

การวิเคราะห์เชิงเบสส์สำหรับตัวแบบการถดถอยเชิงเส้นเชิงเดียว



นางสาววีรพา ฐานะปรัชญ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ


คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2542

ISBN 974-334-078-5

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

BAYESIAN ANALYSIS FOR SIMPLE LINEAR REGRESSION MODEL



Miss Weerapa Thanaprach

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Statistics

Department of Statistics

Faculty of Commerce and Accountancy

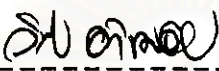
Chulalongkorn University

Academic Year 1999


ISBN 974-334-078-5

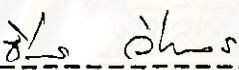
หัวข้อวิทยานิพนธ์ การวิเคราะห์เชิงเบสส์สำหรับตัวแบบการถดถอยเชิงเส้นเชิงเดียว
โดย นางสาววิรพา ฐานะปรัชญ์
ภาควิชา สถิติ
อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร.ธีระพร วีระถาวร


คณะพาณิชย์ศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์
ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต


----- คนบดัดคณะพาณิชย์ศาสตร์และการบัญชี
(ผู้ช่วยศาสตราจารย์ ดร.วิรัช อภิเมธีธำรง)

คณะกรรมการสอบวิทยานิพนธ์


----- ประธานกรรมการ
(รองศาสตราจารย์ ศิริพร สาเกตทอง)


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


----- กรรมการ
(รองศาสตราจารย์ ผกาวดี ศิริรังษี)


----- กรรมการ
(รองศาสตราจารย์ ดร.สุพล ตุงศ์วัฒนา)

สถิตยศาสตร์และการบัญชี
จุฬาลงกรณ์มหาวิทยาลัย

วิธการ ฐานะปรักษณ์ : การวิเคราะห์เชิงเบสส์สำหรับตัวแบบการถดถอยเชิงเส้นเชิงเดี่ยว
(BAYESIAN ANALYSIS FOR SIMPLE LINEAR REGRESSION MODEL)

อ.ที่ปรักษา : รศ. ดร.ธีระพร วีระถาวร , 222 หน้า. ISBN 974-334-078-5.

การวิจัยครั้งนี้มีวัตถุประสงค์ประการแรกเพื่อเปรียบเทียบค่าประมาณสัมประสิทธิ์การถดถอยของตัวแบบการถดถอยเชิงเส้นเชิงเดี่ยวจาก 4 วิธี คือ วิธีกำลังสองน้อยสุด (OLS) วิธีเชิงเบสส์เมื่อใช้การแจกแจงก่อนที่ไม่ให้ข้อมูล (UNI) วิธีเชิงเบสส์เมื่อใช้การแจกแจงก่อนที่ให้ข้อมูล (NOR) และวิธีเชิงเบสส์เมื่อใช้การแจกแจงก่อนของเจฟฟรีย์ส (JEF) เกณฑ์ที่ใช้ในการตัดสินใจ คือ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (AMSE) และวัตถุประสงค์ประการที่สองเพื่อศึกษาวิธี NOR ในกรณีการหาค่า Z ที่เหมาะสมซึ่งทำให้วิธี NOR มีประสิทธิภาพใกล้เคียงกับวิธี OLS และ UNI มากที่สุด โดยที่ค่า Z เป็นค่าที่ทำให้ค่าเฉลี่ยก่อน (prior mean) เบี่ยงเบนจากค่าสัมประสิทธิ์การถดถอย Z เท่าของส่วนเบี่ยงเบนมาตรฐานก่อน (prior standard deviation) ซึ่งวัตถุประสงค์ทั้งสองประการดังกล่าวจะศึกษาเมื่อขนาดตัวอย่างเท่ากับ 10 30 50 และ 100 ความคลาดเคลื่อนมีการแจกแจงปกติซึ่งมีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.1 0.3 0.5 0.7 และ 0.9 ตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติด้วยค่าเฉลี่ยเท่ากับ 1 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.1 0.15 0.2 0.25 0.3 0.5 0.7 และ 0.9 (เปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผัน (CV(X)) เท่ากับ 10% 15% 20% 25% 30% 50% 70% และ 90% ตามลำดับ) ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคมอนติคาร์โลซึ่งกระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์

จากการพิสูจน์ทางพีชคณิตแสดงให้เห็นว่าการแจกแจงภายหลัง (posterior distribution) ที่ได้จากวิธีเชิงเบสส์เมื่อใช้การแจกแจงก่อนของเจฟฟรีย์ส (Jeffreys's prior distribution) เหมือนกันกับการแจกแจงภายหลังที่ได้จากวิธีเชิงเบสส์เมื่อใช้การแจกแจงก่อนที่ไม่ให้ข้อมูล (noninformative prior distribution) ดังนั้น ผู้วิจัยจึงเปรียบเทียบค่าประมาณสัมประสิทธิ์การถดถอยของ 3 วิธี คือ วิธี OLS วิธี UNI และวิธี NOR

