การเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

โดย
นางสาวชญาณัฐ โพธิ์นอก

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2552

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ

โดย
นางสาวชญาณัฐ โพธิ์นอก

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2552

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

COMPARISONS OF PARAMETER ESTIMATION APPROACHES OF GENERALIZED LINEAR MIXED MODELS FOR ORDINAL RESPONSES

By

Chayanat Phonork

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

Department of Statistics

Graduate School

SILPAKORN UNIVERSITY

2009

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง "การเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ" เสนอโดย นางสาวชญาณัฐ โพธิ์นอก เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์

…….....

(รองศาสตราจารย์ ดร. ศิริชัย ชินะตังกูร)

คณบดีบัณฑิตวิทยาลัย

วันที่.....เดือน..... พ.ศ.....

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รองศาสตราจารย์ วีรานันท์ พงศาภักดี

คณะกรรมการตรวจสอบวิทยานิพนธ์

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

..... ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. กมลชนก พานิชการ)

...../...../.....

..... กรรมการ

(รองศาสตราจารย์ ดร. กมล บุษบา)

...../...../.....

..... กรรมการ

(รองศาสตราจารย์ วีรานันท์ พงศาภักดี)

...../...../.....

50304202 : สาขาวิชาสถิติประยุกต์
คำสำคัญ : ตัวแปรตอบสนองแบบอันดับ/ ลอจิตสะสม/ ตัวแบบมาร์จินัล/ ตัวแบบอิทธิพลสุ่ม/ Generalized Estimating Equations/ Marginal Quasi-Likelihood/ Adaptive Gaussian Quadrature/ Penalized Quasi-Likelihood

ชญาณัฐ โพธิ์นอก : การเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ. อาจารย์ที่ปรึกษาวิทยานิพนธ์ : รศ. วีรานันท์ พงศาภักดี. 87 หน้า.

ตัวแบบผสมเชิงเส้นนัยทั่วไป (GLMMs) เป็นตัวแบบที่นิยมใช้สำหรับข้อมูลที่มีความสัมพันธ์กัน วิธีประมาณพารามิเตอร์สำหรับตัวแบบนี้ค่อนข้างซับซ้อน โดยเฉพาะสำหรับตัวแปรตอบสนองแบบอันดับ งานวิจัยนี้มีจุดประสงค์ 2 ประการ ประการแรก คือ เพื่อเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบผสมเชิงเส้นนัยทั่วไป ภายใต้ตัวแบบลอจิตสะสมแบบมาร์จินัล ระหว่างวิธี Generalized Estimating Equation (GEE) และ Marginal Quasi-Likelihood (MQL) และตัวแบบลอจิตสะสมแบบอิทธิพลสุ่ม ระหว่างวิธี Adaptive Gaussian Quadrature (AGQ) และ Penalized Quasi-Likelihood (PQL) ประการที่สอง คือ เพื่อประเมินผลกระทบของขนาดตัวอย่าง (จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่าง) และสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ โดยพิจารณาจากความเอนเอียง ความคลาดเคลื่อนกำลังสองเฉลี่ยและสัดส่วนของช่วงความเชื่อมั่น 95% ที่ครอบคลุมค่าจริง จากการจำลองแบบข้อมูล จำนวน 1,000 ครั้ง

ผลการวิจัยพบว่าสำหรับตัวแบบมาร์จินัล วิธี MQL มีความถูกต้องในการประมาณพารามิเตอร์สูงกว่าวิธี GEE เล็กน้อย เมื่อสหสัมพันธ์ในกลุ่มมีขนาดกลางและขนาดใหญ่ แต่มีความถูกต้องในการประมาณพารามิเตอร์ใกล้เคียงกับวิธี GEE และใกล้เคียงกับวิธี ML สำหรับตัวแบบอิทธิพลคงที่ เมื่อสหสัมพันธ์ในกลุ่มมีขนาดเล็ก สำหรับตัวแบบอิทธิพลสุ่ม พบว่าวิธี AGQ มีความถูกต้องมากกว่าวิธี PQL สรุปโดยรวมทั้งตัวแบบอิทธิพลคงที่ ตัวแบบมาร์จินัล และตัวแบบอิทธิพลสุ่ม พบว่าวิธี AGQ เป็นวิธีที่มีความถูกต้องมากที่สุด นอกจากนี้ยังพบว่าเมื่อขนาดตัวอย่างเพิ่มขึ้นส่งผลให้ความถูกต้องในการประมาณพารามิเตอร์สูงขึ้น โดยเฉพาะวิธี AGQ และ PQL และเมื่อสหสัมพันธ์ในกลุ่มเพิ่มขึ้นส่งผลให้ความถูกต้องในการประมาณพารามิเตอร์ของวิธี GEE และ MQL ลดลงอย่างมาก แต่ลดลงเพียงเล็กน้อยสำหรับวิธี AGQ และ PQL

ภาควิชาสถิติ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ปีการศึกษา 2552

ลายมือชื่อนักศึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์ .....

