

การศึกษาเปรียบเทียบตัวประมาณวิเคราะ

นางสาวเจษฎาพร อุกฤษณ์วิบูลย์ชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

ภาควิชาสถิติ

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ.ศ. 2533

ISBN 974-577-718-8

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

016590

I10811283

A COMPARATIVE STUDY ON RIDGE ESTIMATORS

Miss Jedsadaporn Yuttanawiboonchai

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science

Department of Statistics

Graduate School

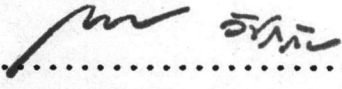
Chulalongkorn University

1990

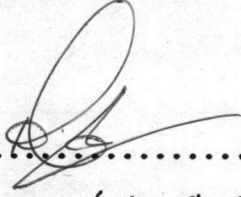
ISBN 974-577-718-8

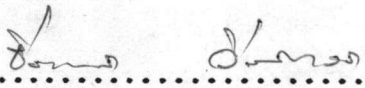
หัวข้อวิทยานิพนธ์ : การศึกษาเปรียบเทียบตัวประมาณวิคัจ
ชื่อนิสิต : นางสาวเจษฎาพร สุทนต์วิบูลย์ชัย
ภาควิชา : สถิติ
อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร. ชีระพร วีระถาวร

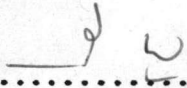
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้เนบวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของ
การศึกษาตามหลักสูตรปริญญาโทบัณฑิต

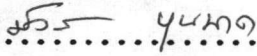

..... คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร. ถาวร วัชรากัย)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(รองศาสตราจารย์ ส่องศรี พิทยารัตน์)


..... อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ ดร. ชีระพร วีระถาวร)


..... กรรมการ
(รองศาสตราจารย์ ดร. สรชัย นิสาลบุตร)


..... กรรมการ
(รองศาสตราจารย์ มัลลิกา บุนนาค)

พิมพ์ต้นฉบับบทความวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

เจษฎาพร มุขธนวิบูลย์ชัย : การศึกษาเปรียบเทียบตัวประมาณริดจ์ (A COMPARATIVE STUDY ON RIDGE ESTIMATORS) อาจารย์ที่ปรึกษา : ผศ.ดร.ธีระพร วีระถาวร, 225 หน้า. ISBN 974-577-718-8

การวิจัยครั้งนี้มีวัตถุประสงค์ที่จะศึกษาเปรียบเทียบตัวประมาณริดจ์ โดยเปรียบเทียบวิธี Hoerl-Kennard-Baldwin (HKB) วิธี TZE-SAN-LEE (TZE-SAN-LEE) และวิธี McDonald-Galarneau (McD&G) เกิดการเปรียบเทียบที่ใช้คือ อัตราส่วนของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง โดยที่การแจกแจงของค่าความคลาดเคลื่อนที่ศึกษา คือ การแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ 0.03 และ 1.0 การแจกแจงแบบปกติปลอมที่มีสเกลแฟกเตอร์เท่ากับ 3 และ 10 เปอร์เซนต์การปลอมปนเท่ากับ 5 และ 10 การแจกแจงแบบลอการิธึมที่มี $\mu = 0$ และ $6^2 = .10$, $.30$ และ $.70$ การแจกแจงแบบไบนอมัลที่มี $\beta = 1$ $\alpha = 1, 2$ และ 10 โดยกำหนดขนาดตัวอย่าง = 10, 30, 50 และ 100 สำหรับจำนวนตัวแปรอิสระเท่ากับ 3 จะกำหนดระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น $.70, .90$ และ $.99$ และสำหรับจำนวนตัวแปรอิสระเท่ากับ 5 จะกำหนดระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น $(.70, .30), (.99, .90)$ และ $(.99, .99)$ ในการทดลองนี้ทำการจำลองเหตุการณ์ต่าง ๆ ด้วยเครื่องคอมพิวเตอร์โดยใช้เทคนิคมอนติคาร์โลและการกระทำซ้ำ 200 ครั้งในแต่ละกรณี