ผลการวิจัยจากวัตถุประสงค์ประการแรกพบว่า วิธี OLS และวิธี UNI มีประสิทธิภาพใกล้เคียงกัน โดยที่วิธี OLS จะมีประสิทธิภาพดีกว่าวิธี UNI เสมอ ซึ่งวิธี OLS มีประสิทธิภาพดีเมื่อ CV(X) มีค่าสูง ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าต่ำและขนาดตัวอย่างสูง ส่วนวิธี NOR มีประสิทธิภาพดีเมื่อ CV(X) มีค่าต่ำ ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าสูงและขนาดตัวอย่างต่ำ โดยประสิทธิภาพของตัวประมาณสัมประสิทธิ์การถดถอยของทั้งสามวิธีมีแนวโน้มเพิ่มขึ้นเมื่อ CV(X) และขนาดตัวอย่างเพิ่มขึ้น และมีแนวโน้มลดลงเมื่อส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามเพิ่มขึ้น

ผลการวิจัยจากวัตถุประสงค์ประการที่สองพบว่า ค่า Z ที่เหมาะสมแปรผันตามส่วนเบี่ยงเบนมาตรฐานของตัวแปรตาม แต่แปรผกผันกับ CV(X) และขนาดตัวอย่าง วิธี NOR มีแนวโน้มดีกว่าวิธี OLS และ UNI เมื่อค่า Z ที่เหมาะสมมีแนวโน้มเพิ่มขึ้น แต่วิธี NOR มีแนวโน้มไม่ดีกว่าวิธี OLS และ UNI เมื่อค่า Z ที่เหมาะสมมีแนวโน้มลดลง

ภาควิชา..... สถิติ
สาขาวิชา..... สถิติ
ปีการศึกษา..... 2542

ลายมือชื่อ..... วิธการ ฐานะปรักษณ์
ลายมือชื่ออาจารย์ที่ปรักษา..... อ.ธีระพร วีระถาวร
ลายมือชื่ออาจารย์ที่ปรักษาร่วม.....

กิตติกรรมประกาศ

ขอขอบพระคุณรองศาสตราจารย์ ดร.ธีระพร วีระถาวร เป็นอย่างสูงในความกรุณาที่ให้ความช่วยเหลือและแก้ไขข้อบกพร่องต่างๆ ในวิทยานิพนธ์เล่มนี้ให้สามารถสำเร็จสมบูรณ์ และในความเมตตาที่มีต่อศิษย์ในการอบรมให้มีความคิดและปรับปรุงตนเองไปในทางที่ดีมาโดยตลอด ขอขอบพระคุณผู้ช่วยศาสตราจารย์ ร้อยเอกมานพ วรภักดิ์ ที่ให้คำแนะนำทางด้านการจำลองปัญหาด้วยการใช้โปรแกรมคอมพิวเตอร์ซึ่งเป็นส่วนสำคัญที่จะได้มาซึ่งผลการวิจัยที่สมบูรณ์ และขอขอบพระคุณคณะกรรมการตอบวิทยานิพนธ์ทุกท่านที่ได้กรุณาตรวจทานและแก้ไขให้วิทยานิพนธ์เล่มนี้มีความสมบูรณ์ยิ่งขึ้น ตลอดจนอาจารย์ภาควิชาสถิติทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ให้แก่ศิษย์

สุดท้ายนี้ ขอขอบพระคุณบิดามารดาที่ให้โอกาสทางการศึกษา และขอบคุณสมาชิกทุกคนในครอบครัวและเพื่อนๆ ทุกคนที่เป็นกำลังใจเสมอมา

วีรพา ฐานะปรีชญ์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญ	ช
สารบัญตาราง	ซ
สารบัญรูป	ฅ
สารบัญแผนผัง	ญ
บทที่	
1 บทนำ	1
1.1 ความสำคัญและที่มาของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	2
1.3 สมมติฐานของการวิจัย	2
1.4 ขอบเขตของการวิจัย	2
1.5 เกณฑ์การตัดสินใจ	5
1.6 ประโยชน์ที่คาดว่าจะได้รับ	5
1.7 ตัวสถิติที่ใช้ในการวิจัย	5
2 ทฤษฎีและตัวสถิติที่เกี่ยวข้อง	8
2.1 ตัวแปรตามกำลังสองน้อยสุด	8
2.2 การแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนที่ให้ข้อมูลและการแจกแจงก่อนที่ไม่ให้ ข้อมูล ในกรณีที่มี σ^2 ทราบค่า	11
2.3 การแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนที่ให้ข้อมูลและการแจกแจงก่อนที่ไม่ให้ ข้อมูล ในกรณีที่มี σ^2 ไม่ทราบค่า	16
2.4 การแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนของเจฟฟรีย์	20
3 วิธีดำเนินการวิจัย	23
3.1 แผนการทดลอง	23
3.2 ขั้นตอนการวิจัย	24
3.3 การประมาณค่าสัมประสิทธิ์การถดถอยจากแต่ละวิธี	24
3.4 การหาค่า Z ที่เหมาะสม	25

สารบัญ (ต่อ)