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

50304202 : MAJOR : APPLIED STATISTICS
KEY WORDS : ORDINAL RESPONSE/ CUMULATIVE LOGITS / MARGINAL MODELS/ RANDOM EFFECTS MODELS/ GENERALIZED ESTIMATING EQUATIONS/ MARGINAL QUASI-LIKELIHOOD/ ADAPTIVE GAUSSIAN QUADRATURE/ PENALIZED QUASI-LIKELIHOOD

CHAYANAT PHONORK : COMPARISONS OF PARAMETER ESTIMATION APPROACHES OF GENERALIZED LINEAR MIXED MODELS FOR ORDINAL RESPONSES. THESIS ADVISOR : ASSOC. PROF. VEERANUN PONGSAPUKDEE. 87 pp.

The Generalized Linear Mixed Models (GLMMs) are popular models for correlated data. Parameter estimation approaches for these models are complicated, especially those for ordinal response variables. There are two objectives of this research. First, to compare the parameter estimation approaches for GLMMs using cumulative logits based on marginal and random effects models. In marginal model, two approaches, Generalized Estimating Equations (GEE) and Marginal Quasi-Likelihood (MQL) are performed. In random effects model, the other two approaches, Penalized Quasi-Likelihood (PQL) and Adaptive Gaussian Quadrature (AGQ) are compared. Second, to assess the influence of sample size (number of clusters and cluster size) and the intra-cluster correlations to the accuracy of parameter estimates from the bias, mean square error and the coverage from 1,000 datasets.

The research results indicate that the better approach for the marginal model is MQL compared to the GEE when the intra-cluster correlation is medium and large but they are approximately the same when the intra-cluster correlation is small. For the random effects model, the AGQ is better than the PQL. Generally conclusion, the AGQ is the best for all models. In addition, as the number of cluster and the cluster size are increased, the accuracy of parameter estimates is improved. Moreover, the increasing of the intra-cluster correlation affects the GEE and the MQL estimates but such impact is only little for the AGQ and the PQL estimates.

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

Department of Statistics Graduate School, Silpakorn University Academic Year 2009
Student's signature .....
Thesis Advisor's signature .....

กิตติกรรมประกาศ

การเรียบเรียงวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงด้วยดี ผู้วิจัยได้รับความอนุเคราะห์จากรองศาสตราจารย์วีรานันท์ พงศาภักดี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้ความรู้และคำแนะนำในการเรียบเรียงและแก้ไขข้อบกพร่องของวิทยานิพนธ์ ด้วยความรัก ความเมตตาและความเอาใจใส่อย่างมาก

ขอกราบขอบพระคุณผู้ช่วยศาสตราจารย์ ดร. กมลชนก พานิชการ ประธานกรรมการตรวจสอบวิทยานิพนธ์ และรองศาสตราจารย์ ดร. กมล บุษบา กรรมการผู้ทรงคุณวุฒิตรวจสอบวิทยานิพนธ์ ที่ให้ความกรุณาแนะนำข้อบกพร่อง เพื่อแก้ไขปรับปรุงให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบคุณบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากรที่มอบทุนอุดหนุนวิทยานิพนธ์ และขอบพระคุณคณาจารย์ภาควิชาสถิติ ที่พิจารณาให้ได้รับทุน ซึ่งเป็นส่วนหนึ่งที่ทำให้มีโอกาสได้ใช้เวลาส่วนใหญ่ทุ่มเทให้กับงานวิจัยในครั้งนี้

สุดท้ายนี้ขอกราบขอบพระคุณบิดา มารดา คณาจารย์ ที่ให้กำลังใจและให้การสนับสนุนทางด้านการศึกษา ตลอดจนเพื่อน ๆ ที่ให้คำแนะนำ มีการแลกเปลี่ยนความรู้ ในช่วงเวลาของการศึกษาและการเรียบเรียงวิทยานิพนธ์ตลอดมา

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

สารบัญ

		หน้า
บทคัดย่อภาษาไทย		ง
บทคัดย่อภาษาอังกฤษ		จ
กิตติกรรมประกาศ		ฉ
สารบัญตาราง		ฌ
สารบัญภาพ		ญ
บทที่		
1	บทนำ	1
	ความเป็นมาและความสำคัญของปัญหา	1
	วัตถุประสงค์ของการวิจัย	5
	สมมติฐานของการวิจัย	5
	ขอบเขตของการวิจัย	6
	นิยามศัพท์เฉพาะ	8
	ประโยชน์ที่คาดว่าจะได้รับ	8
2	ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	9
	ทฤษฎีที่เกี่ยวข้อง	9
	ตัวแบบเชิงเส้นนัยทั่วไป	9
	ตัวแบบอิทธิพลคงที่และการประมาณพารามิเตอร์ของตัวแบบ	10
	ตัวแบบมาร์จินัลและการประมาณพารามิเตอร์ของตัวแบบ	12
	ตัวแบบผสมเชิงเส้นนัยทั่วไป	16
	ตัวแบบอิทธิพลสุ่มและการประมาณพารามิเตอร์ของตัวแบบ	16
	งานวิจัยที่เกี่ยวข้อง	21
3	วิธีดำเนินงานวิจัย	25
	ข้อมูลที่ใช้ในการวิจัย	25
	เครื่องมือที่ใช้ในการวิจัย	25
	สถิติที่ใช้ในการวิจัย	25
	ขั้นตอนการจำลองแบบข้อมูล	26
	วิธีการวิเคราะห์ข้อมูล	28

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

บทที่		หน้า
4	ผลการวิจัย	29
	การเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัล	30
	การเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่ม	42
	การศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่าง และสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์	54
5	สรุป อภิปรายผล และข้อเสนอแนะ	70
	สรุปผลการวิจัย	70
	อภิปรายผลการวิจัย	71
	ข้อเสนอแนะของการวิจัย	73
	ข้อเสนอแนะในการวิจัยครั้งต่อไป	73
บรรณานุกรม		75
ภาคผนวก		78
	โปรแกรมคอมพิวเตอร์ที่ใช้ในการวิจัย	79
ประวัติผู้วิจัย		87

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

สารบัญตาราง

ตารางที่		หน้า
1	ชนิดของตัวแบบเชิงเส้นนัยทั่วไป สำหรับการวิเคราะห์เชิงสถิติ	30
2	ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.05	30
3	ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่า Coverage สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.05	32
4	ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.20	34
5	ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่า Coverage สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.20	36
6	ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.40	38
7	ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่า Coverage สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.40	40
8	ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.05	42
9	ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่า Coverage สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.05	44
10	ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.20	46
11	ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่า Coverage สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.20	48
12	ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.40	50
13	ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่า Coverage สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.40	52

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

สารบัญภาพ

ภาพที่ หน้า

1 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_1 54

2 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการการประมาณพารามิเตอร์ α_2 .. 55

3 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_1 56

4 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_2 57

5 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ σ_u^2 58

6 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_1 59

7 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_2 60

8 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_1 61

9 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_2 62

10 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ σ_u^2 63

11 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_1 64

12 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_2 65

13 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_1 66

14 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_2 67

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ หน้า

15 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ในการศึกษาผลกระทบ
ของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ σ_u^2 68

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

บทที่ 1

บทนำ

1. ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันมีการนำความรู้และวิธีการทางสถิติไปประยุกต์ใช้กับงานด้านต่าง ๆ อย่างกว้างขวางมากขึ้น ทั้งนี้เป็นเพราะว่าวิธีการทางสถิตินั้นมีระเบียบวิธีการดำเนินการอย่างเป็นระบบภายใต้เหตุและผล ดังนั้นการนำไปประยุกต์ใช้ในด้านต่าง ๆ จึงจำเป็นต้องเลือกวิธีที่เหมาะสมกับวัตถุประสงค์ของงานวิจัยและข้อมูลที่สนใจศึกษา

