

การประมาณค่าพารามิเตอร์ในการวิเคราะห์ความถดถอยพหุนาม
เมื่อมีความคลาดเคลื่อนเกิดขึ้นในตัวแปรอิสระ



นางสาวธัญจิรัตน์ เมฆบัณฑิตกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2546

ISBN 974-17-3622-3

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

121226878

PARAMETERS-ESTIMATION OF POLYNOMIAL REGRESSION MODELS
WITH ERRORS IN INDEPENDENT VARIABLES

Miss Thitirat Mekbunditkul

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 2003

ISBN 974-17-3622-3

หัวข้อวิทยานิพนธ์

การประมาณค่าพารามิเตอร์ในการวิเคราะห์ความถดถอยพหุนามเมื่อมีความคลาดเคลื่อนเกิดขึ้นในตัวแปรอิสระ

โดย

นางสาวธิวิรัตน์ เมฆบัณฑิตกุล

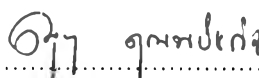
สาขาวิชา

สถิติ

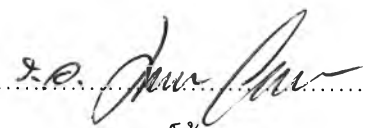
อาจารย์ที่ปรึกษา

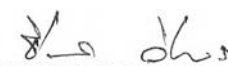
รองศาสตราจารย์ ดร.ธีระพร วีระถาวร

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....  คณบดีคณะพาณิชยศาสตร์และการบัญชี
(ผู้ช่วยศาสตราจารย์ ดร.ดนุชา คุณพนิชกิจ)

คณะกรรมการสอบวิทยานิพนธ์

.....  ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ร้อยเอก มานพ วรภักดิ์)

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

.....  กรรมการ
(รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา)

วิทยุทัศน์ เมฆบัณฑิตกุล : การประมาณค่าพารามิเตอร์ในการวิเคราะห์ความถดถอยพหุนามเมื่อมีความคลาดเคลื่อนเกิดขึ้นในตัวแปรอิสระ (PARAMETERS-ESTIMATION OF POLYNOMIAL REGRESSION MODELS WITH ERRORS IN INDEPENDENT VARIABLES) อ. ที่ปรึกษา : รองศาสตราจารย์ ดร.ธีระพร วีระถาวร , 169 หน้า. ISBN 974-17-3622-3.

การวิจัยครั้งนี้มีวัตถุประสงค์ที่จะเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในตัวแบบการวิเคราะห์ความถดถอยพหุนาม กรณีที่มีความคลาดเคลื่อนในตัวแปรอิสระ โดยจะเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ 3 วิธี ได้แก่ วิธีกำลังสองน้อยสุดสามัญ (Ordinary Least Squares Method (OLS)) วิธีกำลังสองน้อยสุดปรับปรุง (Adjusted Least Squares Method (ALS)) วิธีกำลังสองน้อยสุดถ่วงน้ำหนัก (Weighted Least Squares Method (WLS)) เกณฑ์การเปรียบเทียบที่ใช้ คือ ค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (Average Relative Root Mean Squares Error (ARRMSE)) และใช้อัตราส่วนผลต่างของค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (Ratio of Different Average Relative Root Mean Squares Error (RDARRMSE)) เป็นเกณฑ์ประกอบการตัดสินใจ การแจกแจงของความคลาดเคลื่อนสุ่มในตัวแปรตามที่ใช้ศึกษาคือการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 0.1, 0.3, 0.5, 0.7 และ 1.0 การแจกแจงของความคลาดเคลื่อนในตัวแปรอิสระที่ใช้ศึกษาคือการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 0.1, 0.3, 0.5 และ 0.7 ขนาดตัวอย่างเป็น 15, 30, 50, 100 และ 200 กำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (highest degree of independent variables for dependent variable building in model (MB)) เป็น 2, 3, 4, 5 และ 6 สำหรับข้อมูลที่ใช้ในการวิจัยใช้โปรแกรมคอมพิวเตอร์จำลองด้วยเทคนิคมอนติคาร์โลกระทำซ้ำ 1000 ครั้ง ในแต่ละสถานการณ์ ซึ่งผลการวิจัยสรุปได้ดังนี้

ค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (ARRMSE) แปรผันตามปัจจัยต่อไปนี้จากมากไปหาน้อย กำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (MB) ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระ (σ_ϵ) และส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มในตัวแปรตาม (σ_δ) แต่แปรผกผันกับขนาดตัวอย่าง (n) ซึ่งอิทธิพลของปัจจัยดังกล่าวเป็นดังนี้

1. กรณีที่ขนาดตัวอย่างมีค่าน้อย ($n = 15$)

ถ้ากำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับสร้างตัวแปรตามในตัวแบบถดถอยพหุนามมีค่าน้อย ($MB = 2$) ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มในตัวแปรตามและส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระมีค่าน้อย ($\sigma_\delta \leq 0.3$ และ $\sigma_\epsilon \leq 0.3$) วิธี OLS จะเป็นวิธีที่ดีที่สุด รองลงมาคือวิธี ALS และ WLS ตามลำดับ ยกเว้นกรณีที่มีส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มในตัวแปรตามมีค่าเท่ากับ 0.3 และส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระมีค่าเท่ากับ 0.3 วิธี ALS จะเป็นวิธีที่ดีที่สุด รองลงมาคือวิธี WLS และ OLS ตามลำดับ และถ้าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มในตัวแปรตาม และส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระมีค่าปานกลางถึงมาก ($\sigma_\delta \geq 0.5$ และ $\sigma_\epsilon \geq 0.5$) วิธี WLS จะเป็นวิธีที่ดีที่สุด รองลงมาคือวิธี ALS และ OLS ตามลำดับ

2. กรณีที่ขนาดตัวอย่างมีค่าปานกลางถึงมาก ($n > 15$)

ถ้ากำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับสร้างตัวแปรตามในตัวแบบถดถอยพหุนามมีค่ามาก ($MB > 2$) วิธี WLS จะเป็นวิธีที่ดีที่สุด รองลงมาคือ วิธี ALS และ OLS ตามลำดับ สำหรับทุกค่าของส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มในตัวแปรตาม (σ_δ) และส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในตัวแปรอิสระ (σ_ϵ)

ภาควิชา สถิติ

สาขาวิชา สถิติ

ปีการศึกษา 2546

ลายมือชื่อนิสิต..... *Thitane*

ลายมือชื่ออาจารย์ที่ปรึกษา..... *Dr. Theeraporn*

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4482263726 : MAJOR STATISTICS

KEY WORD : POLYNOMIAL REGRESSION/ ADJUSTED LEAST SQUARES/ WEIGHTED LEAST SQUARES/
 ERRORS IN INDEPENDENT VARIABLES

THITIRAT MEKBUNDITKUL : PARAMETERS-ESTIMATION OF POLYNOMIAL REGRESSION MODELS WITH
 ERRORS IN INDEPENDENT VARIABLES. THESIS ADVISOR : ASSOC.PROF.THEERAPORN VERATHAWORN,
 Ph.D. 169 pp. ISBN 974-17-3622-3

The purpose of this research is to compare the method of parameter estimations in polynomial regression models with errors in independent variables by Ordinary least squares method (OLS) , Adjusted least squares method (ALS) and Weighted least squares method (WLS). The criterion of comparison is average relative root mean squares error (ARRMSE) and use ratio of different average relative root mean squares error (RDARRMSE) to support decision. The distribution of random errors are normal distribution with mean equal to 0 and standard deviation equal to 0.1, 0.3, 0.5, 0.7 and 1.0, respectively. For the distribution of errors in independent variables are normal distribution with mean equal to 0 and standard deviation equal to 0.1, 0.3, 0.5 and 0.7, respectively. The sample sizes used in this study are 15, 30, 50, 100 and 200, respectively, highest degree of independent variables for dependent variable building in model (MB) are 2, 3, 4, 5 and 6, respectively. The data for this experiment are generated through the Monte Carlo simulation technique and repeating 1000 times for each case. The results of this research are as follow :

The average relative root mean squares error vary with , most to least , respectively , highest degree of independent variables for dependent variable building in model (MB) , standard deviation of errors in independent variables (σ_g) , standard deviation of random errors (σ_e) but converse to sample size (n).

1. In case of sample size is low ($n = 15$).

If MB is low (MB = 2) , standard deviation of random errors (σ_e) and standard deviation of errors in independent variables (σ_g) are low ($\sigma_e \leq 0.3$ and $\sigma_g \leq 0.3$) the OLS method is the best, the ALS method and the WLS method , respectively. Except in case standard deviation of random errors (σ_e) and standard deviation of errors in independent variables (σ_g) equal to 0.3 the ALS method is the best , the WLS method and the OLS method , respectively. When standard deviation of random errors (σ_e) and standard deviation of errors in independent variables (σ_g) are medium to high ($\sigma_e \geq 0.5$ and $\sigma_g \geq 0.5$) , the WLS method is the best, the ALS method and the OLS method , respectively

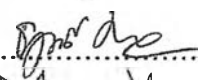
2. In case of sample size is medium to high ($n > 15$).

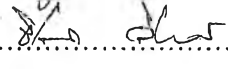
If MB is medium to high (MB > 2) , for all standard deviation of random errors (σ_e) and standard deviation of errors in independent variables (σ_g), the WLS method is the best, the ALS method and the OLS method , respectively.

Department Statistics

Field of study Statistics

Academic year 2003

Student's signature.....

Advisor's signature.....

Co-advisor's signature.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความช่วยเหลืออย่างดียิ่งจาก รองศาสตราจารย์ ดร.ธีระพร วีระถาวร อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำปรึกษา แนะนำ และแก้ไขข้อบกพร่องต่าง ๆ เป็นอย่างดีมาโดยตลอด ผู้วิจัยใคร่ขอกราบขอบพระคุณเป็นอย่างสูง

ผู้เขียนขอกราบขอบพระคุณ คณะกรรมการสอบวิทยานิพนธ์ ซึ่งประกอบด้วย ผู้ช่วยศาสตราจารย์ ร้อยเอกมานพ วรภักดิ์ ที่กรุณาให้คำปรึกษาด้านการสร้างข้อมูลเป็นอย่างดี อีกทั้งได้กรุณาช่วยตรวจและแก้ไขวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น และ รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา ที่ได้กรุณาช่วยตรวจและแก้ไขวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น

ผู้เขียนขอขอบคุณเจ้าหน้าที่ห้องสมุด คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ที่ได้อำนวยความสะดวกในด้านตำราซึ่งใช้ค้นคว้าประกอบการทำวิทยานิพนธ์ และขอขอบคุณ พี่ ๆ เพื่อน ๆ ทุกคน ที่ให้ความช่วยเหลือ และเป็นกำลังใจให้ผู้เขียน

สุดท้ายนี้ผู้เขียนขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และขอขอบคุณ พี่สาว และ พี่ชายของผู้เขียน ที่ให้กำลังใจและส่งเสริมสนับสนุนด้านการเรียนของผู้เขียนตลอดมา

ธัญรัตน์ เมฆบัณชิตกุล

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญตาราง.....	ฅ
สารบัญรูป.....	ฎ
บทที่ 1 บทนำ.....	1
ที่มาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ของการวิจัย.....	4
ข้อตกลงเบื้องต้น.....	4
สมมติฐานทางการวิจัย.....	4
ขอบเขตการวิจัย.....	4
ประโยชน์ของการวิจัย.....	6
เกณฑ์การตัดสินใจ.....	6
บทที่ 2 สถิติที่ใช้ในการวิจัย.....	8
การวิเคราะห์ความถดถอยพหุนาม.....	8
การประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณโดยวิธีกำลังสองน้อยสุดสามัญ...	13
การประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณโดยวิธีกำลังสองน้อยสุดปรับปรุง..	14
การประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณโดยวิธีกำลังสองน้อยสุด	
ถ่วงน้ำหนัก.....	18
บทที่ 3 วิธีดำเนินการวิจัย.....	21
การหาข้อสรุปโดยใช้การจำลองของข้อมูลด้วยวิธีมอนติคาร์โล.....	21
แผนการทดลอง.....	22
ขั้นตอนในการศึกษาวิจัย.....	22
บทที่ 4 ผลการวิจัย.....	33

สารบัญ (ต่อ)

บทที่	หน้า
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	115
ปัจจัยที่มีผลต่อค่าเฉลี่ยราคของค่าตลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ ของแต่ละวิธี.....	115
การเปรียบเทียบค่าเฉลี่ยราคของค่าตลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ ของแต่ละวิธี.....	117
ข้อเสนอแนะ.....	117
รายการอ้างอิง.....	126
ภาคผนวก.....	127
ประวัติผู้เขียนวิทยานิพนธ์.....	169

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.19 แสดงค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (<i>ARRMSE</i>) และค่าอัตราส่วนผลต่างของค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (<i>RDARRMSE</i>) เมื่อ $\sigma_{\epsilon} = 1.0$ และ $\sigma_{\delta} = 0.5$	91
4.20 แสดงค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (<i>ARRMSE</i>) และค่าอัตราส่วนผลต่างของค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ (<i>RDARRMSE</i>) เมื่อ $\sigma_{\epsilon} = 1.0$ และ $\sigma_{\delta} = 0.7$	94

สารบัญรูป

รูปที่	หน้า
4.1 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.1$ และ $\sigma_{\delta} = 0.1$	38
4.2 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.1$ และ $\sigma_{\delta} = 0.3$	41
4.3 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.1$ และ $\sigma_{\delta} = 0.5$	43
4.4 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.1$ และ $\sigma_{\delta} = 0.7$	47
4.5 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.3$ และ $\sigma_{\delta} = 0.1$	51
4.6 แสดงการเปรียบเทียบค่า <i>RRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.3$ และ $\sigma_{\delta} = 0.3$	54
4.7 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.3$ และ $\sigma_{\delta} = 0.5$	57
4.8 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.3$ และ $\sigma_{\delta} = 0.7$	60
4.9 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.5$ และ $\sigma_{\delta} = 0.1$	63
4.10 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.5$ และ $\sigma_{\delta} = 0.3$	66
4.11 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.5$ และ $\sigma_{\delta} = 0.5$	69
4.12 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.5$ และ $\sigma_{\delta} = 0.7$	72
4.13 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.7$ และ $\sigma_{\delta} = 0.1$	75
4.14 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีการที่ $\sigma_{\varepsilon} = 0.7$ และ $\sigma_{\delta} = 0.3$	78

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.15 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีกรณีที่ $\sigma_{\varepsilon} = 0.7$ และ $\sigma_{\delta} = 0.5$	81
4.16 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีกรณีที่ $\sigma_{\varepsilon} = 0.7$ และ $\sigma_{\delta} = 0.7$	84
4.17 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีกรณีที่ $\sigma_{\varepsilon} = 1.0$ และ $\sigma_{\delta} = 0.1$	87
4.18 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีกรณีที่ $\sigma_{\varepsilon} = 1.0$ และ $\sigma_{\delta} = 0.3$	90
4.19 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีกรณีที่ $\sigma_{\varepsilon} = 1.0$ และ $\sigma_{\delta} = 0.5$	93
4.20 แสดงการเปรียบเทียบค่า <i>ARRMSE</i> และแนวโน้มของทั้ง 3 วิธีกรณีที่ $\sigma_{\varepsilon} = 1.0$ และ $\sigma_{\delta} = 0.7$	96