

การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส  
เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา

นางสาววรรณ จันทรโอกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาสถิติ ภาควิชาสถิติ  
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย  
ปีการศึกษา 2549  
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A COMPARISON OF BAYESIAN SELECTION METHODS FOR  
BEST REGRESSION EQUATION WITH CONJUGATE  
GAMMA PRIOR DISTRIBUTION

Miss Watsamon Janokul

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 2006

Copyright of Chulalongkorn University

**491821**

หัวข้อวิทยานิพนธ์

การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส  
เมื่อใช้การแจกแจงก่อนแบบกลุ่มสับยุคแกมมา

โดย

นางสาววรรษมน จันทรโถกุล

สาขาวิชา

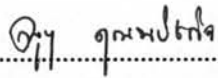
สถิติ

อาจารย์ที่ปรึกษา

รองศาสตราจารย์ ดร.ธีระพร วีระถาวร

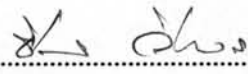
---

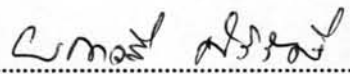
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัย  
ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

  
.....คณบดีคณะพาณิชยศาสตร์และการบัญชี  
(รองศาสตราจารย์ ดร.คณูชา คุณพนิชกิจ)

คณะกรรมการสอบวิทยานิพนธ์

  
.....ประธานกรรมการ  
(รองศาสตราจารย์ ดร.สุพถ คุรงค์วัฒนา)

  
.....อาจารย์ที่ปรึกษา  
(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)

  
.....กรรมการ  
(รองศาสตราจารย์ ผกาวดี ศิริรัมย์)

วรรณมน จันทร์โอกุล : การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา (A COMPARISON OF BAYESIAN SELECTION METHODS FOR BEST REGRESSION EQUATION WITH CONJUGATE GAMMA PRIOR DISTRIBUTION) อ.ที่ปรึกษา : รศ. ดร. ธีระพร วีระถาวร, 126 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา โดยจะเปรียบเทียบวิธีการคัดเลือกสมการถดถอย 3 วิธี ได้แก่ วิธีการเฉลี่ยตัวของเบส์โดยการหาค่าประกะอบของตัวแบบด้วยเทคนิคอนติคาร์โลเมื่อใช้ลูกโซ่มาร์คอฟ ( $BMA_{MC^3}$ ) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (OPM) และวิธีการถดถอยแบบขั้นบันได (SR) เกณฑ์ที่ใช้ในการตัดสินใจคือเกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) โดยกำหนดจำนวนตัวแปรอิสระที่ศึกษา คือ 3 5 10 และ 15 ขนาดตัวอย่างที่ศึกษา คือ 15 30 50 และ 100 ค่าส่วนเบี่ยงเบนมาตรฐานของค่าความคลาดเคลื่อนที่ศึกษา คือ 0.25 0.50 และ 2.50 ค่าคงที่สำหรับวิธี  $BMA_{MC^3}$  และวิธี OPM ที่ศึกษามี 4 ระดับ คือ (1,5) (1,10) (10,100) และ (10,500) ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคอนติคาร์โลกระทำซ้ำ 500 รอบในแต่ละสถานการณ์ ซึ่งผลการวิจัยได้ข้อสรุปดังนี้

1. ค่า AMSE ของทั้ง 3 วิธีเรียงลำดับจากน้อยไปหามาก ได้แก่ วิธี  $BMA_{MC^3}$  OPM และ SR ตามลำดับในทุกสถานการณ์ ซึ่งวิธี  $BMA_{MC^3}$  จะมีค่า AMSE ต่ำกว่าวิธี OPM และวิธี SR จะมีค่า AMSE สูงกว่าวิธี  $BMA_{MC^3}$  และวิธี OPM ในทุกสถานการณ์
2. ปัจจัยที่มีผลต่อค่า AMSE ของทุกวิธี คือจำนวนตัวแปรอิสระ ขนาดตัวอย่าง และค่าส่วนเบี่ยงเบนมาตรฐานของค่าความคลาดเคลื่อน โดยที่ค่า AMSE จะแปรผันตามจำนวนตัวแปรอิสระ และค่าส่วนเบี่ยงเบนมาตรฐานของค่าความคลาดเคลื่อน แต่จะแปรผกผันกับขนาดตัวอย่าง
3. ปัจจัยที่มีผลต่อค่า AMSE ของ 2 วิธี คือค่าคงที่สำหรับวิธี  $BMA_{MC^3}$  และวิธี OPM จะส่งผลต่อวิธี  $BMA_{MC^3}$  และวิธี OPM ซึ่งเป็นวิธีการภายใต้แนวทางของเบส์เท่านั้น โดยค่า AMSE จะแปรผันตามค่าคงที่สำหรับวิธี  $BMA_{MC^3}$  และวิธี OPM

ภาควิชา.....สถิติ.....      ลายมือชื่อนิสิต..... วรรณมน จันทร์โอกุล .....