ข้อมูลที่มีตัวแปรตอบสนองแบบอันดับพบในงานวิจัยหลายด้าน เช่นด้านการแพทย์ ตัวแปรตอบสนองคืออาการของโรค (เบา, ปานกลาง, รุนแรง) ด้านการตลาด ตัวแปรตอบสนองคือความพึงพอใจของลูกค้า (น้อยที่สุด, น้อย, ปานกลาง, มาก, มากที่สุด) หรือด้านการศึกษา ตัวแปรตอบสนองคือความสามารถของนักเรียน (อ่อน, ปานกลาง, เก่ง) ตัวแบบที่นิยมใช้สำหรับตัวแปรตอบสนองแบบอันดับคือ "ตัวแบบลอจิตสะสม" ซึ่งเป็นตัวแบบอิทธิพลคงที่ตัวแบบหนึ่งภายใต้ชื่อ "ตัวแบบเชิงเส้นนัยทั่วไป" (Generalized Linear Models: GLMs) โดยมีข้อสมมติว่าค่าสังเกตแต่ละค่าเป็นอิสระต่อกัน

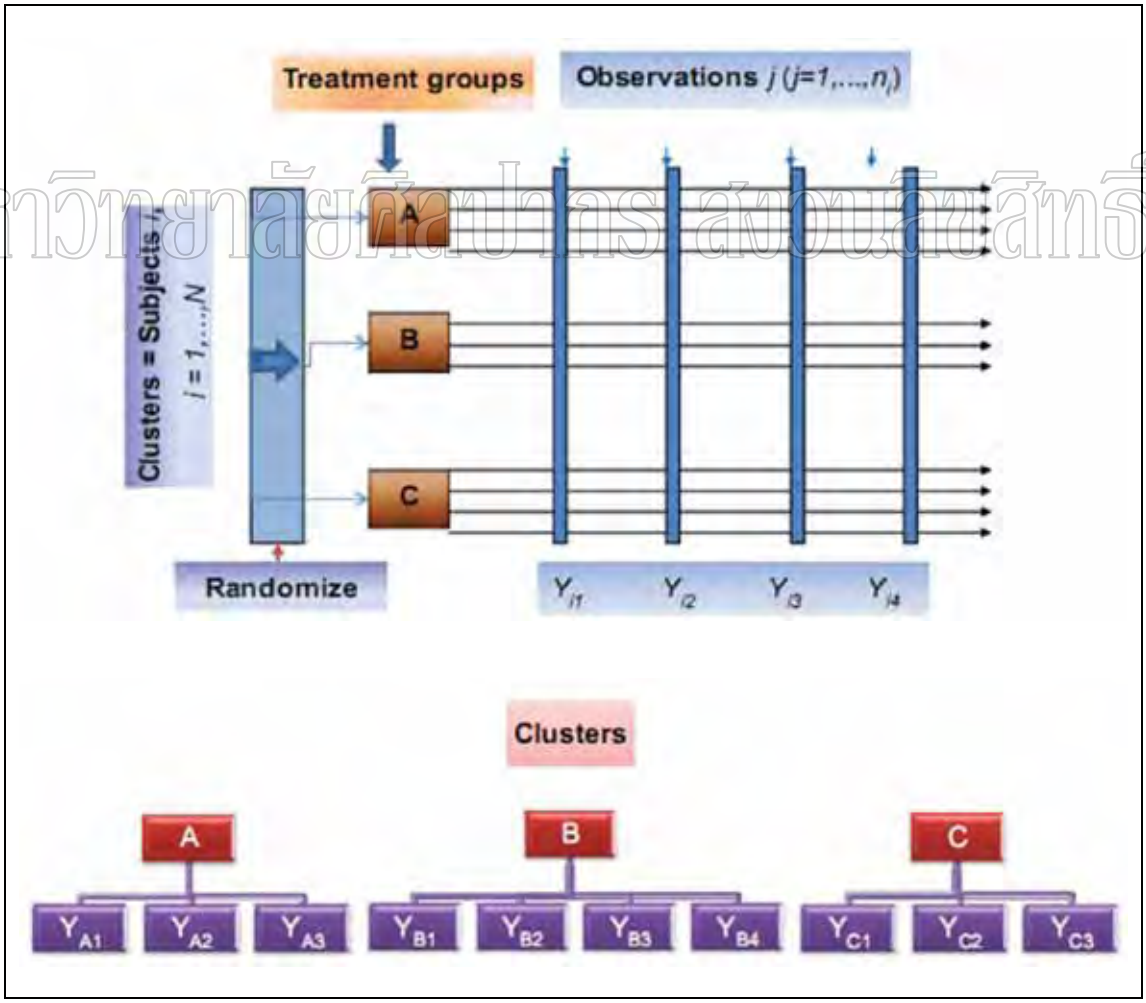
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ในทางปฏิบัติมักพบว่าค่าสังเกตไม่เป็นอิสระต่อกัน อาจเนื่องมาจากการที่หน่วยสังเกตถูกวัดซ้ำในช่วงเวลาหนึ่ง ข้อมูลลักษณะนี้เรียกว่า "Repeated measure data" ดังแผนภาพที่ 1 หรือเป็นค่าสังเกตจากหน่วยสังเกตที่ซ้อนอยู่ในกลุ่ม (เช่น ห้องเรียน ละแวกบ้าน ครัวเรือน สถานที่ทำงาน หรือครอกสัตว์) ข้อมูลลักษณะนี้เรียกว่า "Clustered data" ดังแผนภาพที่ 1 ซึ่งคาดว่าค่าสังเกตจากกลุ่มเดียวกันจะมีความสัมพันธ์กัน เช่นอาการของโรคของผู้ป่วยแต่ละคนมีการประเมินในแต่ละครั้งที่พบแพทย์ ดังนั้นค่าสังเกตที่มาจากผู้ป่วยคนหนึ่ง ๆ น่าจะมีความสัมพันธ์กัน หรือความสามารถของนักเรียนซึ่งวัดจากนักเรียนแต่ละห้องเรียน ค่าสังเกตจากนักเรียนแต่ละคนในห้องเรียนหนึ่ง ๆ น่าจะมีความสัมพันธ์กัน

การวิเคราะห์ข้อมูลที่มีความสัมพันธ์กันนี้ โดยใช้ตัวแบบเชิงเส้นนัยทั่วไป ซึ่งเป็นกลุ่มของตัวแบบอิทธิพลคงที่ ภายใต้ข้อสมมติว่าค่าสังเกตแต่ละค่าเป็นอิสระกันนั้น อาจจะให้ผลสรุปที่ผิดพลาดได้ (Agresti 2000 : 491)

ทางเลือกของปัญหานี้ก็คือการใช้ตัวแบบผสมเชิงเส้นนัยทั่วไป (Generalized Linear Mixed Models: GLMMs) ซึ่งขยายจากตัวแบบตัวแบบเชิงเส้นนัยทั่วไป โดยนำความสัมพันธ์ระหว่างค่าตอบสนองในกลุ่มมาพิจารณาด้วย ซึ่งมีกลุ่มของตัวแบบสำหรับค่าตอบสนองที่มีความสัมพันธ์กันอยู่ 2 กลุ่มคือตัวแบบมาร์จินัลหรือเรียกว่า Population-averaged models และตัวแบบอิทธิพลสุ่มหรือเรียกว่า Cluster-specific models หรือ Subject-specific models การเลือกใช้ตัวแบบใดนั้นขึ้นอยู่กับความต้องการสรุปผลในระดับประชากรหรือระดับกลุ่ม

จากข้อมูลที่ประกอบด้วยกลุ่มที่เป็นอิสระกัน N กลุ่ม โดยที่กลุ่ม i มีค่าสังเกต n_i ค่าสังเกต กำหนดให้ Y_{ij} แทนค่าตอบสนองที่ j, $j = 1,\ldots, n_i$ ในกลุ่ม i, $i = 1,\ldots, N$ ซึ่งมีค่าตกอยู่ในระดับที่ k, $k = 1, 2,\ldots, c$ เมื่อ c แทนจำนวนระดับของตัวแปรตอบสนอง และ $\boldsymbol{x}_{ij} = (x_{1ij},\ldots,x_{pij})'$ แทนเวกเตอร์ของตัวแปรอธิบาย p ตัว สำหรับค่าตอบสนองที่ j ในกลุ่ม i



มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

แผนภาพที่ 1 ลักษณะของ Repeated measure data (บน) และ Clustered data (ล่าง)

สำหรับตัวแบบอิทธิพลคงที่ซึ่งไม่คำนึงความสัมพันธ์ระหว่างค่าตอบสนองในกลุ่ม ตัวแบบลอจิตสะสมจะอยู่ในรูป

$$\text{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij})] = \alpha_k + \boldsymbol{x}'_{ij}\boldsymbol{\beta}, \quad k = 1,\ldots,c-1$$

โดยที่ α_k, $\alpha_1 < \alpha_2 < \ldots < \alpha_{c-1}$ คือค่าเริ่มต้นหรือจุดตัดสำหรับระดับที่ k ของตัวแปรตอบสนองจำนวน $c-1$ ค่า และ $\boldsymbol{\beta}$ คือเวกเตอร์ของพารามิเตอร์อิทธิพลคงที่สำหรับตัวแปรอธิบาย p ตัว

สำหรับตัวแบบมาร์จินัล จะกำหนดตัวแบบเหมือนกับตัวแบบอิทธิพลคงที่ แต่จะพิจารณาความสัมพันธ์ระหว่างค่าตอบสนองในกลุ่มไว้ในเมทริกซ์ความแปรปรวนร่วมของตัวแปรตอบสนองที่เรียกว่า Working covariance matrix

สำหรับตัวแบบอิทธิพลสุ่ม จะพิจารณาความสัมพันธ์ระหว่างค่าตอบสนองในกลุ่ม โดยเพิ่มเทอมของอิทธิพลสุ่ม $\boldsymbol{u}_i$ เข้าไปในตัวพยากรณ์เชิงเส้นของตัวแบบลอจิตสะสมจึงเรียกว่าตัวแบบลอจิตสะสมแบบอิทธิพลสุ่ม ซึ่งมีรูปแบบคือ

$$\text{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i)] = \alpha_k + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{u}_i, \quad k = 1,\ldots,c-1$$

โดยที่ $\boldsymbol{u}_i$ แทนเวกเตอร์ของอิทธิพลสุ่ม r ตัว สำหรับกลุ่ม i ซึ่งเป็นอิสระต่อกันจากการแจกแจงแบบปกติ ที่มีค่าเฉลี่ย $\boldsymbol{0}$ และเมทริกซ์ความแปรปรวนร่วม $\boldsymbol{\Sigma}_u$ และ $\boldsymbol{z}'_{ij}$ คือเวกเตอร์ของตัวแปรอธิบายสำหรับอิทธิพลสุ่ม กำหนดให้ $\boldsymbol{\beta}^* = (\alpha_1,\ldots,\alpha_{c-1},\boldsymbol{\beta})$ ฟังก์ชันภาวะน่าจะเป็นแบบมีเงื่อนไขของกลุ่ม i คือ

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

$$f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta}) = \prod_{j=1}^{n_i}\prod_{k=1}^{c} \pi_k^u(\boldsymbol{x}_{ij})^{y_{ijk}}$$

โดยที่ $y_{ijk} = 1$ ถ้า $Y_{ij} = k$ และ $y_{ijk} = 0$ ถ้า $Y_{ij} \neq k$ และ

$$\pi_k^u(\boldsymbol{x}_{ij}) = P(Y_{ij} = k \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i) = P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i) - P(Y_{ij} \le k-1 \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i)$$

ฟังก์ชันภาวะน่าจะเป็นแบบมาร์จินัลของกลุ่ม i จะอยู่ในรูป

$$L_i(\boldsymbol{\beta}^*, \boldsymbol{\Sigma}_u) = \int_{\boldsymbol{u}_i} f(\boldsymbol{Y}_i \mid \boldsymbol{u}_i; \boldsymbol{\beta}^*) \times g(\boldsymbol{u}_i; \boldsymbol{\Sigma}_u) d\boldsymbol{u}_i$$

เมื่อ $g(\boldsymbol{u}_i; \boldsymbol{\Sigma}_u)$ คือฟังก์ชันความหนาแน่นของ $\boldsymbol{u}_i$ ดังนั้นฟังก์ชันภาวะน่าจะเป็นมีรูปแบบคือ

$$L(\boldsymbol{\beta}^*, \boldsymbol{\Sigma}_u) = \prod_{i=1}^{N} L_i(\boldsymbol{\beta}^*, \boldsymbol{\Sigma}_u)$$

จะเห็นว่าในการประมาณพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดสำหรับตัวแบบอิทธิพลสุ่มมีความยุ่งยากซับซ้อนเนื่องจากฟังก์ชันภาวะน่าจะเป็นติดเทอมของอินทิกรัล ทำให้การคำนวณยาก จึงต้องมีการประมาณค่าอินทิกรัลก่อนที่จะหาค่าสูงสุดของฟังก์ชัน ซึ่งมีวิธีการประมาณอินทิกรัลอยู่หลายวิธี

การประมาณพารามิเตอร์ของตัวแบบสำหรับค่าสังเกตที่มีความสัมพันธ์กันมีการพัฒนาอย่างช้า ๆ เมื่อเทียบกับตัวแบบสำหรับค่าสังเกตที่เป็นอิสระกัน เนื่องจากการคำนวณซับซ้อน จึงมีผู้เสนอวิธีประมาณพารามิเตอร์ของตัวแบบหลายวิธี ดังนี้ สำหรับตัวแบบมาร์จินัล Liang & Zeger (1986) เสนอวิธี Generalized Estimating Equations (GEE) ต่อมา Goldstein (1991) เสนอวิธี Marginal Quasi-Likelihood (MQL) สำหรับตัวแบบอิทธิพลสุ่ม Green (1987) เสนอวิธี Penalized Quasi-Likelihood (PQL) ต่อมา Liu & Pierce (1994) เสนอวิธี Gaussian Quadrature (GQ) ต่อมา Pinheiro & Bates (1995) ได้ปรับปรุงวิธี Gaussian Quadrature และเรียกว่า “Adaptive Gaussian Quadrature ” หรือ AGQ และมีหลายงานวิจัยที่สนใจเปรียบเทียบวิธีประมาณพารามิเตอร์เหล่านี้ เช่น Breslow & Clayton (1993) ทำการเปรียบเทียบวิธี MQL และ PQL สำหรับตัวแปรตอบสนองแบบทวิภาค โดยการจำลองแบบข้อมูล ผลการศึกษาพบว่าวิธี PQL และ MQL ให้ค่าประมาณพารามิเตอร์ที่ต่ำกว่าค่าจริง ต่อมา Tutz & Hennevogl (1996) ทำการเปรียบเทียบวิธีประมาณพารามิเตอร์ที่แตกต่างกันหลายวิธีของตัวแบบอิทธิพลสุ่ม สำหรับตัวแปรตอบสนองแบบอันดับ ประกอบด้วยวิธี GQ วิธี Monte Carlo EM Algorithm (MCEM) และวิธี EM Algorithm based on posterior mode โดยใช้ข้อมูลตัวอย่าง 2 ชุด ต่อมา Spiess & Hamerle (2000) ทำการจำลองข้อมูลตัวแปรตอบสนองแบบทวิภาค เพื่อเปรียบเทียบวิธี GQ และ GEE ซึ่งพบว่าวิธี GQ มีประสิทธิภาพสูงกว่าวิธี GEE เมื่อสหสัมพันธ์มีค่ามาก แต่มีประสิทธิภาพต่ำกว่าวิธี GEE เมื่อสหสัมพันธ์มีค่าน้อยและจำนวนกลุ่มมีขนาดเล็ก ต่อมา Kuss & McLerran (2007) แสดงการเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลที่มีอยู่ในโปรแกรม SAS® คือ GEE และ MQL และตัวแบบอิทธิพลสุ่มคือวิธี AGQ และ PQL โดยใช้ข้อมูลตัวอย่าง ผลการศึกษาพบว่าวิธี GEE และ MQL ให้ค่าประมาณที่ใกล้เคียงกัน ส่วนวิธี AGQ และ PQL ให้ค่าประมาณที่แตกต่างกันบ้าง และต่อมา Moineddin, Matheson & Glazier (2007) ได้ทำการจำลองข้อมูลเพื่อประเมินอิทธิพลของขนาดตัวอย่างที่มีต่อความถูกต้องของค่าประมาณพารามิเตอร์ของตัวแบบลอจิสติกทวิภาคด้วยวิธี AGQ กรณีที่ค่าตอบสนองมีความสัมพันธ์กัน นอกจากนี้ยังศึกษาอิทธิพลของสัดส่วนผลตอบสนองและสหสัมพันธ์ในกลุ่มด้วย ผลการศึกษาพบว่า จำนวนกลุ่มตัวอย่างส่งผลทางบวกต่อความถูกต้องของค่าประมาณพารามิเตอร์มากกว่าขนาดกลุ่มตัวอย่าง ส่วนสัดส่วนผลตอบสนองและสหสัมพันธ์ในกลุ่มส่งผลทางลบต่อความถูกต้องของค่าประมาณพารามิเตอร์

เนื่องจากงานวิจัยที่กล่าวมาเป็นการเปรียบเทียบวิธีประมาณพารามิเตอร์ สำหรับตัวแปรตอบสนองแบบอันดับโดยใช้ข้อมูลตัวอย่างเพียงหนึ่งหรือสองชุดเพียงเท่านั้น หรืองานวิจัยที่มีการจำลองข้อมูลก็เป็นการเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแปรตอบสนองแบบทวิภาคเท่านั้น งานวิจัยนี้จึงสนใจที่จะจำลองแบบข้อมูล กรณีที่ข้อมูลที่มีความสัมพันธ์กัน เพื่อเปรียบเทียบ

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

วิธีประมาณพารามิเตอร์ สำหรับตัวแปรตอบสนองแบบอันดับ โดยใช้ตัวแบบลอจิตสะสมแบบมาร์จินัล 2 วิธี ได้แก่ Generalized Estimating Equations (GEE) และ Marginal Quasi-Likelihood (MQL) และตัวแบบอิทธิพลสุ่ม 2 วิธี ได้แก่ Penalized Quasi-Likelihood (PQL) และ Adaptive Gaussian Quadrature (AGQ) แล้วนำมาเปรียบเทียบกับวิธีภาวะน่าจะเป็นสูงสุด (ML) สำหรับตัวแบบอิทธิพลคงที่ด้วย ซึ่งวิธีเหล่านี้เป็นวิธีที่นิยมใช้และมีอยู่ในโปรแกรม SAS® โดยพิจารณาประสิทธิภาพของวิธีประมาณพารามิเตอร์และความถูกต้องของค่าประมาณพารามิเตอร์จากความเอนเอียง (Bias) ความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และสัดส่วนของช่วงความเชื่อมั่น 95% ของค่าประมาณพารามิเตอร์ที่ครอบคลุมค่าจริง (Coverage) จากการจำลองแบบข้อมูล 1,000 ครั้ง

2. วัตถุประสงค์ของการศึกษา

งานวิจัยนี้มีวัตถุประสงค์ 3 ประการ ดังนี้

1. เพื่อเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบมาร์จินัล ระหว่างวิธี Generalized Estimating Equations (GEE) และ Marginal Quasi-Likelihood (MQL)

2. เพื่อเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบอิทธิพลสุ่ม ระหว่างวิธี Penalized Quasi-Likelihood (PQL) และ Adaptive Gaussian Quadrature (AGQ)

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

3. เพื่อศึกษาผลกระทบของขนาดตัวอย่าง (จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่าง) และสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ของตัวแบบมาร์จินัลและตัวแบบอิทธิพลสุ่ม สำหรับตัวแปรตอบสนองแบบอันดับ

3. สมมติฐานของการวิจัย

งานวิจัยนี้มีสมมติฐานของการวิจัย 3 ประการ ดังนี้

1. สำหรับตัวแบบมาร์จินัล วิธี Marginal Quasi-Likelihood (MQL) เป็นวิธีประมาณพารามิเตอร์ที่เหมาะสมกว่าวิธี Generalized Estimating Equations (GEE) สำหรับตัวแปรตอบสนองแบบอันดับ

2. สำหรับตัวแบบอิทธิพลสุ่ม วิธี Adaptive Gaussian Quadrature (AGQ) เป็นวิธีประมาณพารามิเตอร์ที่เหมาะสมกว่าวิธี Penalized Quasi-Likelihood (PQL) สำหรับตัวแปรตอบสนองแบบอันดับ

3. จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่างส่งผลกระทบทางบวก ส่วนสหสัมพันธ์ภายในกลุ่มส่งผลกระทบทางลบต่อความถูกต้องในการประมาณพารามิเตอร์ของตัวแบบอิทธิพลคงที่ ตัวแบบมาร์จินัล และตัวแบบอิทธิพลสุ่ม สำหรับตัวแปรตอบสนองแบบอันดับ

4. ขอบเขตของการวิจัย

งานวิจัยนี้ใช้การจำลองแบบข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของวิธีประมาณพารามิเตอร์ โดยขอบเขตการวิจัย ประกอบด้วยขอบเขตการจำลองแบบข้อมูล ขอบเขตวิธีประมาณพารามิเตอร์และขอบเขตการวัดประสิทธิภาพของวิธีประมาณพารามิเตอร์ ซึ่งมีรายละเอียดดังนี้

4.1 ขอบเขตการจำลองแบบข้อมูล

การจำลองแบบข้อมูลกำหนดเงื่อนไขของขนาดตัวอย่าง (จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่าง) สหสัมพันธ์ของข้อมูลในกลุ่ม ตัวแปรอธิบายและตัวแปรตอบสนอง ดังนี้

1. กำหนดจำนวนกลุ่มตัวอย่าง (N) เท่ากับ 50, 100 และ 250 ที่มีขนาดกลุ่มตัวอย่าง (n_i) เท่ากับ 5, 7 และ 10

2. กำหนดตัวแปรอธิบาย มีการแจกแจงแบบเบอร์นูลลี จำนวน 2 ตัวแปร ที่มีความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจเท่