frac{m+1}{ml}\right)^2\frac{2}{m-1}B^2+2\frac{(m+1)(l-1)}{ml^2}\frac{l}{m}\left[\widehat{\mathrm{cov}}\left(s_j^2,\overline{Y}_{.j}^2\right)-2\overline{Y}_{..}\widehat{\mathrm{cov}}\left(s_j^2,\overline{Y}_{.j}\right)\right]$$

นิยาม scale reduction factor

$$R=\frac{\hat{V}}{\sigma^2}$$

และประมาณค่า R โดย

$$\hat{R}=\frac{\hat{V}}{W}$$

นิยาม potential scale reduction factor ดังนี้

$$\frac{d}{d-2}\frac{\hat{V}}{W}$$

ต่อมา Brooks and Gelman (1998) ได้ปรับปรุงค่า potential scale reduction factor ใหม่เป็น

$$\hat{R}_c=\frac{d+3}{d+1}\hat{R}=\frac{d+3}{d+1}\frac{\hat{V}}{W}$$

ซึ่งเรียกว่า corrected scale reduction factor

เนื่องจาก ค่า $\hat{R}_c$ เป็นค่าอัตราส่วนระหว่างตัวประมาณของความแปรปรวน 2 ตัว โดยที่ตัวเศษเป็นตัวประมาณของความแปรปรวนที่มีค่าสูงเกินไป (overestimate) และตัวส่วนเป็นตัวประมาณของความแปรปรวนที่มีค่าต่ำเกินไป (underestimate) ดังนั้นการจำลองจะลู่เข้าเมื่อค่า $\hat{R}_c$ มีค่าเข้าใกล้(และมากกว่า) 1

สำหรับในกรณีหลายตัวแปร ซึ่งเราสนใจที่จะประมาณค่าของเวกเตอร์ของพารามิเตอร์ $\underset{\sim}{Y_{ij}}$ บรูกซ์และเกลแมน (Brooks and Gelman (1998)) ได้นำเสนอค่า multivariate potential scale reduction factor หรือ MPSRF ซึ่งเท่ากับ $\max\limits_{\underset{\sim}{a}} \dfrac{\underset{\sim}{a}^T \hat{V} \underset{\sim}{a}}{\underset{\sim}{a}^T W \underset{\sim}{a}}$ โดยที่

$$\hat{V} = \frac{l-1}{l}W + \left(1+\frac{1}{m}\right)\frac{B}{l}$$

$$W = \frac{1}{m(l-1)}\sum_{j=1}^{m}\sum_{i=1}^{l}\left(\underset{\sim}{Y_{ij}} - \underset{\sim}{\overline{Y}_{.j}}\right)\left(\underset{\sim}{Y_{ij}} - \underset{\sim}{\overline{Y}_{.j}}\right)^T$$

$$\frac{B}{l} = \frac{1}{m-1}\sum_{j=1}^{m}\left(\underset{\sim}{\overline{Y}_{.j}} - \underset{\sim}{\overline{Y}_{..}}\right)\left(\underset{\sim}{\overline{Y}_{.j}} - \underset{\sim}{\overline{Y}_{..}}\right)^T$$

ซึ่งใช้ตรวจสอบในลักษณะเดียวกับค่า potential scale reduction factor แต่เป็นในรูปแบบของหลายตัวแปร การจำลองจะลู่เข้าเมื่อค่า MPSRF มีค่าเข้าใกล้(และมากกว่า) 1 (ในทางปฏิบัติเราอาจทำการจำลองจนค่า MPSRF มีค่าน้อยกว่า 1.2 จึงจะถือว่าลู่เข้า)

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

หมิงฮุ่ย เชน และ บรูซ ชไมเซอร์ (Chen M.-H. and Schmeiser B.W., 1993) ได้ศึกษาเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ในกรณีที่เป็นการแจกแจงแบบปกติของ 2 ตัวแปร

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อทำการเปรียบเทียบความเร็วในการลู่เข้าของลูกโซ่มาร์คอฟมอนติคาร์โลสำหรับการอนุมานเชิงเบส์ เมื่อการแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง ระหว่างการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ ซึ่งข้อมูลที่นำมาใช้ในงานวิจัยนี้ ได้มาจากการจำลอง(Simulation)ด้วยโปรแกรม R ในบทนี้จะเป็นการนำเสนอวิธีดำเนินการวิจัยซึ่งผู้วิจัยได้ดำเนินการตามขั้นตอนต่อไปนี้

1. ศึกษาค้นคว้าเอกสารและข้อมูลที่เกี่ยวข้องกับการวิจัย
2. จำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันให้เป็นไปตามขอบเขตที่กำหนด
3. จำลองข้อมูลด้วยการสุ่มตัวอย่างแบบกิบส์ให้เป็นไปตามขอบเขตที่กำหนด
4. คำนวณค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม
5. คำนวณค่า MPSRF ของ บรูกซ์-เกลแมน
6. วิเคราะห์ และสรุปผลการวิจัย

โดยในขั้นตอนที่ 2-6 มีรายละเอียดของการดำเนินการดังนี้

3.1 การจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน

สำหรับขั้นตอนนี้จะทำการจำลองข้อมูลของพารามิเตอร์ซึ่งเป็นค่าคาดหวังของผลตอบแทน $\underset{\sim}{\alpha} = \left(\alpha^{(1)}, \alpha^{(2)}, ..., \alpha^{(k)}\right)$ โดยกำหนดให้

$$\underset{\sim}{\alpha} \sim N_k\left(\underset{\sim}{0}, \Sigma\right)$$

k แทนจำนวนมิติของ $\underset{\sim}{\alpha}$

n แทนจำนวนรอบที่ทำการจำลองข้อมูล

ขั้นตอนมีดังนี้

1.) กำหนดจำนวนมิติ k และจำนวนรอบที่ทำการจำลอง n

2.) จำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน

2.1) สุ่มเลือกจุดตัวอย่าง $\underset{\sim}{\alpha}_i = \underset{\sim}{\alpha}_1 = \left(\alpha_1^{(1)},...,\alpha_1^{(k)}\right)$ ซึ่ง $\alpha_1^{(1)} \geq \alpha_1^{(2)} \geq ... \geq \alpha_1^{(k)}$

2.2) สุ่มเลือกทิศทาง $\underset{\sim}{d}$ บนพื้นผิวทรงกลมใน $\Re^k$

2.3) สุ่มเลือก λ จากการแจกแจงแบบปกติซึ่งมีค่าเฉลี่ย $\underset{\sim}{d}^T \Sigma^{-1}\left(-\underset{\sim}{\alpha}_i\right) / \underset{\sim}{d}^T \Sigma^{-1} \underset{\sim}{d}$ และส่วนเบี่ยงเบนมาตรฐาน $\sqrt{1/\underset{\sim}{d}^T \Sigma^{-1} \underset{\sim}{d}}$ ซึ่งถูกจำกัดอยู่ในบริเวณ $\left[\inf\ val, \sup\ val\right]$ โดยที่

$$\inf\ val = \max\ \left\{\lambda_{bound}\ \left(\lambda_{bound} < 0\right), -\infty\right\}$$

$$\sup\ val = \min\ \left\{\lambda_{bound}\ \left(\lambda_{bound} > 0\right), \infty\right\}$$

เมื่อ

$$\lambda_{bound} = \alpha_i^{(1)} - \alpha_i^{(2)} / d^{(2)} - d^{(1)},...,\alpha_i^{(k-1)} - \alpha_i^{(k)} / d^{(k)} - d^{(k-1)}$$

2.4) สุ่มเลือกจุด $\underset{\sim}{\alpha}_{i+1} = \underset{\sim}{\alpha}_i + \lambda\, \underset{\sim}{d}$

2.5) เพิ่ม i เป็น $i+1$ และกลับไป 2.2) จนกระทั่งจำนวนจุดตัวอย่างเท่ากับ n

3.2 การจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบกิบส์

สำหรับขั้นตอนนี้จะทำการจำลองข้อมูลของพารามิเตอร์ซึ่งเป็นค่าคาดหวังของผลตอบแทน $\underset{\sim}{\alpha} = \left(\alpha^{(1)}, \alpha^{(2)},..., \alpha^{(k)}\right)$ โดยกำหนดให้

$\underset{\sim}{\alpha} \sim N_k\left(\underset{\sim}{0}, \Sigma\right)$

k แทนจำนวนมิติของ $\underset{\sim}{\alpha}$

n แทนจำนวนรอบที่ทำการจำลองข้อมูล

ขั้นตอนมีดังนี้

1.) กำหนดจำนวนมิติ k และจำนวนรอบที่ทำการจำลอง n

2.) จำลองข้อมูลด้วยการสุ่มตัวอย่างแบบกิบส์

2.1) สุ่มจุดตัวอย่าง $\underset{\sim}{\alpha}_i = \underset{\sim}{\alpha}_1 = \left(\alpha_1^{(1)},...,\alpha_1^{(j)},...,\alpha_1^{(k)}\right)$ ซึ่ง $\alpha_1^{(1)} \geq ... \geq \alpha_1^{(j)} \geq ... \geq \alpha_1^{(k)}$

2.2) สุ่มเลือก j แจกแจงแบบสม่ำเสมอ บน $\{1,..., k\}$

2.3) สุ่ม $\alpha_{i+1}^{(j)}$ จากการแจกแจงแบบปกติซึ่งมีค่าเฉลี่ย $-\left(\sum_{\substack{q=1 \\ q \neq p}}^{k} \alpha_i^{(q)} \rho_{p,q}\right) / \rho_{p,p}$ และส่วนเบี่ยงเบนมาตรฐาน $\sqrt{1/\rho_{p,p}}$ ซึ่งถูกจำกัดอยู่ในบริเวณ $[\inf\ val, \sup\ val]$ โดยที่ $\rho_{p,q}$ เป็นสมาชิกแถวที่ p หลักที่ q ใน Σ^{-1}

$$\inf\ val = \begin{cases} -\infty, p = k \\ \alpha_i^{(p+1)}, p \neq k \end{cases}$$

$$\sup\ val = \begin{cases} \infty, p = 1 \\ \alpha_i^{(p-1)}, p \neq 1 \end{cases}$$

2.4) ให้ $\underset{\sim}{\alpha}_{i+1} = \left(\alpha_i^{(1)},...,\alpha_{i+1}^{(j)},...,\alpha_i^{(k)}\right)$

2.5) เพิ่ม i เป็น $i+1$ และกลับไป 2.2) จนกระทั่งจำนวนจุดตัวอย่างเท่ากับ n

3.3 การคำนวณค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม

หลังจากทำการจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและการสุ่มตัวอย่างแบบกิบส์ตามขั้นตอน 3.1 และ 3.2 แล้ว จะได้จุดตัวอย่าง $\left\{\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2, ..., \underset{\sim}{\alpha}_n\right\}$ จำนวน n จุด จากนั้นทำตามขั้นตอนดังนี้

1.) ทำการตัดข้อมูล (burn in) 10 เปอร์เซนต์แรกทิ้งไป และนำ 90 เปอร์เซนต์ที่เหลือมาแบ่งเป็น 30 กลุ่ม กลุ่มละเท่าๆ กัน

2.) หาค่าเฉลี่ยของตัวอย่างในแต่ละกลุ่ม และนำค่าเฉลี่ยของตัวอย่างในแต่ละกลุ่มมาหาค่าเฉลี่ยอีกครั้งหนึ่ง

3.) คำนวณค่าประมาณของ Monte Carlo standard error ซึ่งเท่ากับ $\dfrac{\hat{\sigma}_g}{\sqrt{(0.9)n}}$

4.) คำนวณค่าครึ่งช่วงความเชื่อมั่น ซึ่งเท่ากับ $t_{0.975,(29)} \dfrac{\hat{\sigma}_g}{\sqrt{(0.9)n}}$

3.4 การคำนวณค่า MPSRF ของ บรูกซ์-เกลแมน

หลังจากการจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและการสุ่มตัวอย่างแบบกิบส์ตามขั้นตอน 3.1 และ 3.2 โดยจำลองข้อมูลเป็นลูกโซ่ที่อิสระกัน 5 ลูกโซ่ แล้ว จะได้จุดตัวอย่างของแต่ละลูกโซ่ $\left\{ \underset{\sim}{\alpha}_{1j}, \underset{\sim}{\alpha}_{2j}, ..., \underset{\sim}{\alpha}_{nj} \right\} : j = 1,...,5$ จำนวนลูกโซ่ละ n จุด จากนั้นทำตามขั้นตอนต่อไปนี้

1.) ทำการตัดข้อมูล (burn in) ครึ่งแรกของทุกๆ ลูกโซ่ทิ้งไป

2.) วิเคราะห์การลู่เข้าด้วย package coda ในโปรแกรม R

3.) คำนวณค่า MPSRF โดยใช้ฟังก์ชัน gelman.diag() ใน package coda

3.5 การวิเคราะห์และสรุปผลการวิจัย

ขั้นตอนมีดังนี้

1.) นำค่าครึ่งช่วงความเชื่อมั่นจากการสุ่มตัวอย่างแบบฮิตแอนด์รัน และค่าครึ่งช่วงความเชื่อมั่นจากการสุ่มตัวอย่างแบบกิบส์ ในกรณีที่จำนวนมิติเท่ากัน มาเปรียบเทียบกัน โดยใช้ค่าครึ่งช่วงความเชื่อมั่นของมิติที่ 1 ($\alpha^{(1)}$) มาเป็นตัวแทนในการเปรียบเทียบ ถ้าค่าครึ่งช่วงความเชื่อมั่นมีค่าน้อยกว่าจะมีประสิทธิภาพมากกว่า