สาขาวิชา.....สถิติ.....      ลายมือชื่ออาจารย์ที่ปรึกษา..... *[Signature]* .....

ปีการศึกษา.....2549.....

## 4782359626 : MAJOR STATISTICS

KEYWORD : BAYESIAN MODEL AVERAGING USING MARKOV CHAIN MONTE CARLO MODEL COMPOSITION / OPTIMAL PREDICTIVE MODEL SELECTION

WATSAMON JANOKUL : A COMPARISON OF BAYESIAN SELECTION METHODS FOR BEST REGRESSION EQUATION WITH CONJUGATE GAMMA PRIOR DISTRIBUTION. THESIS ADVISOR : ASSOC. PROF.

THEERAPRON VERATHAWORN, Ph.D., 126 pp.

The objective of this research is to compare the Bayesian selection methods for best regression equation with conjugate gamma prior distribution. The three methods in this comparison are Bayesian Model Averaging method using Markov Chain Monte Carlo model composition ( $BMA_{MC^3}$ ), Optimal Predictive Model Selection (OPM) and Stepwise Regression method (SR). The criterion of comparison is average of mean square error (AMSE) to compare the efficiency of these three methods. In this study, the numbers of independent variables in regression model are 3 5 10 and 15. The size of the samples are 15 30 50 and 100. The standard deviation of random errors are 0.25 0.50 and 2.50. The constant of Bayesian approach for  $BMA_{MC^3}$  and OPM are (1,5) (1,10) (10,100) and (10,500). All data for this research are generated through the Monte Carlo simulation technique and repeating 500 times for each case. The result of this study can be summarized as follow:

1. The AMSE from the three methods rank from minimum to maximum are  $BMA_{MC^3}$ , OPM and SR for all cases. The  $BMA_{MC^3}$  has AMSE lower than OPM but SR has AMSE higher than  $BMA_{MC^3}$  and OPM for all cases.
2. The factors that affect AMSE of all methods are the numbers of independent variables, sample size and the standard deviation of random errors. The AMSE of all methods directly vary to the numbers of independent variables and the standard deviation of random errors but the AMSE of all methods inversely vary to sample size.
3. The constant of Bayesian approach affected AMSE of both  $BMA_{MC^3}$  and OPM. The AMSE of the two methods directly vary to the constant of Bayesian approach.

Department.....Statistics.....  
 Field of study.....Statistics.....  
 Academic year.....2006.....

Student's signature.....*Watsamon Janokul*.....  
 Advisor's signature.....*Theerapron Verathaworn*.....

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความกรุณาและความช่วยเหลืออย่างดียิ่งของ รองศาสตราจารย์ ดร.ธีระพร วีระถาวร อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ช่วยให้คำแนะนำ ปรึกษาตลอดจนแก้ไขข้อบกพร่องต่างๆ ด้วยดีเสมอมา จนกระทั่งวิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.สุพล คุรงค์วัฒนา ประธานกรรมการ สอบวิทยานิพนธ์และรองศาสตราจารย์ผกาวดี ศิริรังษี กรรมการสอบวิทยานิพนธ์ ที่กรุณาตรวจ แก่วิทยานิพนธ์ฉบับนี้ให้มีความสมบูรณ์ยิ่งขึ้น ทั้งนี้ผู้วิจัยขอกราบขอบพระคุณคณาจารย์ภาควิชา สถิติที่ให้โอกาสทางการศึกษาและประสิทธิประสาทความรู้ให้แก่ผู้วิจัยจนกระทั่งสำเร็จการศึกษา

ท้ายนี้ผู้วิจัยขอกราบขอบพระคุณ คุณพ่อ คุณแม่ ที่ให้กำลังใจและส่งเสริมสนับสนุนด้าน การเรียนของผู้วิจัยตลอดมาและขอขอบคุณเพื่อนๆ ทุกคน ที่ให้ความช่วยเหลือต่างๆ และคอยให้ กำลังใจจนกระทั่งวิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ.....	ฎ
<b>บทที่ 1 บทนำ.....</b>	<b>1</b>
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	5
1.3 ข้อตกลงเบื้องต้น.....	5
1.4 สมมติฐานของการวิจัย.....	5
1.5 ขอบเขตของการวิจัย.....	6
1.6 คำจำกัดความที่ใช้ในการวิจัย.....	7
1.7 ขั้นตอนในการวิจัย.....	7
1.8 ประโยชน์ที่คาดว่าจะได้รับ.....	7
1.9 เกณฑ์ที่ใช้ในการตัดสินใจ.....	8
<b>บทที่ 2 ทฤษฎีและสถิติที่ใช้ในการวิจัย.....</b>	<b>10</b>
2.1 การวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ.....	10
2.2 วิธีการเฉลี่ยตัวแบบของเบส์.....	11
2.3 วิธีการเฉลี่ยตัวแบบของเบส์เป็นการหาค่าประกอบของตัวแบบ ด้วยเทคนิคมอนติคาร์โลเมื่อใช้ลูกโซ่มาร์คอฟ.....	14
2.4 วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด.....	17
2.5 วิธีการถดถอยแบบขั้นบันได.....	19
<b>บทที่ 3 วิธีดำเนินการวิจัย.....</b>	<b>21</b>
3.1 การจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล.....	21
3.2 แผนการทดลอง.....	22
3.3 ขั้นตอนในการดำเนินการวิจัย.....	23

<b>บทที่ 4 ผลการวิจัย</b> .....	34
4.1 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกคัดเลือกสมการถดถอยเชิงเบส เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา ซึ่งกำหนดจำนวนตัวแปรอิสระ เป็น 3 5 10 และ 15 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 โดยค่าคงที่สำหรับวิธี $BMA_{MC^3}$ และวิธี OPM คือ (1,5).....	37
4.2 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกคัดเลือกสมการถดถอยเชิงเบส เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา ซึ่งกำหนดจำนวนตัวแปรอิสระ เป็น 3 5 10 และ 15 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 โดยค่าคงที่สำหรับวิธี $BMA_{MC^3}$ และวิธี OPM คือ (1,10).....	51
4.3 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกคัดเลือกสมการถดถอยเชิงเบส เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา ซึ่งกำหนดจำนวนตัวแปรอิสระ เป็น 3 5 10 และ 15 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 โดยค่าคงที่สำหรับวิธี $BMA_{MC^3}$ และวิธี OPM คือ (10,100).....	65
4.4 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกคัดเลือกสมการถดถอยเชิงเบส เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา ซึ่งกำหนดจำนวนตัวแปรอิสระ เป็น 3 5 10 และ 15 ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 โดยค่าคงที่สำหรับวิธี $BMA_{MC^3}$ และวิธี OPM คือ (10,500).....	79
<b>บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ</b> .....	93
5.1 สรุปผลการวิจัย.....	94
5.2 ข้อเสนอแนะ.....	97
รายการอ้างอิง.....	100
ภาคผนวก .....	101
ประวัติผู้เขียนวิทยานิพนธ์.....	126







## สารบัญญภาพ

รูปที่		หน้า
3.1	แสดงผังงานสำหรับขั้นตอนของวิธีมอนติคาร์โล.....	22
3.2	แสดงผังงานสำหรับขั้นตอนในการดำเนินงานวิจัย.....	24
3.3	แสดงผังงานสำหรับขั้นตอนของวิธีการคัดเลือกตัวแปรของเบส์.....	27
3.4	แสดงผังงานสำหรับขั้นตอนของวิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลเมื่อใช้ลูกโซ่มาร์คอฟ.....	28
3.5	แสดงผังงานสำหรับขั้นตอนของวิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด.....	30
3.6	แสดงผังงานสำหรับขั้นตอนของวิธีการถดถอยแบบขั้นบันได.....	32
4.1	แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี $BMA_{MC}$ และวิธี OPM มีค่าคงที่เหมือนกันคือ (1,5).....	39
4.2	แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี $BMA_{MC}$ และวิธี OPM มีค่าคงที่เหมือนกันคือ (1,5).....	42
4.3	แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี $BMA_{MC}$ และวิธี OPM มีค่าคงที่เหมือนกันคือ (1,5).....	45
4.4	แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี $BMA_{MC}$ และวิธี OPM มีค่าคงที่เหมือนกันคือ (1,5).....	48
4.5	แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี $BMA_{MC}$ และวิธี OPM มีค่าคงที่เหมือนกันคือ (1,10).....	53



รูปที่	หน้า
4.13 แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี $BMA_{MC^3}$ และวิธี OPM มีค่าคงที่เหมือนกัน คือ (10,500).....	81
4.14 แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี $BMA_{MC^3}$ และวิธี OPM มีค่าคงที่เหมือนกัน คือ (10,500).....	84
4.15 แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี $BMA_{MC^3}$ และวิธี OPM มีค่าคงที่เหมือนกัน คือ (10,500).....	87
4.16 แสดงการเปรียบเทียบค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี $BMA_{MC^3}$ และวิธี OPM มีค่าคงที่เหมือนกัน คือ (10,500).....	90
5.1 แสดงแผนผังผลสรุปการเลือกใช้วิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมา.....	96
5.2 แสดงแผนผังผลสรุปการเลือกใช้วิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคแกมมาในเชิงปฏิบัติ.....	99