ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ
245845

การเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างด้วยวิธีชิคแมนที่รันแลสการสุ่มตัวอย่างด้วยวิธีกิบส์สำหรับการวิเคราะห์ข้อถดถอยนิพนธ์

นางสาวไรยา พลเสน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

b 00251806

ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ
245845

การเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างด้วยวิธีฮิตแอนด์รันและการสุ่มตัวอย่างด้วยวิธีกิบส์สำหรับการวิเคราะห์ปัจจัยเชิงเบส์

นางสาวไรยา พลเสน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

5 1 8 1 9 0 0 6 2 6

A COMPARISON ON THE EFFICIENCY OF HIT-AND-RUN SAMPLER AND GIBBS SAMPLER FOR BAYESIAN FACTOR ANALYSIS

Miss Raiya Polsen

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

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างด้วยวิธีฮิตแอนด์รันและการสุ่มตัวอย่างด้วยวิธีกิบส์สำหรับการวิเคราะห์ปัจจัยเชิงเบส์
โดย	นางสาวไรยา พลเสน
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. อัครินทร์ ไพบูลย์พานิช)

ไรยา พลเสน : การเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างด้วยวิธีฮิตแอนด์รันและการสุ่มตัวอย่างด้วยวิธีกิบส์สำหรับการวิเคราะห์ปัจจัยเชิงเบส์. (A COMPARISON ON THE EFFICIENCY OF HIT-AND-RUN SAMPLER AND GIBBS SAMPLER FOR BAYESIAN FACTOR ANALYSIS) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร. เสกสรร เกียรติสุไพบูลย์, 66 หน้า.

245845

การศึกษานี้มีวัตถุประสงค์เพื่อเปรียบเทียบการสุ่มตัวอย่างด้วยวิธีฮิตแอนด์รันและการสุ่มตัวอย่างด้วยวิธีกิบส์สำหรับทำการวิเคราะห์ปัจจัยเมื่อมีจำนวนปัจจัย 1 ปัจจัย เพื่อใช้ในการจัดอันดับหน่วยทดลองจากตัวชี้วัดมากกว่า 1 ตัวซึ่งทั้ง 2 วิธีจัดอยู่ในกลุ่มวิธีลูกโซ่มาร์คอฟมอนติคาร์โล (MCMC) การสุ่มตัวอย่างทั้ง 2 วิธีจะถูกทดสอบในการจัดอันดับประเทศตามความเสี่ยงทางการเมือง-เศรษฐกิจด้วยข้อมูลตัวชี้วัดจาก MCMCpack R Package โดย Martin และ Quinn 2004 โดยใช้คะแนนของปัจจัยจากตัวแบบการวิเคราะห์ปัจจัยมาจัดอันดับ ในแต่ละวิธี MCMC จะประมาณคะแนนเฉลี่ยภายใต้การแจกแจงความน่าจะเป็นภายหลังและค่าคลาดเคลื่อนมาตรฐาน (SE) แล้วใช้ค่าคลาดเคลื่อนมาตรฐานเป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพความเร็วในการลู่เข้าของตัวประมาณจากการสุ่มตัวอย่างทั้ง 2 วิธี

จากการทดสอบประสิทธิภาพด้วยข้อมูลความเสี่ยงทางการเมือง-เศรษฐกิจจาก MCMCpack R Package พบว่าการสุ่มตัวอย่างด้วยวิธีฮิตแอนด์รันมีประสิทธิภาพด้อยกว่าวิธีกิบส์ โดยส่วนใหญ่ของค่าคลาดเคลื่อนมาตรฐานของคะแนนเฉลี่ยของแต่ละประเทศที่คำนวณได้จากการสุ่มตัวอย่างด้วยวิธีฮิตแอนด์รันมีค่ามากกว่าค่าคลาดเคลื่อนมาตรฐานที่คำนวณได้จากวิธีกิบส์ สังเกตได้ว่าการวิเคราะห์ปัจจัยเชิงเบส์ในกรณีที่มีพารามิเตอร์จำนวนมาก และมีขอบเขตของพารามิเตอร์ที่ซับซ้อน การเคลื่อนที่ในแต่ละรอบของวิธีฮิตแอนด์รันพารามิเตอร์มีมิติสูงจะเคลื่อนที่ได้อย่างจำกัด จึงทำให้การสุ่มตัวอย่างด้วยวิธีฮิตแอนด์รันมีประสิทธิภาพในเชิงการลู่เข้าของตัวประมาณด้อยกว่าการสุ่มตัวอย่างด้วยวิธีกิบส์

ภาควิชา สถิติ ลายมือชื่อนิสิต ไรยา พลเสน

สาขาวิชา สถิติ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก เสกสรร เกียรติสุไพบูลย์

ปีการศึกษา 2553

5182900626 : MAJOR STATISTICS

KEYWORDS : BAYESIAN FACTOR ANALYSIS / MARKOV CHAIN MONTE CARLO / HIT-AND-RUN METHOD

RAIYA POLSEN: A COMPARISON ON THE EFFICIENCY OF HIT-AND-RUN SAMPLER AND GIBBS SAMPLER FOR BAYESIAN FACTOR ANALYSIS.

THESIS ADVISOR: ASST.PROF.SEKSAN KIATSUPAIBUL, Ph.D., 66 pp.

245845

The purpose of this study is to compare the efficiency of Hit-and-run sampler and Gibb sampler, the two Markov chain Monte Carlo samplers, for single factor analysis with an application to ranking units by their multiple indicators. The two samplers are tested against each other on a cross-national political-economic risk ranking according to the indicator data provided in MCMCpack R package by Martin and Quinn 2004, where the factor scores of the factor analysis model are used as the ranking measure. Each of the MCMC samplers is employed to estimates the expectations of the factor scores under its posterior distribution, and the standard errors of the factor score estimates are adopted as the measurement of the efficiency of the samplers.

From the experiments with the test data set, Hit-and-run sampler is inferior to Gibbs sampler. The majority of the standard errors of the factor score estimates obtained from Hit-and-run sampler are higher than those obtained from Gibb sampler. From observations, Hit-and-run has difficulties in its transition due to the fact that the Bayesian factor analysis model imposes a large number of constraints in each step of Hit-and-run in high dimensional parameter space.

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

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

Academic Year : 2010

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความช่วยเหลือ คำแนะนำและข้อเสนอแนะ รวมถึงการเอาใจใส่อย่างใกล้ชิดจากผู้ช่วยศาสตราจารย์ ดร.เสกสรร เกียรติสุไพบูลย์ ผู้เขียนขอน้อมกราบขอบพระคุณต่อท่านอาจารย์เป็นอย่างสูง ตลอดทั้งอาจารย์ทุกท่านที่ได้สอนผู้เขียนมา ณ โอกาสนี้ด้วย

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. ธีระพร วีระถาวร ประธานกรรมการ รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา และอาจารย์ ดร. อัครินทร์ ไพบูลย์พานิช กรรมการสอบวิทยานิพนธ์ ที่กรุณาตรวจสอบและแก้ไขให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์มากยิ่งขึ้น รวมทั้งขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ได้ประสิทธิประสาทวิชาความรู้ให้แก่ผู้วิจัย

ขอขอบคุณพี่สุมณฑา ที่มีส่วนสำคัญที่ทำให้เกิดวิทยานิพนธ์ฉบับนี้ ขอขอบคุณรุ่นพี่ และเพื่อน ๆ ทุกคนที่ให้ความร่วมมือและเป็นกำลังใจแก่ผู้เขียน

สุดท้ายนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดา มารดา และครอบครัวที่ได้ให้การสนับสนุนด้านการศึกษาและคอยเป็นกำลังให้ผู้วิจัยมาโดยตลอด

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ง

บทคัดย่อภาษาอังกฤษ..... จ

กิตติกรรมประกาศ..... ฉ

สารบัญ..... ช

สารบัญตาราง..... ฌ

สารบัญภาพ..... ญ

บทที่ 1 บทนำ..... 1

ความเป็นมาและความสำคัญของปัญหา..... 1

วัตถุประสงค์ของการวิจัย..... 2

ขอบเขตของการวิจัย..... 2

คำจำกัดความที่ใช้ในการวิจัย..... 2

ประโยชน์ที่คาดว่าจะได้รับ..... 3

วิธีดำเนินการวิจัย..... 3

ลำดับขั้นตอนในการเสนอผลการวิจัย..... 3

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... 5

ตัวแบบการวิเคราะห์ปัจจัยสำหรับข้อมูลแบบผสมที่มีทั้งตัวแปรแบบต่อเนื่องและตัวแปรเชิงอันดับ..... 5

ลูกโซ่มาร์คอฟ..... 7

ทฤษฎีลูกโซ่มาร์คอฟสำหรับสเปซทั่วไป..... 7

ทฤษฎีบทการลู่เข้าของลูกโซ่มาร์คอฟในกรณี total variation..... 8

การวิเคราะห์สถานะเสถียรภาพของการจำลอง..... 10

เทคนิคลูกโซ่มาร์คอฟมอนติคาร์โล..... 10

การสุ่มตัวอย่างแบบเมโทรโปลิส-เฮสติงส์..... 11

การสุ่มตัวอย่างแบบกิบส์..... 12

การสุ่มตัวอย่างแบบฮิตแอนด์รัน..... 12

วิธีค่าเฉลี่ยกลุ่ม..... 13

เอกสารและงานวิจัยที่เกี่ยวข้อง..... 13

บทที่ 3 วิธีดำเนินการวิจัย..... 15

บทที่ หน้า

ขั้นตอนดำเนินการด้วยวิธีการสุ่มตัวอย่างแบบกิบส์..... 17
ขั้นตอนดำเนินการด้วยวิธีการสุ่มตัวอย่างแบบฮิตแอนด์รัน..... 20
บทที่ 4 ผลการวิเคราะห์ข้อมูล..... 23
บทที่ 5 สรุปผลการวิจัย และข้อเสนอแนะ..... 29
สรุปผลการวิจัย..... 29
อภิปรายผลการวิจัย..... 29
ข้อเสนอแนะ..... 31
รายการอ้างอิง..... 32
ภาคผนวก..... 33
ภาคผนวก ก โปรแกรมสำหรับงานวิจัย..... 34
ภาคผนวก ข ผลการจัดอันดับประเทศตามความเสี่ยงทางการเมือง-เศรษฐกิจ
ค่าประมาณและส่วนเบี่ยงเบนมาตรฐานของ λ และ γ..... 52
ประวัติผู้เขียน..... 66

สารบัญตาราง

ตารางที่		หน้า
4.1	แสดงส่วนเบี่ยงเบนมาตรฐาน (SE) ของคะแนนเฉลี่ยจากการสุ่มตัวอย่างแบบฮิตแอนด์รันและการสุ่มตัวอย่างแบบกิบส์	23

สารบัญภาพ

ภาพที่		หน้า
3.1	แสดงขั้นตอนการสุ่มตัวอย่างแบบกิบส์	19
3.2	แสดงขั้นตอนการสุ่มตัวอย่างแบบฮิตแอนด์รัน	21
4.1	คะแนนเฉลี่ยสะสมของประเทศแคนาดา	25
4.2	คะแนนเฉลี่ยสะสมของประเทศสวิสเซอร์แลนด์	25
4.3	คะแนนเฉลี่ยสะสมของประเทศสิงคโปร์	26
4.4	คะแนนเฉลี่ยสะสมของประเทศไทย	26
4.5	คะแนนเฉลี่ยสะสมของประเทศบังคลาเทศ	27
4.6	คะแนนเฉลี่ยสะสมของประเทศโบลิเวีย	27
5.1	แสดงค่าเฉลี่ยสะสมของ Lambda1 prsexp2	30
5.2	แสดงค่าเฉลี่ยสะสมของ Lambda2 prsexp2	30
5.3	แสดงค่าเฉลี่ยสะสมของ Gamma2 prsexp2	31