2.) นำค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รัน และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ในกรณีที่จำนวนมิติเท่ากัน มาเปรียบเทียบกัน โดยถ้าค่า MPSRF มีค่าน้อยกว่าจะมีประสิทธิภาพมากกว่า

3.) วิเคราะห์ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ ในกรณีที่มีมิติต่างๆ ที่กำหนด โดยใช้เกณฑ์ในการเปรียบเทียบประสิทธิภาพทั้งสองเกณฑ์ และทำการสรุปผล

บทที่ 4

ผลการวิจัย

ในงานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพของการอนุมานเชิงเบส์สำหรับการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ ในกรณีที่การแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง เมื่อจำนวนตัวแปร(มิติ)เป็น 2, 3, 4, 5, 7 และ 10 ตัวแปร และเมทริกซ์ความแปรปรวนร่วม $\sum$ มีค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) เป็น 0, 0.5 และ 0.9 โดยใช้เกณฑ์ในการวัดประสิทธิภาพ 2 เกณฑ์ คือ ค่าครึ่งช่วงความเชื่อมั่นที่คำนวณจากวิธีค่าเฉลี่ยกลุ่ม และ ค่า MPSRF ของ บรูกซ์-เกลแมน

การนำเสนอผลการวิจัยนี้ ได้มีการนำเสนอในรูปแบบตารางและรูปภาพ โดยมีการใช้สัญลักษณ์แทนความหมายต่างๆ ดังนี้

k หมายถึง จำนวนตัวแปร (มิติ) ของ $\underset{\sim}{\alpha}$

n หมายถึง จำนวนรอบที่ทำการจำลองข้อมูลตัวอย่างในแต่ละลูกโซ่

half width หมายถึง ค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม

H-R หมายถึง การสุ่มตัวอย่างแบบฮิตแอนด์รัน

Gibbs หมายถึง การสุ่มตัวอย่างแบบกิบส์

MPSRF หมายถึง ค่า MPSRF ของบรูกซ์-เกลแมน

ซึ่งรายละเอียดของผลการวิจัยมีดังนี้

4.1 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม

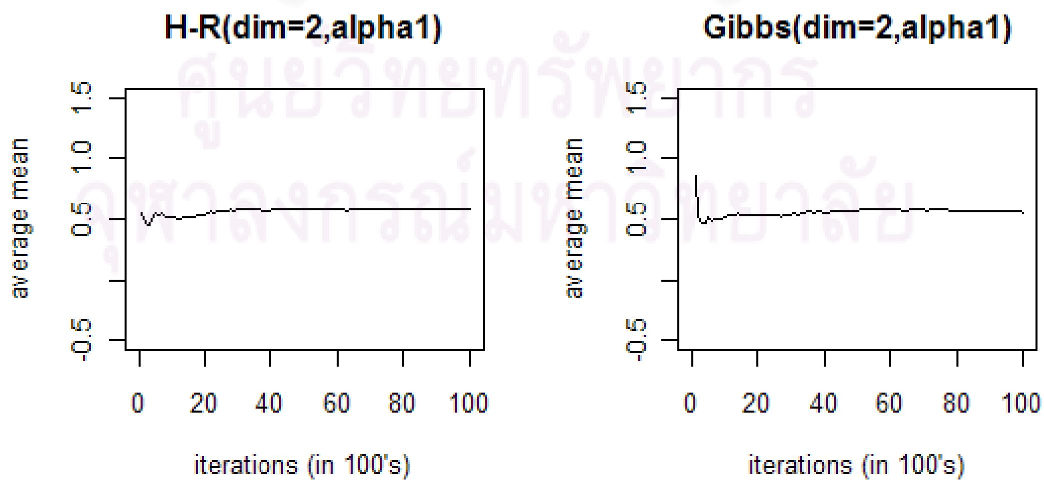
สำหรับการเปรียบเทียบโดยใช้เกณฑ์ของค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่มนี้ จะเลือกเอาค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 ($\alpha^{(1)}$) มาเป็นตัวแทนในการพิจารณาเปรียบเทียบในทุกๆ กรณี และกำหนดจำนวนรอบในการจำลองสูงสุดไม่เกิน 500,000 รอบ

4.1.1 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0

ตารางที่ 4.1 แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0

k	n	Parameter	half width	
			Hit-and-Run	Gibbs
2	10,000	$\alpha^{(1)}$	0.0295	0.0354
3	15,000	$\alpha^{(1)}$	0.0398	0.0527
4	20,000	$\alpha^{(1)}$	0.0394	0.0403
5	25,000	$\alpha^{(1)}$	0.0571	0.0383
7	35,000	$\alpha^{(1)}$	0.0730	0.0497
10	50,000	$\alpha^{(1)}$	0.1046	0.0510

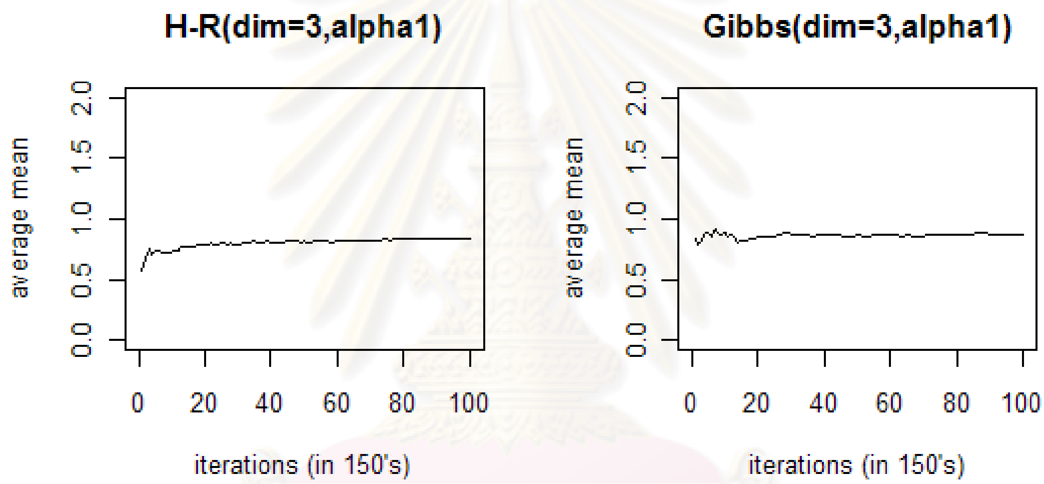
<u>กรณี 2 มิติ</u>



รูปที่ 4.1 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 2 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 10,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.1 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 10,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.1 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0295 น้อยกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0354 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

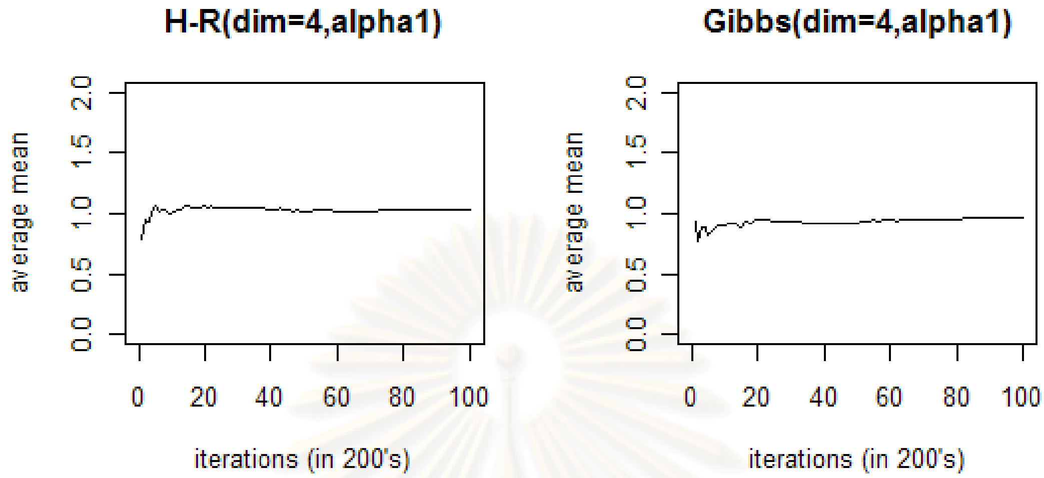
<u>กรณี 3 มิติ</u>



รูปที่ 4.2 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 3 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 15,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.2 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 15,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.1 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0398 น้อยกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0527 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

<u>กรณี 4 มิติ</u>

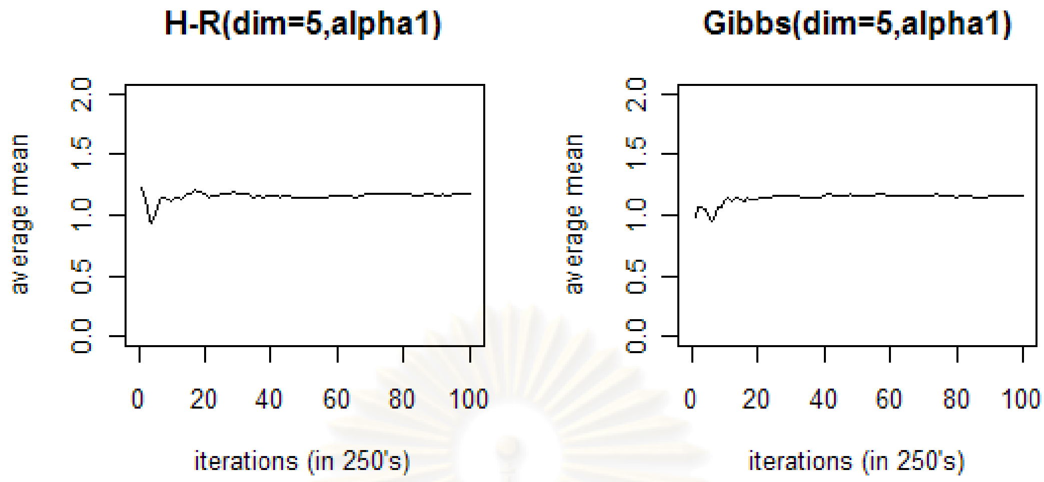


รูปที่ 4.3 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 4 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 20,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.3 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 20,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.1 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0394 น้อยกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0403 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

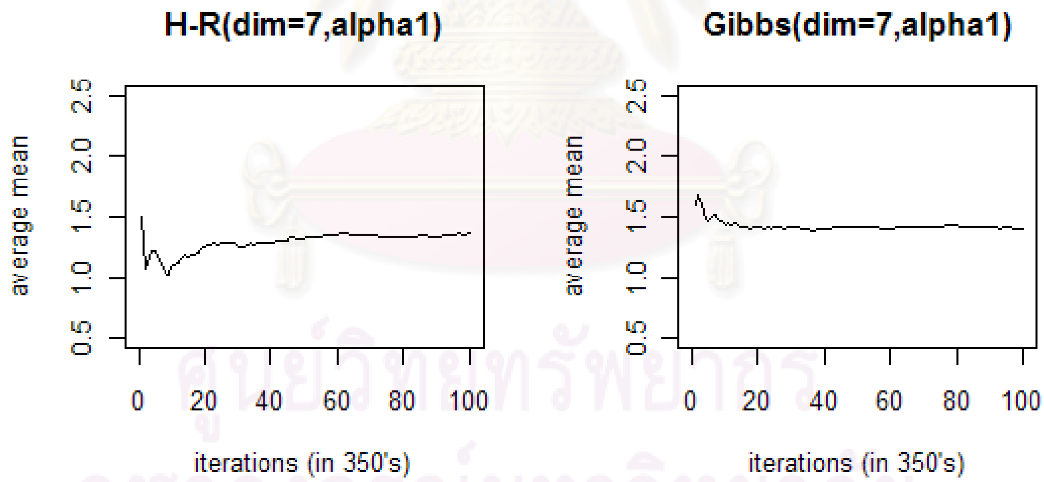
<u>กรณี 5 มิติ</u>

จากรูปที่ 4.4 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 25,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.1 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0571 มากกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0383 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน



รูปที่ 4.4 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 5 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 25,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

<u>กรณี 7 มิติ</u>

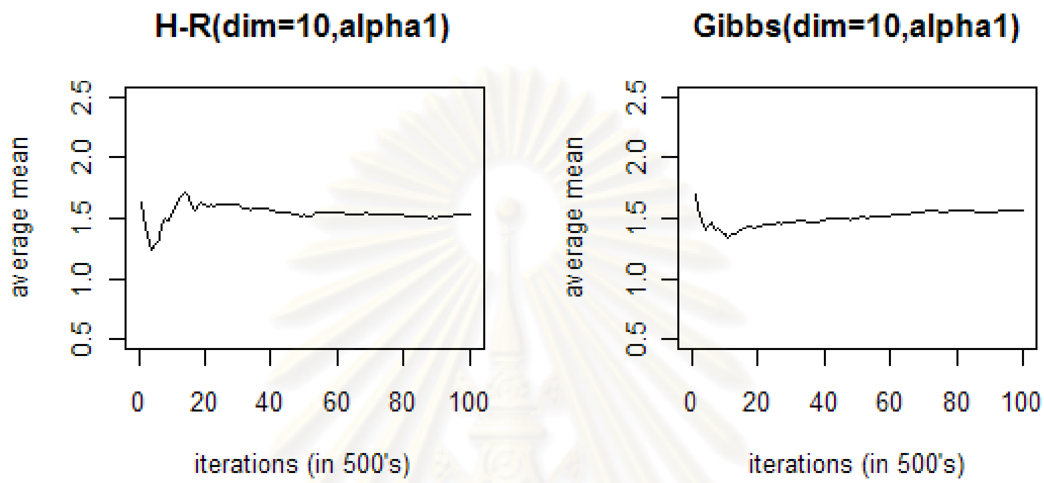


รูปที่ 4.5 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 7 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 35,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.5 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 35,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.1 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0730 มากกว่า

ค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0497 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

<u>กรณี 10 มิติ</u>



รูปที่ 4.6 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 10 มิติ และ $\rho = 0$ เมื่อทำการจำลองข้อมูล 50,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

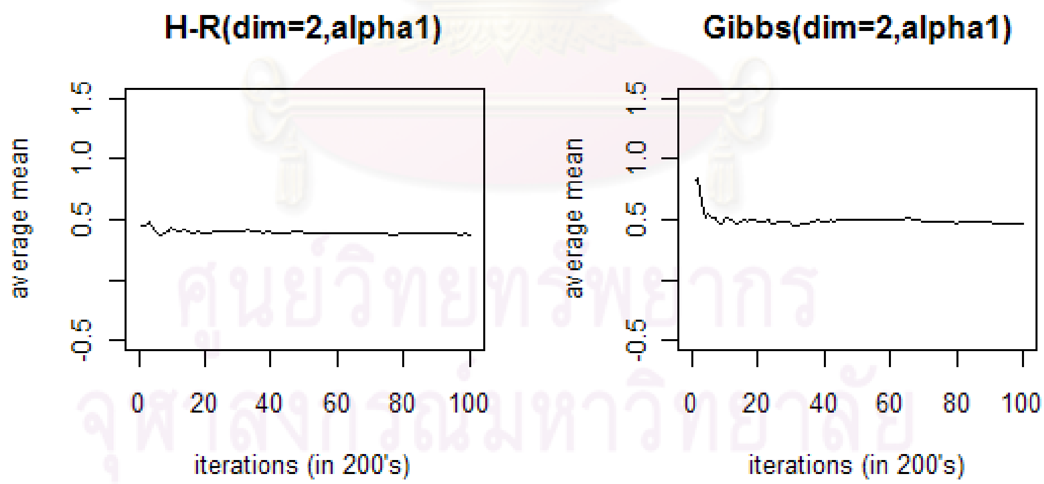
จากรูปที่ 4.6 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 50,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.1 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.1046 มากกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0510 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

4.1.2 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.5

ตารางที่ 4.2 แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น0.5

k	n	Parameter	half width	
			Hit-and-run	Gibbs
2	20,000	$\alpha^{(1)}$	0.0344	0.0541
3	45,000	$\alpha^{(1)}$	0.0522	0.0639
4	80,000	$\alpha^{(1)}$	0.0551	0.0808
5	125,000	$\alpha^{(1)}$	0.0887	0.0861
7	245,000	$\alpha^{(1)}$	0.1314	0.1163
10	500,000	$\alpha^{(1)}$	0.2130	0.1650

<u>กรณี 2 มิติ</u>

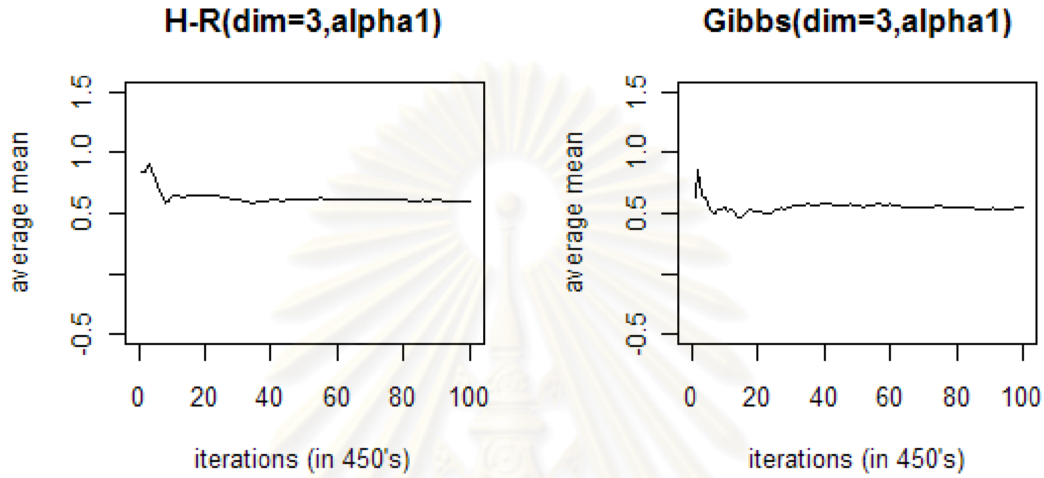


รูปที่ 4.7 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 2 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 20,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.7 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 20,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.2 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0344 น้อยกว่า

ค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0541 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

กรณี 3 มิติ

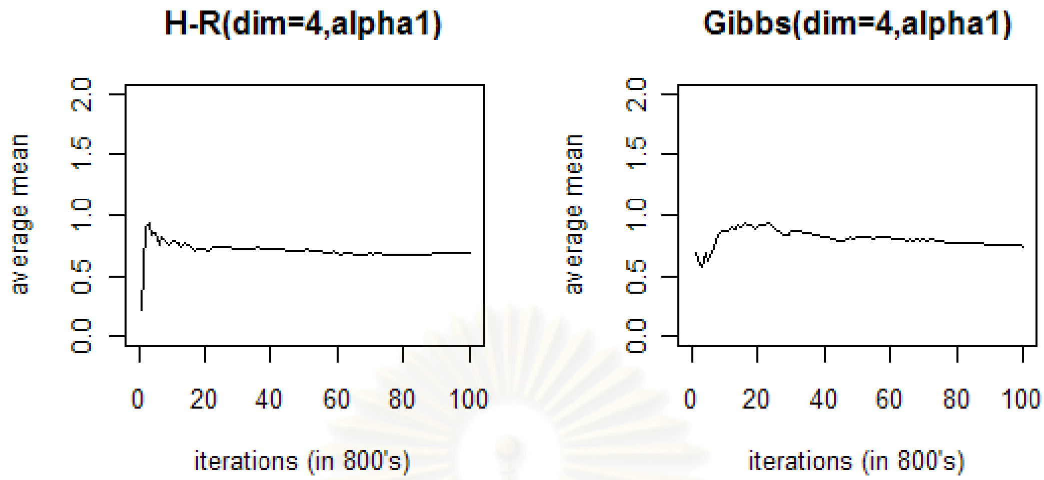


รูปที่ 4.8 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 3 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 45,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.8 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 45,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.2 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0522 น้อยกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0639 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

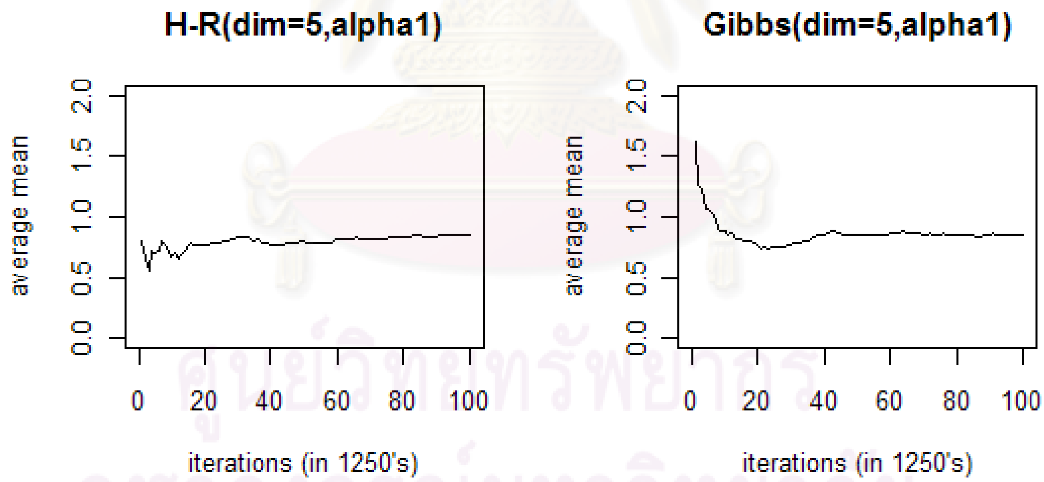
กรณี 4 มิติ

จากรูปที่ 4.9 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 80,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.2 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0551 น้อยกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0808 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์



รูปที่ 4.9 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 4 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 80,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

<u>กรณี 5 มิติ</u>

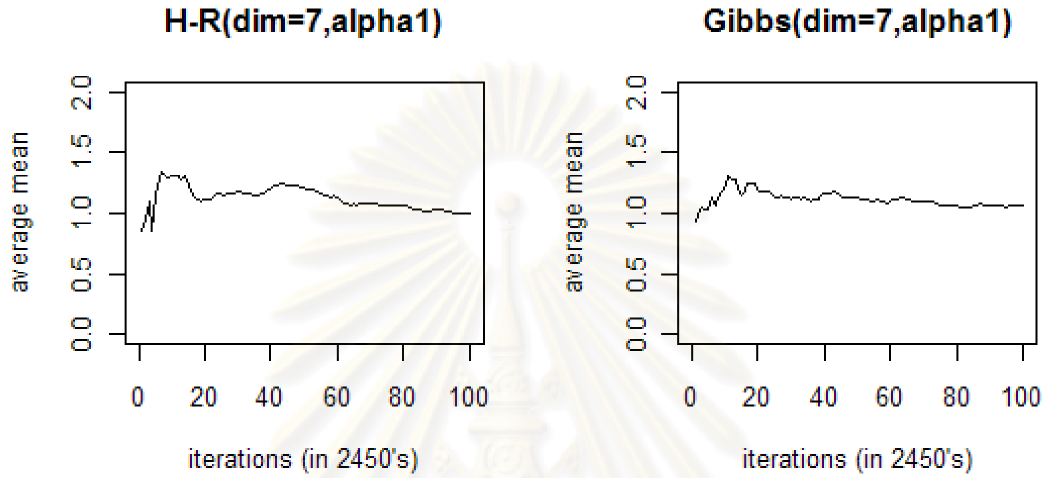


รูปที่ 4.10 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 5 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 125,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.10 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 125,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.2 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0887 มากกว่า

ค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.0861 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

กรณี 7 มิติ

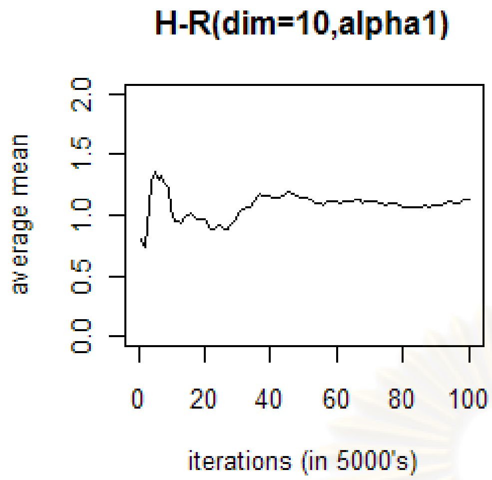


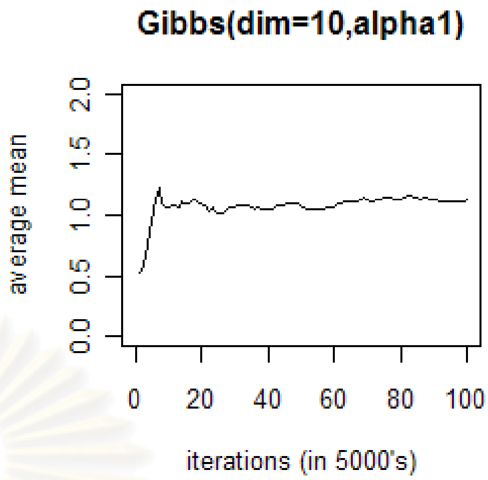
รูปที่ 4.11 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 7 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 245,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.11 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 245,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์เริ่มที่จะลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.2 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.1314 มากกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.1163 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน แต่อย่างไรก็ตาม เนื่องจากค่าทั้ง 2 มีค่าค่อนข้างสูง การเปรียบเทียบในกรณีนี้อาจให้ผลที่ไม่ดีนัก

กรณี 10 มิติ

จากรูปที่ 4.12 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 500,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์เริ่มที่จะลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.2 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.2130 มากกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.1650 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน แต่อย่างไรก็ตาม เนื่องจากค่าทั้ง 2 มีค่าค่อนข้างสูง การเปรียบเทียบในกรณีนี้อาจให้ผลที่ไม่ดีนัก





รูปที่ 4.12 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 10 มิติ และ $\rho = 0.5$ เมื่อทำการจำลองข้อมูล 500,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

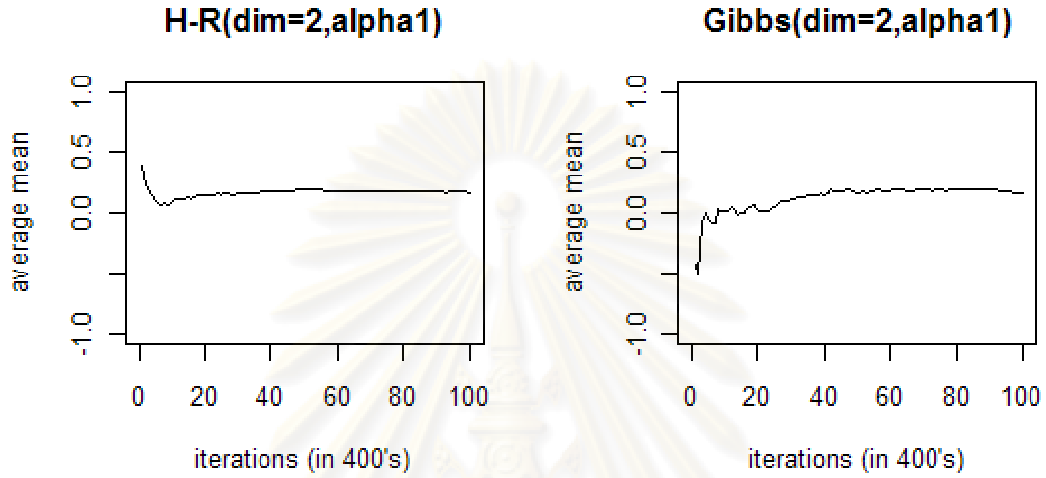
4.1.3 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9

ตารางที่ 4.3 แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9

k	n	Parameter	half width	
			Hit-and-run	Gibbs
2	40,000	$\alpha^{(1)}$	0.0413	0.1048
3	135,000	$\alpha^{(1)}$	0.0556	0.1412
4	320,000	$\alpha^{(1)}$	0.0629	0.1050
5	500,000	$\alpha^{(1)}$	0.1323	0.1144
*7	500,000	$\alpha^{(1)}$	*0.2099	*0.2816
*10	500,000	$\alpha^{(1)}$	*0.2974	*0.3265

* หมายเหตุ ค่าครึ่งช่วงความเชื่อมั่นในกรณี 7 และ 10 มิติ มีค่าสูง ซึ่งหมายความว่าทั้งการจำลองด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์ ยังไม่ลู่เข้า

กรณี 2 มิติ

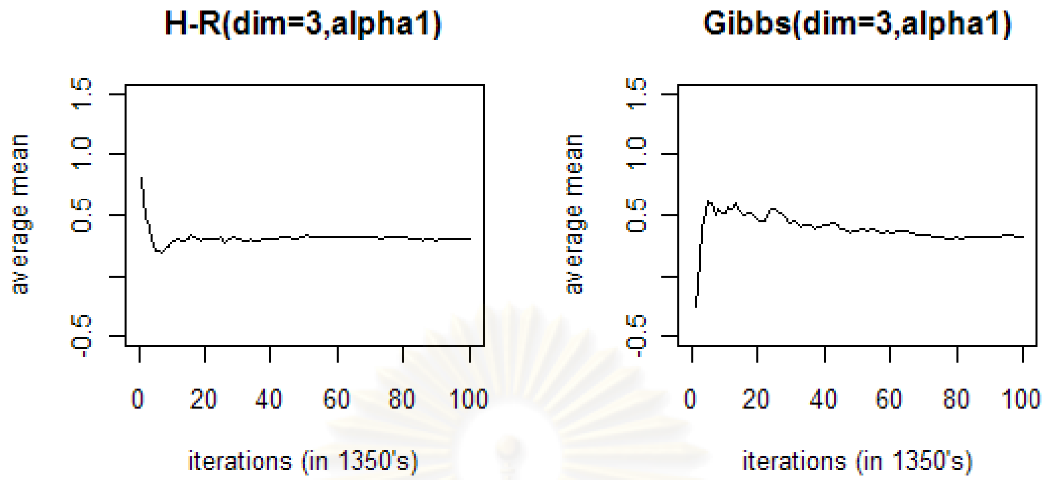


รูปที่ 4.13 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 2 มิติ และ $\rho = 0.9$ เมื่อทำการจำลองข้อมูล 40,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.13 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 40,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.3 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0413 น้อยกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.1048 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

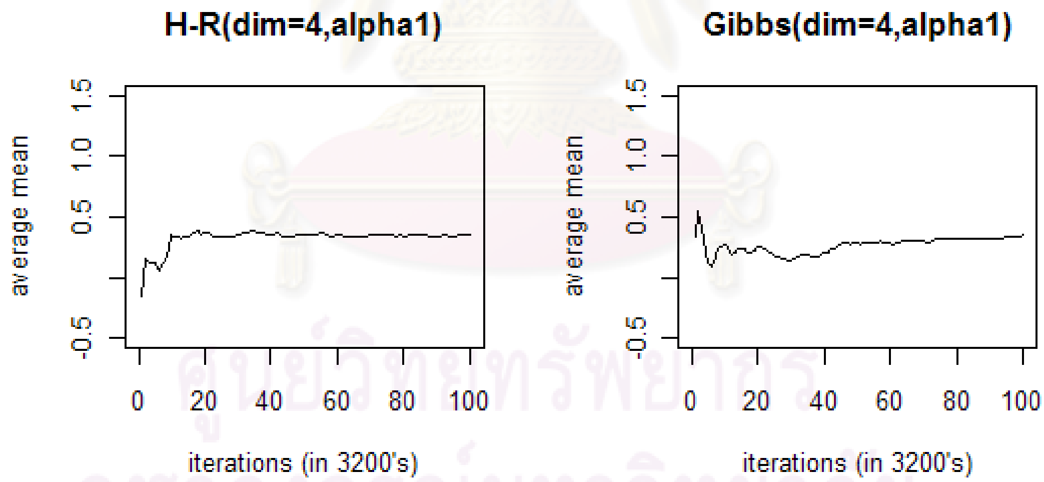
กรณี 3 มิติ

จากรูปที่ 4.14 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 135,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.3 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0556 น้อยกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.1412 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์



รูปที่ 4.14 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 3 มิติ และ $\rho = 0.9$ เมื่อทำการจำลองข้อมูล 135,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

<u>กรณี 4 มิติ</u>

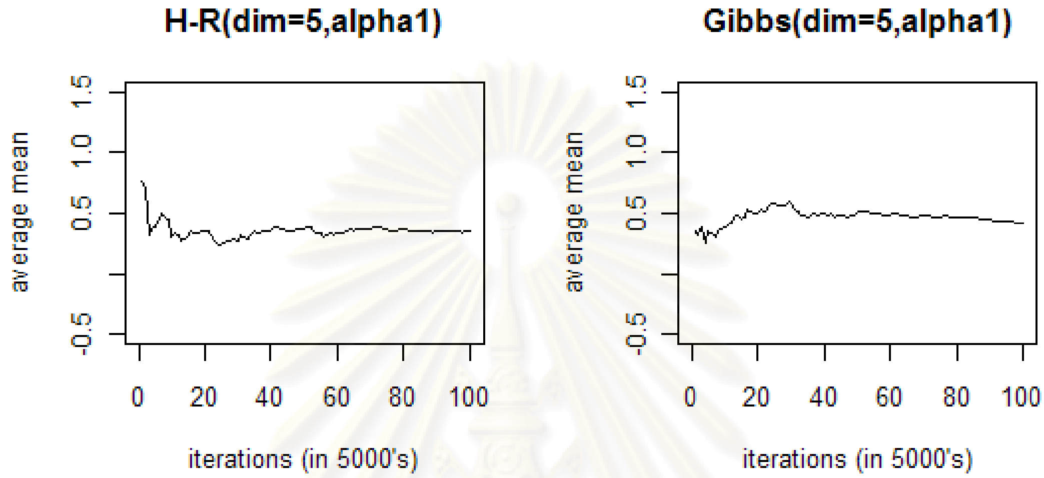


รูปที่ 4.15 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 4 มิติ และ $\rho = 0.9$ เมื่อทำการจำลองข้อมูล 320,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.15 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 320,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.3 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.0629 น้อยกว่า

ค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.1050 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

<u>กรณี 5 มิติ</u>

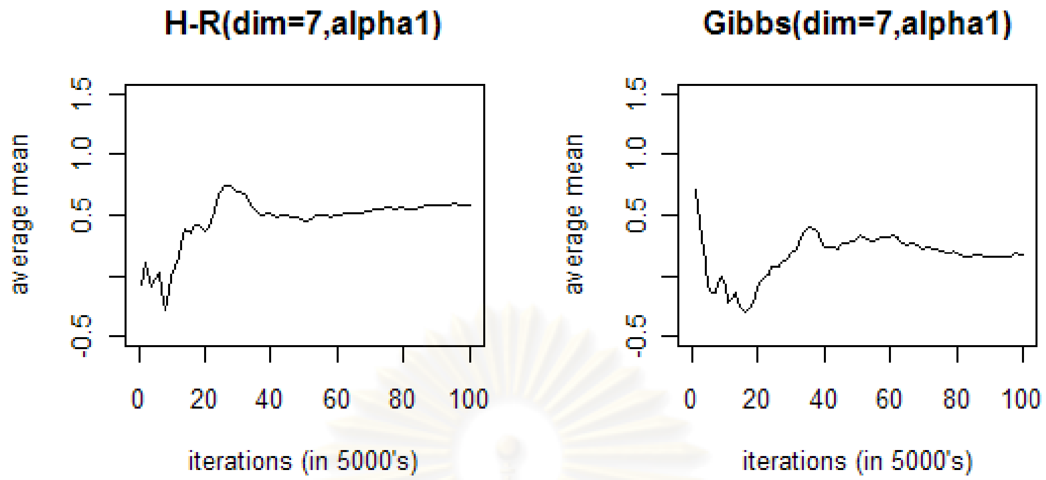


รูปที่ 4.16 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 5 มิติ และ $\rho = 0.9$ เมื่อทำการจำลองข้อมูล 500,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.16 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 500,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ลู่เข้าแล้ว และเมื่อพิจารณาค่าจากในตารางที่ 4.3 พบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.1323 มากกว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์ซึ่งมีค่า 0.1144 ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

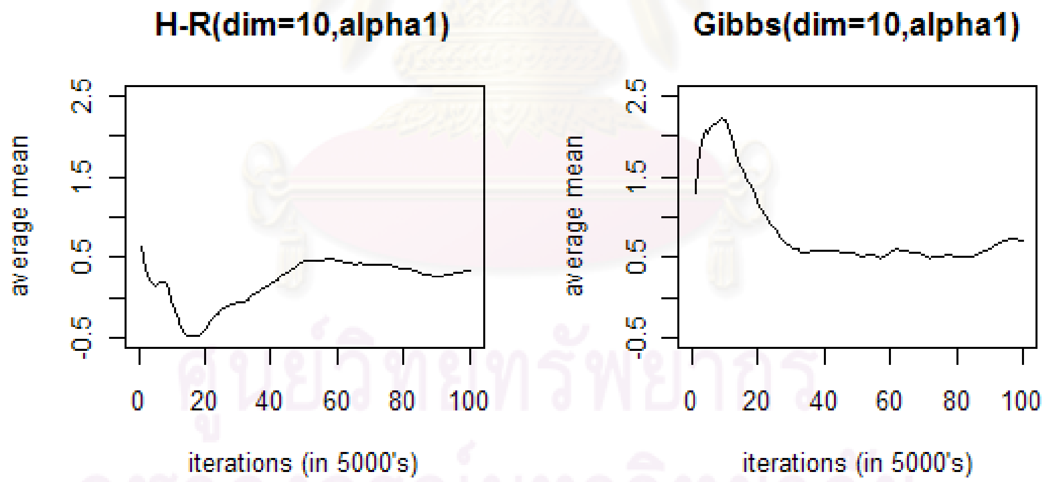
<u>กรณี 7 มิติ</u>

จากรูปที่ 4.17 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 500,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ยังไม่ลู่เข้า และเมื่อพิจารณาค่าจากในตารางที่ 4.3 จะพบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.2099 และค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์มีค่า 0.2816 ซึ่งมีค่าสูงทั้ง 2 วิธี ดังนั้นจึงไม่เหมาะที่จะนำค่าทั้งสองมาเปรียบเทียบกัน



รูปที่ 4.17 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 7 มิติ และ $\rho = 0.9$ เมื่อทำการจำลองข้อมูล 500,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

<u>กรณี 10 มิติ</u>



รูปที่ 4.18 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ในกรณี 10 มิติ และ $\rho = 0.9$ เมื่อทำการจำลองข้อมูล 500,000 รอบ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

จากรูปที่ 4.18 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 500,000 รอบ ทั้งการสุ่มตัวอย่างแบบฮิตแอนด์รัน และการสุ่มตัวอย่างแบบกิบส์ยังไม่ลู่เข้า และเมื่อพิจารณาค่าจากในตารางที่ 4.3 จะพบว่าค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 0.2974

และค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 จากการสุ่มตัวอย่างแบบกิบส์มีค่า 0.3265 ซึ่งมีค่าสูงทั้ง 2 วิธี ดังนั้นจึงไม่เหมาะที่จะนำค่าทั้งสองมาเปรียบเทียบกัน

4.2 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรูกซ์-เกลแมน

สำหรับการเปรียบเทียบโดยใช้เกณฑ์นี้ จะกำหนดจำนวนรอบในการจำลองสูงสุดไม่เกิน 300,000 รอบต่อลูกโซ่

4.2.1 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรูกซ์-เกลแมน เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0

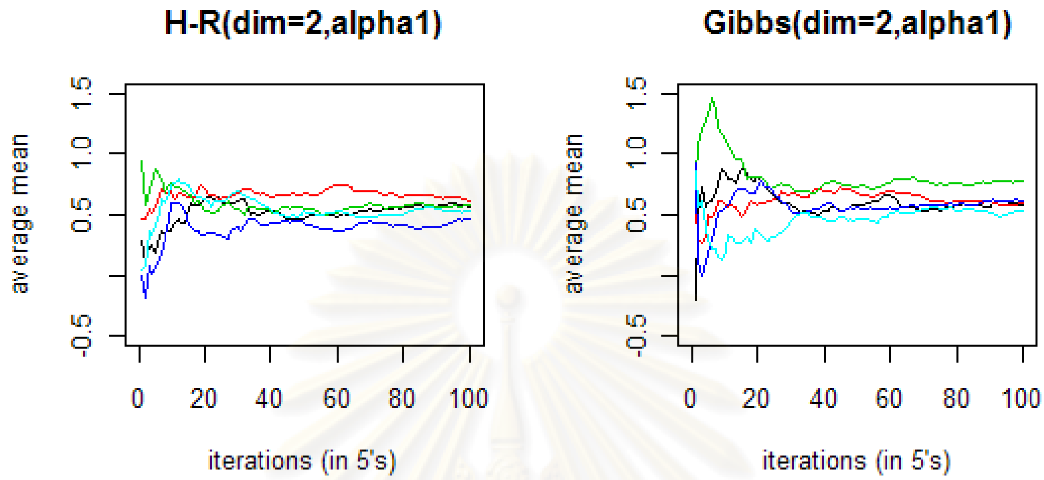
ตารางที่ 4.4 แสดงค่า MPSRF ของ บรูกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0

k	n	MPSRF	
		Hit-and-Run	Gibbs
2	1,000	1.01	1.02
3	2,000	1.03	1.04
4	3,000	1.02	1.03
5	4,000	1.03	1.01
7	6,000	1.09	1.08
10	30,000	1.09	1.02

<u>กรณี 2 มิติ</u>

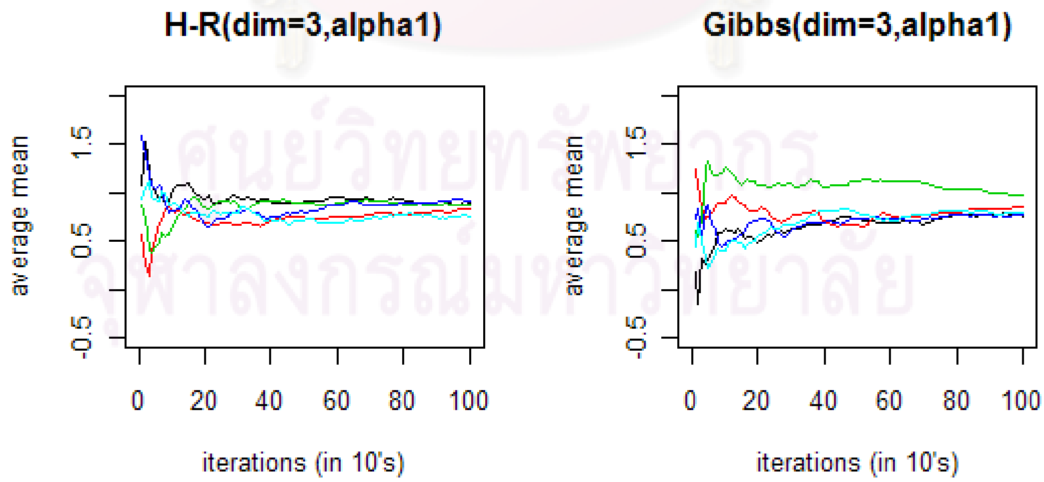
เมื่อพิจารณาจากตารางที่ 4.4 และรูปที่ 4.19 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 1,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.01 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.02 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่ม

ตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์



รูปที่ 4.19 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 1,000 รอบต่อลูกโซ่ ในกรณี 2 มิติ และ $\rho = 0$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

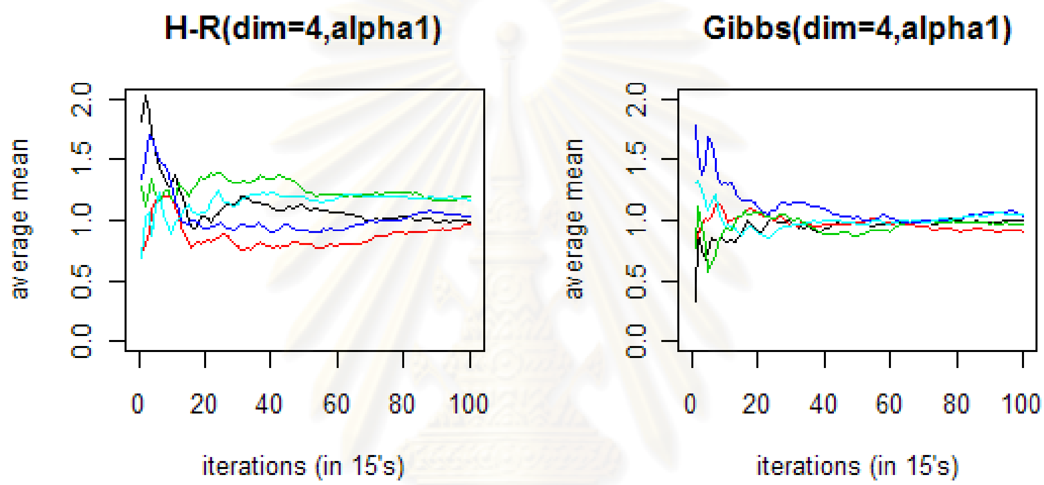
<u>กรณี 3 มิติ</u>



รูปที่ 4.20 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 2,000 รอบต่อลูกโซ่ ในกรณี 3 มิติ และ $\rho = 0$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.4 และรูปที่ 4.20 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 2,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.03 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.04 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

<u>กรณี 4 มิติ</u>



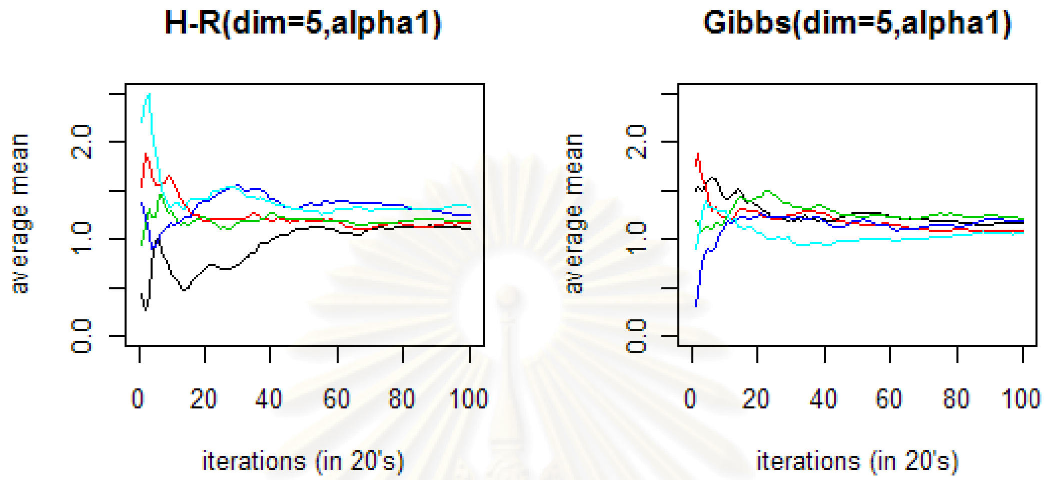
รูปที่ 4.21 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 3,000 รอบต่อลูกโซ่ ในกรณี 4 มิติ และ $\rho = 0$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.4 และรูปที่ 4.21 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 3,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.02 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.03 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

<u>กรณี 5 มิติ</u>

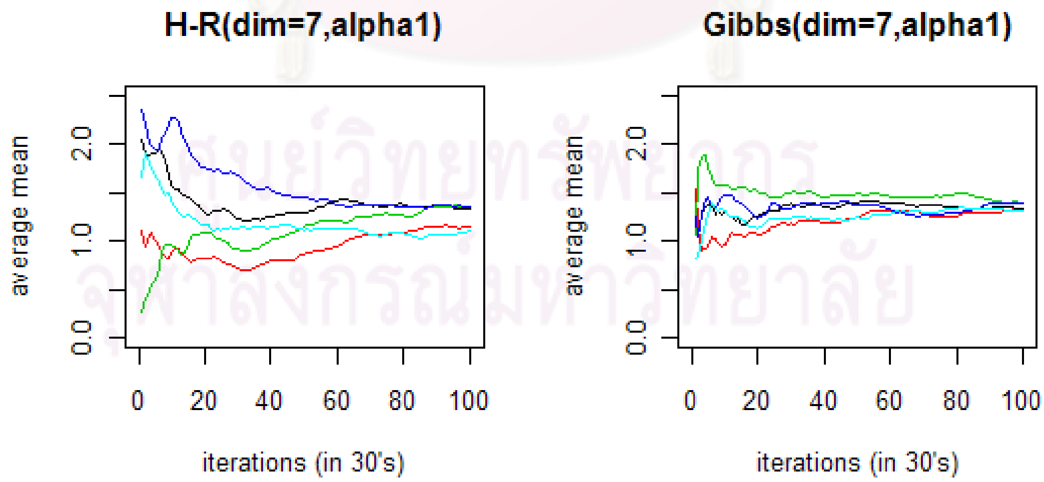
เมื่อพิจารณาจากตารางที่ 4.4 และรูปที่ 4.22 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 4,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.03 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.01 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่ม

ตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน



รูปที่ 4.22 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 4,000 รอบต่อลูกโซ่ ในกรณี 5 มิติ และ $\rho = 0$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

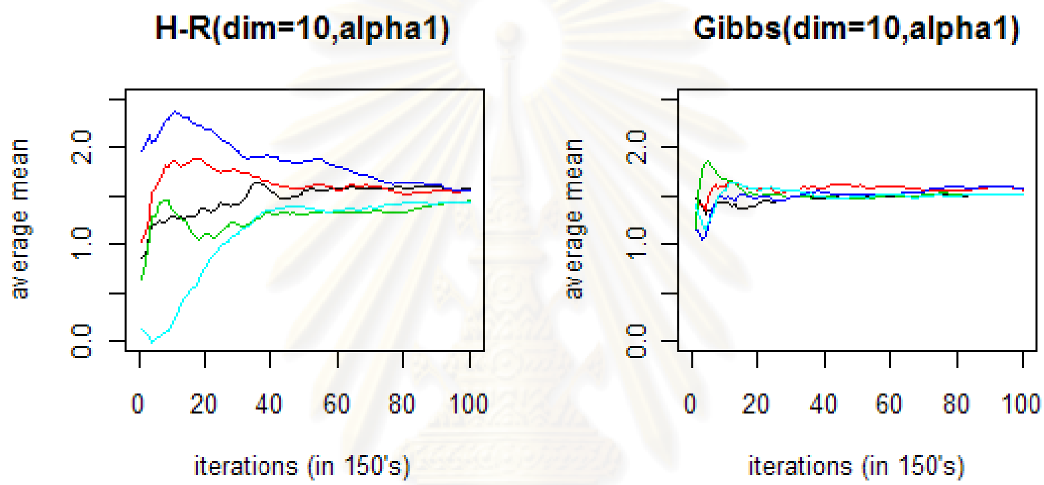
<u>กรณี 7 มิติ</u>



รูปที่ 4.23 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 6,000 รอบต่อลูกโซ่ ในกรณี 7 มิติ และ $\rho = 0$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.4 และรูปที่ 4.23 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 6,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.09 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.08 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

<u>กรณี 10 มิติ</u>



รูปที่ 4.24 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 30,000 รอบต่อลูกโซ่ ในกรณี 10 มิติ และ $\rho = 0$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

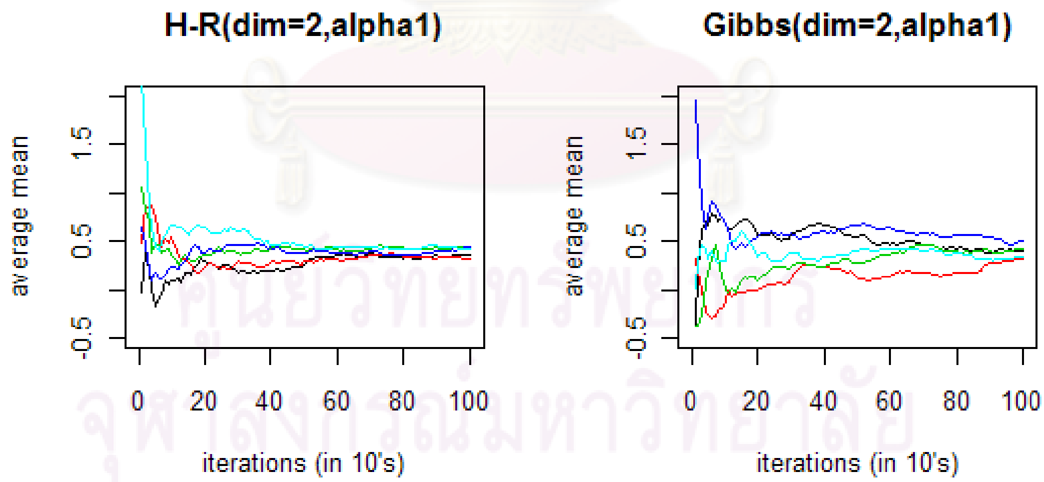
เมื่อพิจารณาจากตารางที่ 4.4 และรูปที่ 4.24 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 30,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.09 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.02 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

4.2.2 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรูกซ์-เกลแมน เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.5

ตารางที่ 4.5 แสดงค่า MPSRF ของ บรูกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.5

k	n	MPSRF	
		Hit-and-Run	Gibbs
2	2,000	1.01	1.02
3	6,000	1.01	1.02
4	12,000	1.01	1.02
5	20,000	1.04	1.01
7	42,000	1.07	1.01
10	300,000	1.08	1.03

กรณี 2 มิติ

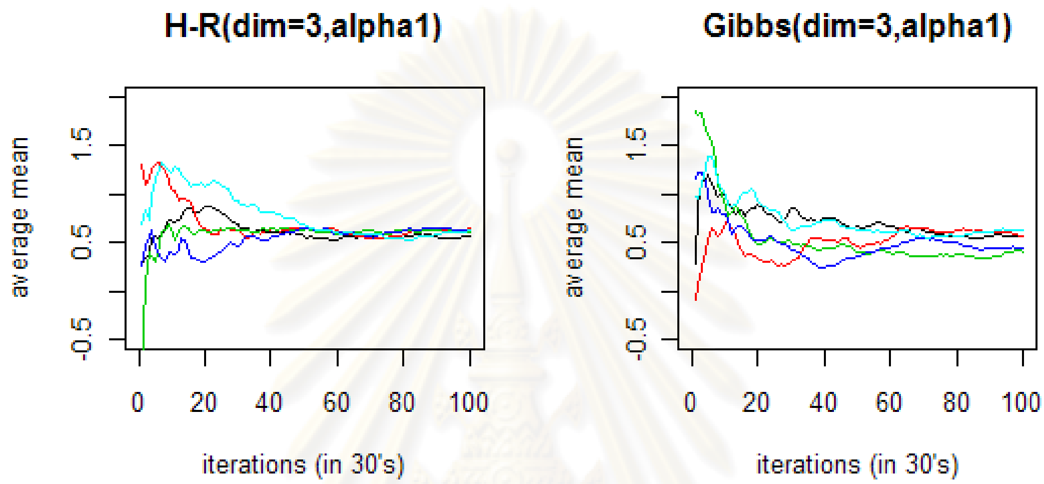


รูปที่ 4.25 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 2,000 รอบต่อลูกโซ่ ในกรณี 2 มิติ และ $\rho = 0.5$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.5 และรูปที่ 4.25 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 2,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.01 และค่า MPSRF

จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.02 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

<u>กรณี 3 มิติ</u>



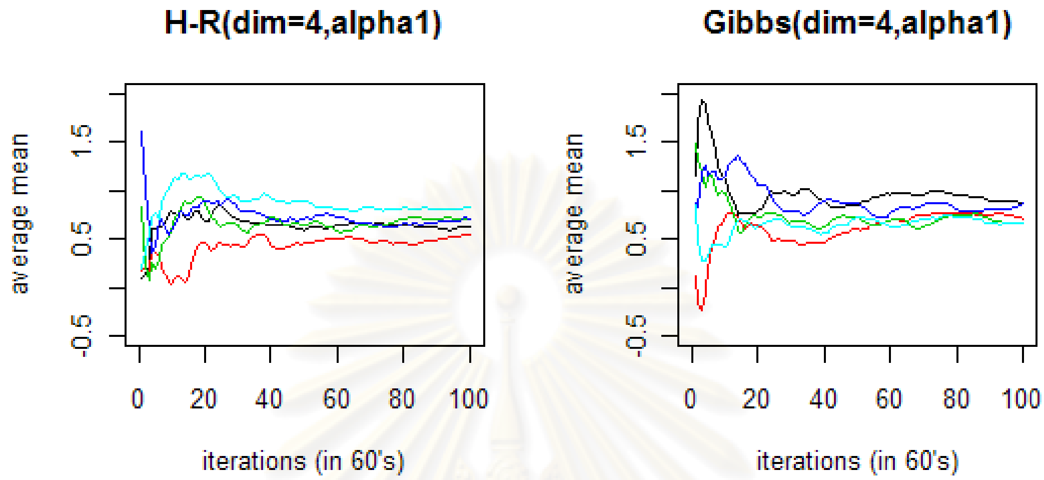
รูปที่ 4.26 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 6,000 รอบต่อลูกโซ่ ในกรณี 3 มิติ และ $\rho = 0.5$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.5 และรูปที่ 4.26 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 6,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.01 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.02 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

<u>กรณี 4 มิติ</u>

เมื่อพิจารณาจากตารางที่ 4.5 และรูปที่ 4.27 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 12,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.01 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.02 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่ม

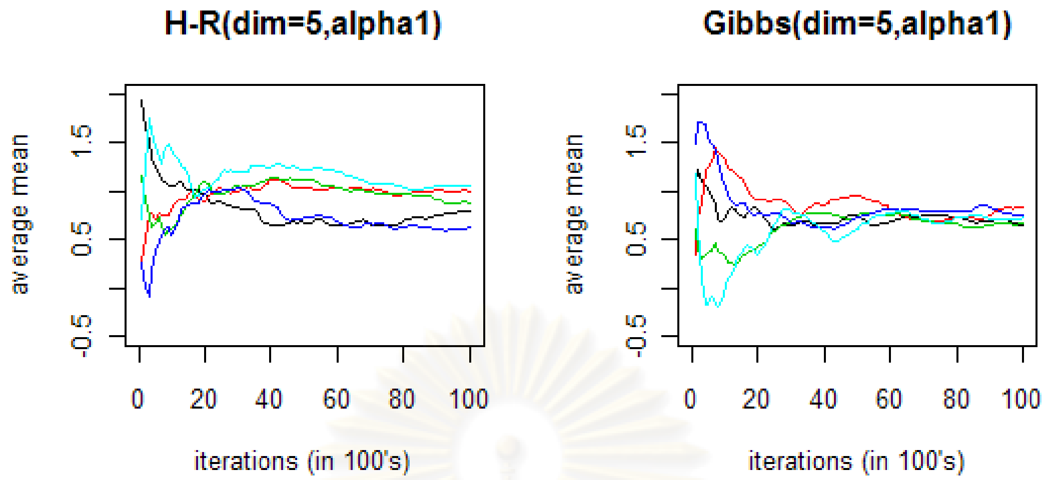
ตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์



รูปที่ 4.27 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 12,000 รอบต่อลูกโซ่ ในกรณี 4 มิติ และ $\rho = 0.5$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

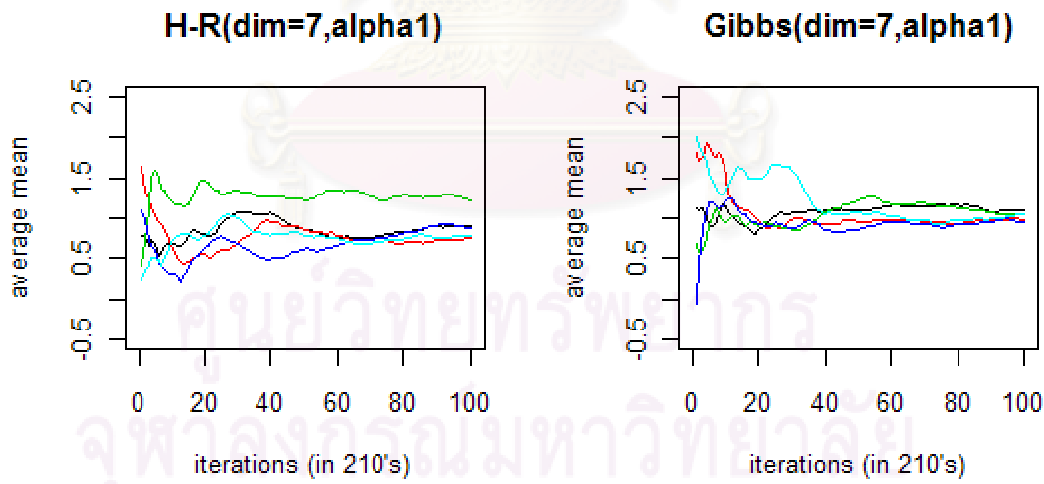
<u>กรณี 5 มิติ</u>

เมื่อพิจารณาจากตารางที่ 4.5 และรูปที่ 4.28 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 20,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.04 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.01 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน



รูปที่ 4.28 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 20,000 รอบต่อลูกโซ่ ในกรณี 5 มิติ และ $\rho = 0.5$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

<u>กรณี 7 มิติ</u>



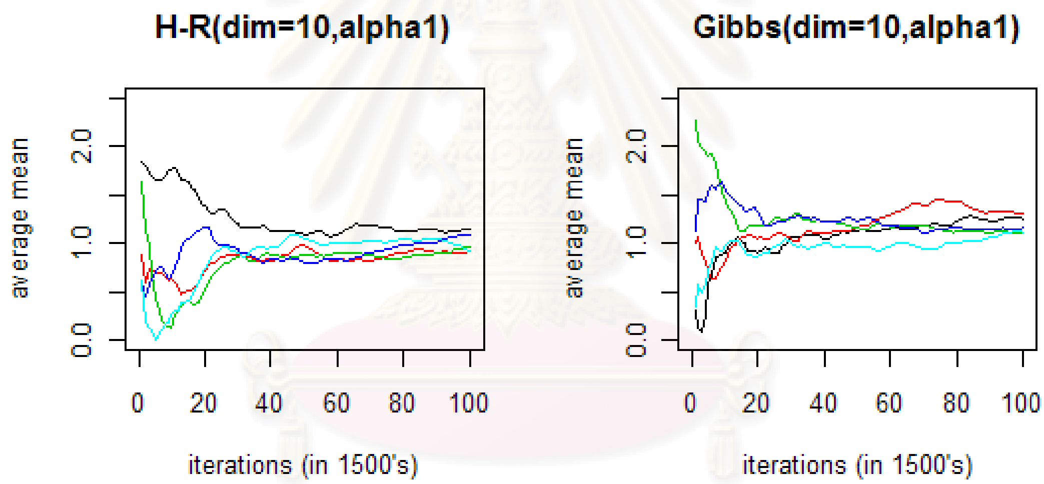
รูปที่ 4.29 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 42,000 รอบต่อลูกโซ่ ในกรณี 7 มิติ และ $\rho = 0.5$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.5 และรูปที่ 4.29 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 42,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.07 และค่า MPSRF

จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.01 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

<u>กรณี 10 มิติ</u>

เมื่อพิจารณาจากตารางที่ 4.5 และรูปที่ 4.30 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 300,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.07 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.03 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน



รูปที่ 4.30 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 300,000 รอบต่อลูกโซ่ ในกรณี 10 มิติ และ $\rho = 0.5$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

4.2.3 ผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรูกซ์-เกลแมน เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9

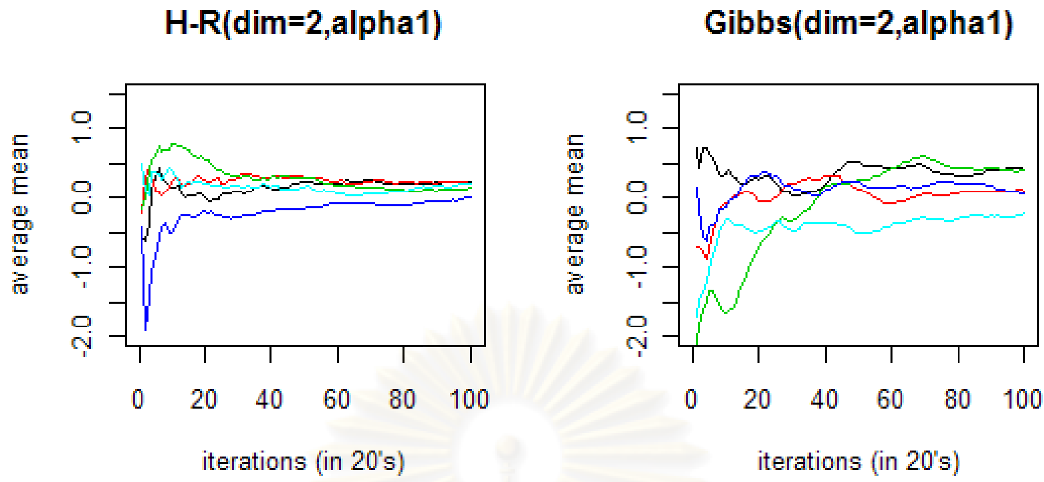
ตารางที่ 4.6 แสดงค่า MPSRF ของ บรูกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.9

k	n	MPSRF	
		Hit-and-Run	Gibbs
2	4,000	1.01	1.06
3	18,000	1.01	1.08
4	48,000	1.00	1.04
5	100,000	1.17	1.08
7	294,000	1.14	1.10
*10	300,000	*1.31	*1.48

* หมายเหตุ ค่า MPSRF ในกรณี 10 มิติ มีค่าสูง ซึ่งหมายความว่าทั้งการจำลองด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์ ยังไม่ลู่เข้า

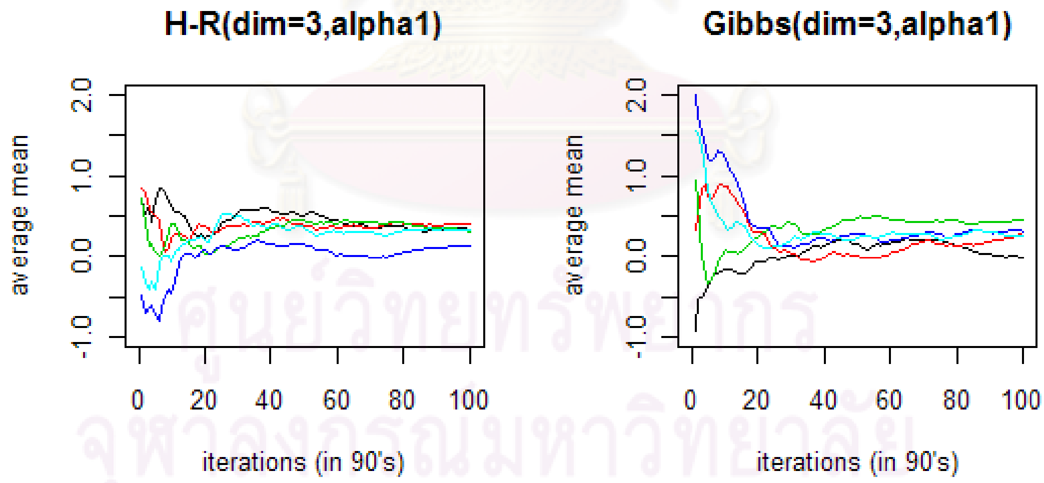
<u>กรณี 2 มิติ</u>

เมื่อพิจารณาจากตารางที่ 4.6 และรูปที่ 4.31 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 4,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.01 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.06 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์



รูปที่ 4.31 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 4,000 รอบต่อลูกโซ่ ในกรณี 2 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

<u>กรณี 3 มิติ</u>

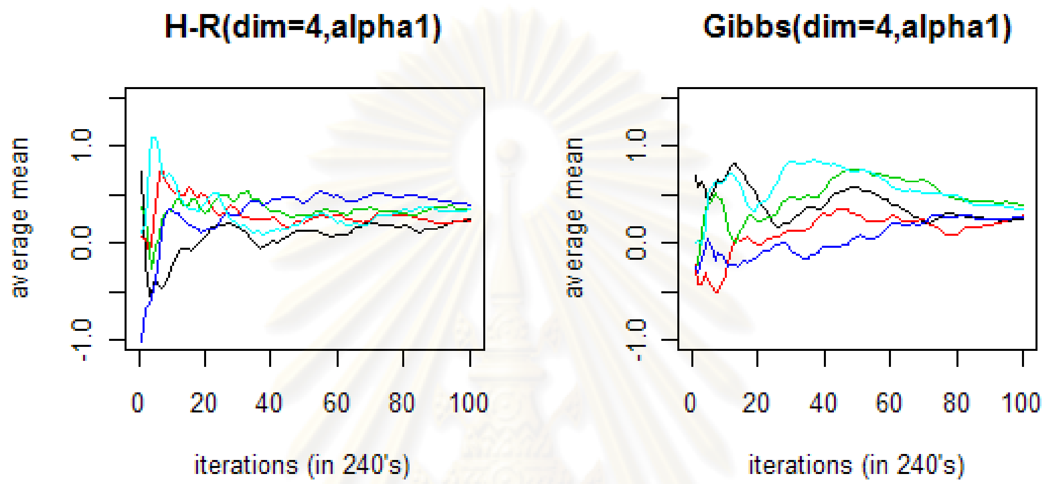


รูปที่ 4.32 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 18,000 รอบต่อลูกโซ่ ในกรณี 3 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.6 และรูปที่ 4.32 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 18,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.01 และค่า MPSRF

จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.08 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

<u>กรณี 4 มิติ</u>

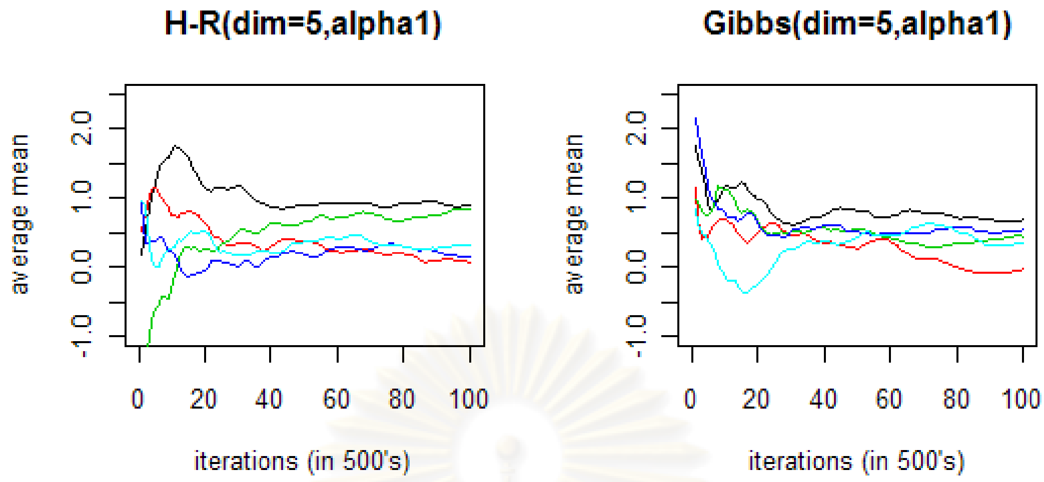


รูปที่ 4.33 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 48,000 รอบต่อลูกโซ่ ในกรณี 4 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.6 และรูปที่ 4.33 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 48,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.00 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.04 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่าน้อยกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้การสุ่มตัวอย่างแบบฮิตแอนด์รันมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์

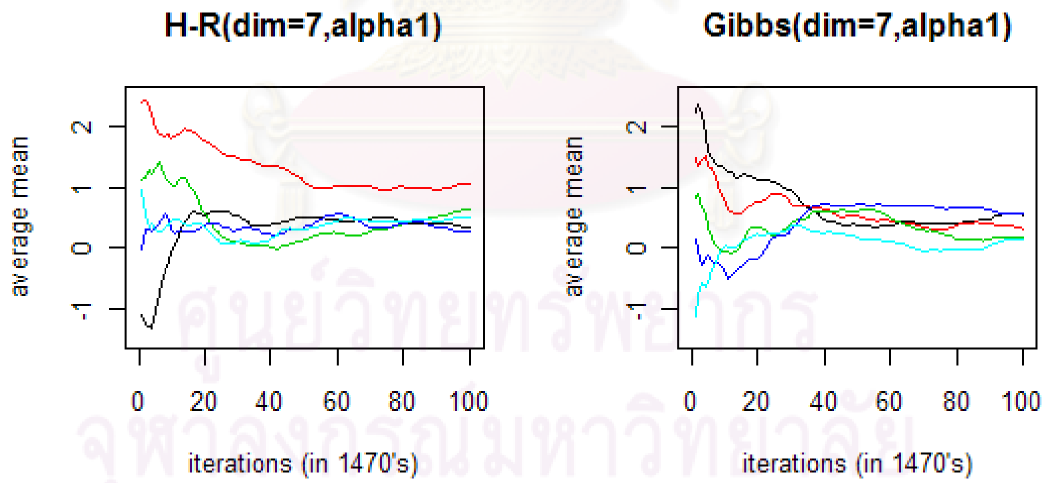
<u>กรณี 5 มิติ</u>

เมื่อพิจารณาจากตารางที่ 4.6 และรูปที่ 4.34 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 100,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.17 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.08 ดังนั้นทั้ง 2 วิธีลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน



รูปที่ 4.34 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 100,000 รอบต่อลูกโซ่ ในกรณี 5 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

<u>กรณี 7 มิติ</u>

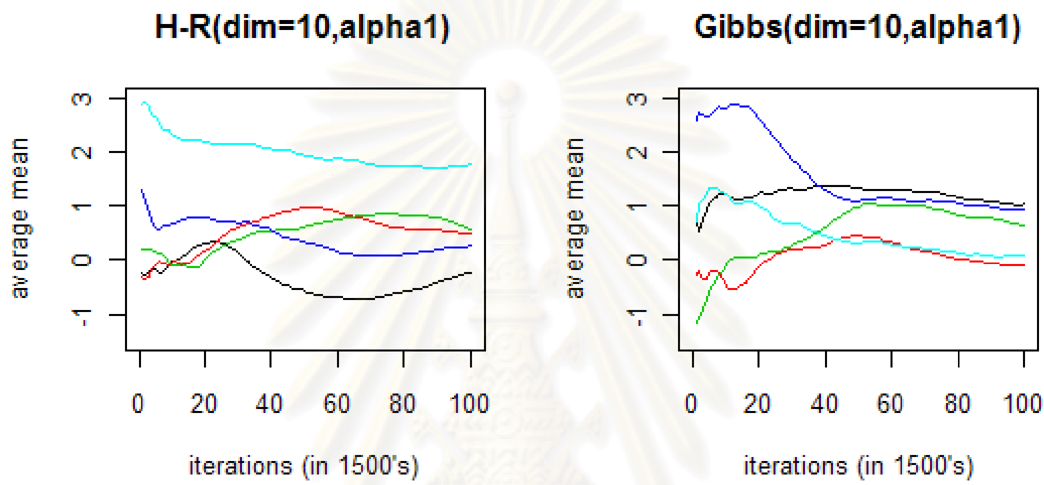


รูปที่ 4.35 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 294,000 รอบต่อลูกโซ่ ในกรณี 7 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.6 และรูปที่ 4.35 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 294,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.14 และค่า

MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.10 ซึ่งทั้ง 2 วิธีเริ่มที่จะลู่เข้าแล้ว และค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่ามากกว่าค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์ ดังนั้นในกรณีนี้ การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน แต่อย่างไรก็ตาม เนื่องจากค่าทั้ง 2 มีค่าค่อนข้างสูง การเปรียบเทียบในกรณีนี้อาจให้ผลที่ไม่ดีนัก

<u>กรณี 10 มิติ</u>



รูปที่ 4.36 แสดงค่าเฉลี่ยของค่าคาดหวังของตัวแปรตัวที่ 1 ของแต่ละลูกโซ่ เมื่อทำการจำลองข้อมูล 300,000 รอบต่อลูกโซ่ ในกรณี 10 มิติ และ $\rho = 0.9$ ด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

เมื่อพิจารณาจากตารางที่ 4.6 และรูปที่ 4.36 พบว่า เมื่อทำการจำลองข้อมูลตัวอย่าง 5 ลูกโซ่ ลูกโซ่ละ 300,000 รอบ ค่า MPSRF จากการสุ่มตัวอย่างแบบฮิตแอนด์รันมีค่า 1.31 และค่า MPSRF จากการสุ่มตัวอย่างแบบกิบส์มีค่า 1.48 ซึ่งมีค่าสูงทั้ง 2 วิธี ดังนั้นทั้ง 2 วิธียังไม่ลู่เข้า จึงไม่เหมาะที่จะนำค่าทั้งสองมาเปรียบเทียบกัน

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อทำการเปรียบเทียบประสิทธิภาพของการอนุมานสถิติเชิงเบส์สำหรับการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ ในกรณีที่การแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง เมื่อจำนวนตัวแปร(มิติ)เป็น 2, 3, 4, 5, 7 และ 10 ตัวแปร และเมทริกซ์ความแปรปรวนร่วม Σ มีค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) เป็น 0, 0.5 และ 0.9 โดยใช้เกณฑ์ในการวัดประสิทธิภาพ 2 เกณฑ์ คือ ค่าครึ่งช่วงความเชื่อมั่นที่คำนวณจากวิธีค่าเฉลี่ยกลุ่ม และ ค่า MPSRF ของ บรูกซ์-เกลแมน ซึ่งสามารถสรุปผลการวิจัยได้ดังนี้

5.1 สรุปผลการวิจัย

5.1.1 สรุปผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม

จากผลการเปรียบเทียบประสิทธิภาพ พบว่า

เมื่อ $\underset{\sim}{\alpha}$ มีจำนวนมิติต่ำๆ เช่น 2, 3 และ 4 มิติ การสุ่มตัวอย่างแบบฮิตแอนด์รันจะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ในเมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$)

เมื่อ $\underset{\sim}{\alpha}$ เริ่มมีจำนวนมิติสูงขึ้น เช่น 5, 7 และ 10 มิติ การสุ่มตัวอย่างแบบกิบส์จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ในเมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$) (ยกเว้นกรณี 7 และ 10 มิติ เมื่อ $\rho = 0.9$ ที่ไม่สามารถเปรียบเทียบกันได้)

5.1.2 สรุปผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรูกซ์-เกลแมน

จากผลการเปรียบเทียบประสิทธิภาพพบว่า

เมื่อ $\underset{\sim}{\alpha}$ มีจำนวนมิติต่ำๆ เช่น 2, 3 และ 4 มิติ การสุ่มตัวอย่างแบบฮิตแอนด์รันจะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ใน

เมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$)

เมื่อ $\underset{\sim}{\alpha}$ เริ่มมีจำนวนมิติสูงขึ้น เช่น 5, 7 และ 10 มิติ การสุ่มตัวอย่างแบบกิบส์จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ในเมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$) (ยกเว้นกรณี 10 มิติ เมื่อ $\rho = 0.9$ ที่ไม่สามารถเปรียบเทียบกันได้)

ตารางที่ 5.1 แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 และค่า MPSRF ของบรูกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

k	rho	half width		MPSRF	
		Hit-and-Run	Gibbs	Hit-and-Run	Gibbs
2	0	0.0295	0.0354	1.01	1.02
	0.5	0.0344	0.0541	1.01	1.02
	0.9	0.0413	0.1048	1.01	1.06
3	0	0.0398	0.0527	1.03	1.04
	0.5	0.0522	0.0639	1.01	1.02
	0.9	0.0556	0.1412	1.01	1.08
4	0	0.0394	0.0403	1.02	1.03
	0.5	0.0551	0.0808	1.01	1.02
	0.9	0.0629	0.1050	1.00	1.04

ตารางที่ 5.1 (ต่อ) แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 และค่า MPSRF ของบรูกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

k	rho	half width		MPSRF	
		Hit-and-Run	Gibbs	Hit-and-Run	Gibbs
5	0	0.0571	0.0383	1.03	1.01
	0.5	0.0887	0.0861	1.04	1.01
	0.9	0.1323	0.1144	1.17	1.08
7	0	0.0730	0.0497	1.09	1.08
	0.5	0.1314	0.1163	1.07	1.01
	0.9	*0.2099	*0.2816	1.14	1.10
10	0	0.1046	0.0510	1.09	1.02
	0.5	0.2130	0.1650	1.08	1.03
	0.9	*0.2974	*0.3265	*1.31	*1.48

* หมายเหตุ ทั้งการจำลองด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์ ยังไม่ลู่เข้า จึงไม่สามารถทำการเปรียบเทียบในกรณีนั้นได้

5.2 อภิปรายผลการวิจัย

จากผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยใช้เกณฑ์ในการเปรียบเทียบ 2 เกณฑ์ คือ วัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม และ วัดจากค่า MPSRF ของ บรูกซ์-เกลแมน พบว่าได้ผลการเปรียบเทียบในทำนองเดียวกันทั้ง 2 เกณฑ์ นั่นคือ ในกรณีที่ $\underset{\sim}{\alpha}$ มีจำนวนมิติต่ำๆ เช่น 2, 3 และ 4 มิติ การสุ่มตัวอย่างแบบฮิตแอนด์รันจะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ ซึ่งผลที่ได้นี้สอดคล้องกับผลจากงานวิจัยของ Chen M.-H. and Schmeiser B.W. (1993) ซึ่งได้ทำการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ในกรณี

ที่เป็นการแจกแจงแบบปกติของ 2 ตัวแปร(ไม่ตัดหาง) ส่วนในกรณีที่ $\underset{\sim}{\alpha}$ มีจำนวนมิติเพิ่มสูงขึ้น เช่น 5, 7 และ 10 มิติ จากผลที่ได้กลับพบว่า การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน ซึ่งสาเหตุที่เป็นเช่นนี้อาจเป็นเพราะ ในกรณีที่เราศึกษานั้น ค่าพารามิเตอร์ที่เป็นค่าคาดหวัง $\underset{\sim}{\alpha} = \left(\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(k)}\right)$ มีการเรียงอันดับอย่างสมบูรณ์ พารามิเตอร์แต่ละตัวจะถูกจำกัดค่าด้วยพารามิเตอร์ตัวที่อยู่ติดกัน สำหรับการสุ่มตัวอย่างแบบฮิตแอนด์รัน ในการจำลองแต่ละรอบจะทำการสุ่มเลือกทิศทาง ซึ่งค่าของตัวอย่างสามารถเปลี่ยนพร้อมกันได้ทุกมิติ แต่ว่าในแต่ละมิติค่าของตัวอย่างจะเปลี่ยนได้ในระยะทางที่จำกัด ยิ่งจำนวนมิติมาก ข้อจำกัดก็จะยิ่งมาก การเปลี่ยนแปลงค่าในการจำลองแต่ละรอบ ถึงแม้จะเปลี่ยนพร้อมๆ กันได้ทุกมิติก็ตาม แต่ในแต่ละมิติจะเปลี่ยนได้เล็กน้อยเท่านั้น เมื่อเทียบกับการสุ่มตัวอย่างแบบกิบส์ ซึ่งการจำลองแต่ละรอบ ค่าของตัวอย่างจะเปลี่ยนได้ครั้งละหนึ่งมิติโดยที่ค่าในมิติอื่นคงที่ แต่ค่าที่เปลี่ยนในมิตินั้นอาจจะเปลี่ยนได้ในระยะทางที่มากกว่า ดังนั้นเมื่อจำนวนมิติมีค่าสูงๆ จึงทำให้การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน

5.3 ข้อเสนอแนะ

1. เมื่อต้องการประมาณค่าคาดหวังของผลตอบแทน ในกรณีที่พารามิเตอร์ที่เป็นค่าคาดหวัง มีการเรียงอันดับอย่างสมบูรณ์ และเวกเตอร์ค่าคาดหวังมีจำนวนมิติต่ำกว่า 5 มิติ ควรเลือกใช้วิธีการสุ่มตัวอย่างแบบฮิตแอนด์รัน แต่ถ้าเวกเตอร์ค่าคาดหวังมีจำนวนมิติตั้งแต่ 5 มิติขึ้นไป ควรเลือกใช้วิธีการสุ่มตัวอย่างแบบกิบส์

2. ในการวิจัยครั้งนี้ ได้ทำการกำหนดจำนวนรอบในการจำลองข้อมูลที่แน่นอน แล้วจึงทำการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์ โดยเปรียบเทียบที่มิติเดียวกัน และจำนวนรอบที่จำลองข้อมูลเท่ากัน ซึ่งในบางกรณี การจำลองข้อมูลด้วยวิธีทั้ง 2 วิธี อาจจะยังไม่ลู่เข้าดีนัก จึงทำให้การเปรียบเทียบประสิทธิภาพในกรณีนั้นอาจเกิดความผิดพลาดขึ้นได้ ซึ่งการแก้ปัญหาดังกล่าวอาจทำได้โดย เพิ่มจำนวนรอบในการจำลองข้อมูลในกรณีนั้น หรืออาจทำการกำหนดเกณฑ์ขั้นต่ำที่ยอมรับได้ว่าลู่เข้า แล้วเปรียบเทียบประสิทธิภาพจากจำนวนรอบในการจำลองข้อมูลจนเข้าเกณฑ์ขั้นต่ำนั้น

รายการอ้างอิง

ภาษาไทย

เสกสรร เกียรติสุไพบูลย์. เอกสารประกอบการสอนวิชาการจำลองแบบเชิงสถิติ. กรุงเทพมหานคร: จุฬาลงกรณ์มหาวิทยาลัย, 2552.

ภาษาอังกฤษ

Anant Chiarawongse, Seksan Kiatsupaibul, Sunti Tirapat and Benjamin Van Roy. Portfolio Selection with Qualitative Input. (2009).(working paper)

Bélisle C.J.P.,Romeijn H.E. and Smith R.L. Hit-and-Run Algorithms for Generating Multivariate Distributions. Mathematics of Operations Research 18 (1993) : 255-266.

Brooks S.P. and Gelman A.. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7, 4 (1998) : 434-455.

Chen M.-H. and Schmeiser B.W. Performance of the Gibbs, Hit-and-Run, and Metropolis Samplers. Journal of Computational and Graphical Statistics 2, 3 (1993) : 251-272.

Flegal J.M.,Haran M. and Jones G.L. Markov Chain Monte Carlo: Can We Trust the Third Significant Figure?. Statistical Science 23, 2 (2008) : 250-260.

Gelman A. and Rubin D.B. Inference from iterative simulation using multiple sequences. Statistical Science 7, 4 (1992) : 457-511.

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

Lovász L. Hit-and-Run Mixes Fast. Mathematical Programming 86 (1998) : 443-461.

Romeijn H.E. and Smith R.L. Simulated annealing for constrained global optimization. Journal of Global Optimization 5 (1994) : 101-126.

Smith R.L.. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. Operations Research 32 (1984) : 1296-1308.

ภาคผนวก

ตัวอย่างการใช้โปรแกรม R ในการดำเนินงานวิจัย

การจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และคำนวณค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม กรณี 2 มิติ และค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0

```
##สร้างฟังก์ชัน hitandrun
hitandrun<-function(x0,dim,sigmainv){
        d0<-rnorm(dim,0,1)
        d<-d0/sqrt(sum(d0*d0))
        mutrun<-(d%*%sigmainv%*%matrix(-x0,nrow=dim))
                /(d%*%sigmainv%*%matrix(d,nrow=dim))
        vartrun<-1/(d%*%sigmainv%*%matrix(d,nrow=dim))
        sdtrun<-sqrt(vartrun)
        lambdabound<-(x0[1:dim-1]-x0[2:dim])/(d[2:dim]-d[1:dim-1])
        pvec<-c(lambdabound[lambdabound >0],Inf)
        nvec<-c(lambdabound[lambdabound <0],-Inf)
        infval<-max(nvec)
        supval<-min(pvec)
        lamda<-qnorm((pnorm(supval,mean=mutrun,sd=sdtrun)
               -pnorm(infval,mean=mutrun,sd=sdtrun))*runif(1)
               +pnorm(infval,mean=mutrun,sd=sdtrun),mean=mutrun,sd=sdtrun)
        x0<-x0+lamda*d
        x0
}

maxround<-10000;          ##กำหนดจำนวนรอบในการจำลองข้อมูล 10,000 รอบ
burnin<-0.1*maxround      ##burn in 10 %
x<-c(0.41,-0.10)
dim<-length(x);
sigma<-matrix(0,dim,dim);diag(sigma)<-1;      ##กำหนดค่าสัมประสิทธิ์สหสัมพันธ์เป็น0
sigmainv<-solve(sigma)
```

```
xarray<-x;
for (i in 1:maxround){
        x<-hitandrun(x,dim,sigmainv);
        xarray<-rbind(xarray,x)
}
mub<-c()
##กำหนดจำนวนbatches=30
for(i in 1:30){
        mub<-rbind(mub, colMeans(xarray[((i-1)*((0.9*maxround)/30)
              +burnin+1):(i*((0.9*maxround)/30)+burnin),]))
}
##คำนวณค่าครึ่งช่วงความเชื่อมั่น
error<-qt(0.975,29)*sd(mub)/sqrt(30)
mu<-colMeans(mub)
```

การจำลองข้อมูลด้วยการสุ่มตัวอย่างแบบกิบส์ และคำนวณค่า MPSRF กรณี 2 มิติ และค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0

```
#สร้างฟังก์ชัน gibbs
gibbs<-function(x0,dim,sigmainv){
        nchoose<-floor(dim*runif(1))+1
        mutrun<--sum((x0[-nchoose])*(sigmainv[nchoose,][-nchoose]))
                /sigmainv[nchoose,nchoose]
        vartrun<-1/(sigmainv[nchoose,nchoose])
        sdtrun<-sqrt(vartrun)
        infval<-ifelse(nchoose==dim,-Inf,x0[nchoose+1])
        supval<-ifelse(nchoose==1,Inf,x0[nchoose-1])
        x[nchoose]<-qnorm((pnorm(supval,mean=mutrun,sd=sdtrun)
                   -pnorm(infval,mean=mutrun,sd=sdtrun))*runif(1)
                   +pnorm(infval,mean=mutrun,sd=sdtrun),mean=mutrun,sd=sdtrun)
        x<-c(x0[-(nchoose:dim)],x[nchoose],x0[-(1:nchoose)])
```

```
        x0<-x
}

xseed<-rnorm(2,0,1)
x<-xseed[c(order(-xseed))];
dim<-length(x);
sigma<-matrix(0,dim,dim);diag(sigma)<-1;    ##กำหนดค่าสัมประสิทธิ์สหสัมพันธ์เป็น0
sigmainv<-solve(sigma)
xarray<-x;

gb<-function(maxround,burnin){
for (i in 1:maxround){
        x<-gibbs(x,dim,sigmainv);
        xarray<-rbind(xarray,x)
}
xuse<-xarray[(burnin+1):maxround,]
}
##สร้างลูกโซ่ 5 ลูกโซ่ แต่ละลูกโซ่ burn in 50%
gb.draws1<-mcmc(gb(1000,500))
gb.draws2<-mcmc(gb(1000,500))
gb.draws3<-mcmc(gb(1000,500))
gb.draws4<-mcmc(gb(1000,500))
gb.draws5<-mcmc(gb(1000,500))

gb.list<-mcmc.list(list(gb.draws1,gb.draws2,gb.draws3,gb.draws4,gb.draws5))
## คำนวณค่า MPSRF จากฟังก์ชัน gelman.diag() ใน package coda
gelman.diag(gb.list)
```

ประวัติผู้เขียนวิทยานิพนธ์

นายพงศ์ศักดิ์ สถิตรุ่งพรชัย เกิดวันที่ 13 สิงหาคม 2524 สำเร็จการศึกษาระดับปริญญาบัณฑิต หลักสูตรวิทยาศาสตรบัณฑิต (วท.บ.) จากภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2548 และเข้าศึกษาต่อในระดับปริญญามหาบัณฑิต หลักสูตรสถิติศาสตรมหาบัณฑิต ที่คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2551