ผลการศึกษาเปรียบเทียบอัตราส่วนของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองสามารถสรุปได้ดังนี้

1. กรณีที่ความคลาดเคลื่อนมีการแจกแจงแบบปกติและปกติปลอมปน โดยส่วนใหญ่วิธี TZE-SAN-LEE ให้ผลดีสำหรับจำนวนตัวแปรอิสระเท่ากับ 3 ยกเว้นเมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติวิธี HKB จะให้ผลดีที่ระดับความสัมพันธ์ $.70$ และวิธี HKB ยังให้ผลดีสำหรับจำนวนตัวแปรอิสระเท่ากับ 5 ผลสรุปได้ว่าจำนวนตัวแปรอิสระ ระดับความสัมพันธ์ ขนาดตัวอย่างและความแปรปรวนมีอิทธิพลต่อค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของทั้ง 3 วิธี โดยที่จำนวนตัวแปรอิสระมีอิทธิพลมากที่สุด

2. กรณีที่ความคลาดเคลื่อนมีการแจกแจงแบบลอการิธึมและไบนอมัล วิธี TZE-SAN-LEE ให้ผลดีสำหรับจำนวนตัวแปรอิสระเท่ากับ 3 ส่วนวิธี HKB ให้ผลดีสำหรับตัวแปรอิสระเท่ากับ 5 วิธี McD&G มีแนวโน้มให้ผลดีเพิ่มขึ้นสำหรับขนาดตัวอย่างเท่ากับ 10 เมื่อระดับความสัมพันธ์เพิ่มขึ้น และทั้งสามวิธีให้ผลใกล้เคียงกันเมื่อความคลาดเคลื่อนมีการแจกแจงแบบลอการิธึม ($6^2 = .10$) และมีการแจกแจงแบบไบนอมัล ($\alpha = 10$) ผลสรุปได้ว่าจำนวนตัวแปรอิสระ ระดับความสัมพันธ์และขนาดตัวอย่างมีอิทธิพลต่อค่าเฉลี่ยความคลาดเคลื่อนของทั้งสามวิธี โดยที่จำนวนตัวแปรอิสระมีอิทธิพลมากที่สุด

ภาควิชา สถิติ
สาขาวิชา สถิติ
ปีการศึกษา 2532

ลายมือชื่อนิสิต มุขธนวิบูลย์ชัย
ลายมือชื่ออาจารย์ที่ปรึกษา ธีระพร วีระถาวร

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

JEDSADAPORN YUTTANAWIBOONCHAI : A COMPARATIVE STUDY ON RIDGE ESTIMATORS. THESIS ADVISOR : ASSIS. PROF. THEERAPORN VERATHAWORN, Ph.D. 225 PP. ISBN 974-577-718-8

The Objective of this thesis is to compare ridge estimators by comparing Hoerl-Kennard-Baldwin method (HKB), TZE-SAN-LEE method (TZE-SAN-LEE) and McDonald-Galarnean method (McD&G). They are compared by using the ratio of average value of mean square error. By studying the residual distributions from normal distribution with mean of 0 and variance of .03 and 1.0, contaminated normal distribution with scale factors of 3 and 10 and percent contaminations of 5 and 10, lognormal distribution with $\mu = 0$ and $\sigma^2 = .10, .30$ and $.70$, and Weibull distribution with $\beta = 1$ and $\alpha = 1, 2$ and 10 . This study used sample sizes of 10, 30, 50 and 100. The correlations among the independent variables are .70, .90 and .99 for the number of independent variables of 3 and (.70, .30), (.99, .90) and (.99, .99) for the number of independent variables of 5. The data is obtained through simulation using Monte Carlo technique, and repeating 200 times for each case.

The results of this study are as follows :

1. In case, residuals have normal and contaminated normal distribution; generally, TZE-SAN-LEE method gives the best result for the number of independent variables of 3 except when residuals have normal distribution with correlation of .70, HKB method gives the best result. HKB method also gives the best result for the number of independent variables of 5. Thus, the number of independent variables, correlation, sample sizes and variance are factors of mean square error which the number of independent variables in the main factor.

2. In case, residuals have lognormal and Weibull distribution; TZE-SAN-LEE method gives the best result for the number of independent variables of 3, HKB method gives the best result for the number of independent variables of 5 and McD&G method gives the best result when the number of sample size is 10 and correlation increases. These methods give close results when residuals have lognormal distribution with σ^2 of .10 and Weibull distribution with α of 10. Thus, the number of independent variables, correlation and sample sizes are factors of mean square error of these methods which the number of independent variables is the main factor.

ภาควิชา สถิติ
สาขาวิชา สถิติ
ปีการศึกษา 2532

ลายมือชื่อนิสิต (พิมพ์) ทวี ทรัพย์วณิช
ลายมือชื่ออาจารย์ที่ปรึกษา [Signature] [Signature]

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาของ ผศ.ดร. ชีระพร วีระถาวร
อาจารย์ประจำภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ที่ให้
คำปรึกษาและแนะนำตลอดจนแก้ไขข้อบกพร่องต่างๆอย่างดีมาโดยตลอด ผู้วิจัยขอกราบขอบพระคุณ
เป็นอย่างสูงไว้ ณ. โอกาสนี้

ขอขอบคุณ คุณศนิवार รอดบุญธรรม และคุณสุวรรณ เกลี้ยงปรามิต สำหรับงานด้านการ
พิมพ์ และขอขอบคุณเพื่อนทุกคนที่มีส่วนช่วยเหลือในการทำวิทยานิพนธ์ครั้งนี้

ขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และพี่ๆ ที่ส่งเสริมและสนับสนุนการเรียนของผู้วิจัย
เสมอมา และขอกราบขอบพระคุณอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทความรู้แก่ผู้วิจัยมา
โดยตลอด

เจษฎาพร ยุทธนวิบูลย์ชัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญตาราง	ฅ
สารบัญรูป	ฉ
บทที่ 1 บทนำ	
1.1 ความสำคัญและความเป็นมาของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	4
1.3 สมมติฐานของการวิจัย	5
1.4 ขีดกลางเบื้องต้น	5
1.5 ขอบเขตของการวิจัย	5
1.6 ประโยชน์ที่คาดว่าจะได้รับ	10
บทที่ 2 ตัวสถิติและผลงานวิจัยที่เกี่ยวข้อง	
2.1 ตัวแบบทั่วไป	11
2.2 การประมาณค่าสัมประสิทธิ์การถดถอยพหุโดยวิธีวิธีกำลังกำลัง	14
2.3 ตัวประมาณวิธี	20
2.4 การแปลงข้อมูลโดยการแปลงที่ใช้การยกกำลัง	21
2.5 ผลงานวิจัยที่เกี่ยวข้อง	23
บทที่ 3 วิธีดำเนินการวิจัย	
3.1 แผนการทดลอง	25
3.2 ขั้นตอนในการวิจัย	26
3.3 การประมาณค่าสัมประสิทธิ์การถดถอยพหุโดยวิธีวิธีกำลังกำลัง	35
3.4 การหาค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของสัมประสิทธิ์การถดถอยพหุ โดยวิธีวิธีกำลังกำลัง	35

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการวิจัย	
4.1 การเปรียบเทียบตัวประมาณริตจ์ในกรณีที่มีความคลาดเคลื่อนมีการ แจกแจงแบบปกติ	44
4.2 การเปรียบเทียบตัวประมาณริตจ์ในกรณีที่มีความคลาดเคลื่อนมีการ มีการแจกแจงแบบปกติปลอมปน	60
4.3 การเปรียบเทียบตัวประมาณริตจ์ในกรณีที่มีความคลาดเคลื่อนมีการ มีการแจกแจงแบบลอกนอร์มอล	89
4.4 การเปรียบเทียบตัวประมาณริตจ์ในกรณีที่มีความคลาดเคลื่อนมีการ มีการแจกแจงแบบไวบูลล์	111
บทที่ 5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ	
5.1 สรุปผลการวิจัย	141
5.2 การอภิปรายผล	147
5.3 ข้อเสนอแนะ	147
บรรณานุกรม	149
ภาคผนวก ก	151
ภาคผนวก ข	160
ประวัติผู้เขียน	214

สารบัญตาราง

ตารางที่	หน้า
1.1 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติ จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .03$	45
1.2 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติ จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = 1$	48
1.3 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติ จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .03$	51
1.4 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติ จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = 1$	54
1.5 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 3 เปอร์เซนต์การปลอมปน = 5 และ สเกลแฟกเตอร์ = 3	61
1.6 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 3 เปอร์เซนต์การปลอมปน = 10 และ สเกลแฟกเตอร์ = 3	64
1.7 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 3 เปอร์เซนต์การปลอมปน = 5 และ สเกลแฟกเตอร์ = 10	67
1.8 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 3 เปอร์เซนต์การปลอมปน = 10 และ สเกลแฟกเตอร์ = 10	70
1.9 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 5 เปอร์เซนต์การปลอมปน = 5 และ สเกลแฟกเตอร์ = 3	73

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
1.10 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 5 เปอร์เซนต์การปลอมปน = 10 และ สเกลแฟกเตอร์ = 3	76
1.11 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 5 เปอร์เซนต์การปลอมปน = 5 และ สเกลแฟกเตอร์ = 10	79
1.12 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จำนวนตัวแปรอิสระ = 5 เปอร์เซนต์การปลอมปน = 10 และ สเกลแฟกเตอร์ = 10	82
1.13 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบลอกนอร์มอล จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .10$	90
1.14 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบลอกนอร์มอล จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .30$	93
1.15 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบลอกนอร์มอล จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .70$	96
1.16 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบลอกนอร์มอล จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .10$	99

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
1.17 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ลอกนอร์มอล จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .30$	102
1.18 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ลอกนอร์มอล จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\mu = 0$ และ $\sigma^2 = .70$	105
1.19 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ไวบูลล์ จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\beta = 1$ และ $\alpha = 1$	112
1.20 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ไวบูลล์ จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\beta = 1$ และ $\alpha = 2$	114
1.21 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ไวบูลล์ จำนวนตัวแปรอิสระ = 3 พารามิเตอร์ $\beta = 1$ และ $\alpha = 10$	117
1.22 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ไวบูลล์ จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\beta = 1$ และ $\alpha = 1$	120
1.23 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ไวบูลล์ จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\beta = 1$ และ $\alpha = 2$	123
1.24 การเปรียบเทียบตัวประมาณวิเศษในกรณีที่มีความคลาดเคลื่อนมีการแจกแจงแบบ ไวบูลล์ จำนวนตัวแปรอิสระ = 5 พารามิเตอร์ $\beta = 1$ และ $\alpha = 10$	126

สารบัญรูป

รูปที่		หน้า
1.5.1	แสดงเส้นโค้งการแจกแจงแบบปกติ $\mu = 0$ และ $\sigma^2 = 0.03$ และ 1.0	6
1.5.2	แสดงเส้นโค้งการแจกแจงแบบปกติปลอมปน μ, p และ c	7
1.5.3	แสดงเส้นโค้งการแจกแจงแบบลอกนอร์มอล $\mu = 0$ และ $\sigma^2 = 0.1,$ 0.3 และ 0.7	8
1.5.4	แสดงเส้นโค้งการแจกแจงแบบไวบูลล์ $\beta = 1$ และ $\alpha = 1, 2$ และ 10	9
3.1	แสดงเส้นโค้งการแจกแจงแบบปกติปลอมปน μ, p และ c	30
3.2	แสดงเส้นโค้งการแจกแจงแบบลอกนอร์มอล $\mu = 0$ และ $\sigma^2 = 0.1,$ 0.3 และ 0.7	31
3.3	แสดงเส้นโค้งการแจกแจงแบบไวบูลล์ $\beta = 1$ และ $\alpha = 1, 2$ และ 10	32