บทที่	หน้า
3.5 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองและการเปรียบเทียบประสิทธิภาพของตัวประ- มาณสัมประสิทธิ์การถดถอย	26
4 ผลการวิจัย	31
4.1 การเปรียบเทียบตัวประมาณสัมประสิทธิ์การถดถอย เมื่อตัวประมาณเบสที่ใช้การ แจกแจงก่อนที่ให้ข้อมูลจะศึกษาในกรณีเฉพาะ	32
4.2 การหาตัวประมาณเบสที่ใช้การแจกแจงก่อนที่ให้ข้อมูลที่ไม่ดีที่สุด แต่ดีกว่าตัวประ- มาณกำลังสองน้อยสุดและตัวประมาณเบสที่ใช้การแจกแจงก่อนที่ไม่ให้ข้อมูล	105
5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ	171
5.1 สรุปผลการวิจัย	171
5.1.1 ผลสรุปจากการเปรียบเทียบตัวประมาณสัมประสิทธิ์การถดถอย เมื่อตัวประ- มาณเบสที่ใช้การแจกแจงก่อนที่ให้ข้อมูลศึกษาในกรณีเฉพาะ	171
5.1.2 ผลสรุปจากการหาตัวประมาณเบสที่ใช้การแจกแจงก่อนที่ให้ข้อมูลที่ ไม่ดีที่สุด แต่ดีกว่าตัวประมาณกำลังสองน้อยสุดและตัวประมาณเบสที่ใช้การแจกแจง ก่อนที่ไม่ให้ข้อมูล	173
5.2 การอภิปรายผล	174
5.3 ข้อเสนอแนะ	175
รายการอ้างอิง	182
ภาคผนวก	183
ประวัติผู้วิจัย	222

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.2.12 การศึกษาตัวประมาณแบบสที่้ใช้การแจกแจงก่อนที่้ให้ข้อมูล เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติโดยมี $CV(X) = 25%$ เมื่อไม่ทราบค่าความแปรปรวน โดยศึกษาในกรณีการหาค่า Z ที่เหมาะสม	147
4.2.13 การศึกษาตัวประมาณแบบสที่้ใช้การแจกแจงก่อนที่้ให้ข้อมูล เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติโดยมี $CV(X) = 30%$ เมื่อไม่ทราบค่าความแปรปรวน โดยศึกษาในกรณีการหาค่า Z ที่เหมาะสม	150
4.2.14 การศึกษาตัวประมาณแบบสที่้ใช้การแจกแจงก่อนที่้ให้ข้อมูล เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติโดยมี $CV(X) = 50%$ เมื่อไม่ทราบค่าความแปรปรวน โดยศึกษาในกรณีการหาค่า Z ที่เหมาะสม	153
4.2.15 การศึกษาตัวประมาณแบบสที่้ใช้การแจกแจงก่อนที่้ให้ข้อมูล เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติโดยมี $CV(X) = 70%$ เมื่อไม่ทราบค่าความแปรปรวน โดยศึกษาในกรณีการหาค่า Z ที่เหมาะสม	156
4.2.16 การศึกษาตัวประมาณแบบสที่้ใช้การแจกแจงก่อนที่้ให้ข้อมูล เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติโดยมี $CV(X) = 90%$ เมื่อไม่ทราบค่าความแปรปรวน โดยศึกษาในกรณีการหาค่า Z ที่เหมาะสม	159
5.1.1 วิธีที่้ให้ประสิทธิภาพดีที่้สุดในแต่ละสถานการณ์ ในกรณีที่้ σ^2 ทราบค่า	172
5.1.2 วิธีที่้ให้ประสิทธิภาพดีที่้สุดในแต่ละสถานการณ์ ในกรณีที่้ σ^2 ไม่ทราบค่า	173

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.2.13 การเปลี่ยนแปลงของค่า Z เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติ โดยมี $CV(X) = 30\%$ โดยที่ไม่ทราบค่าความแปรปรวน	168
4.2.14 การเปลี่ยนแปลงของค่า Z เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติ โดยมี $CV(X) = 50\%$ โดยที่ไม่ทราบค่าความแปรปรวน	168
4.2.15 การเปลี่ยนแปลงของค่า Z เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติ โดยมี $CV(X) = 70\%$ โดยที่ไม่ทราบค่าความแปรปรวน	169
4.2.16 การเปลี่ยนแปลงของค่า Z เมื่อตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากการแจกแจงปกติ โดยมี $CV(X) = 90\%$ โดยที่ไม่ทราบค่าความแปรปรวน	169



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญแนมฝั่ง

แนมฝั่งที่	หน้า
1 การเขียนโปรแกรมในการหาค่าประมาณสั่มประสิทธิ์การถดถอยและค่าเฉลี่ยของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง.....	29
2 การเลือกวิธีประมาณค่าสั่มประสิทธิ์การถดถอยที่ให้ค่า AMSE ต่ำสุดในสถานการณ์ต่างๆ เมื่อ σ^2 ทราบค่า.....	175
3 การเลือกวิธีประมาณค่าสั่มประสิทธิ์การถดถอยที่ให้ค่า AMSE ต่ำสุดในสถานการณ์ต่างๆ เมื่อ σ^2 ไม่ทราบค่า.....	178



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย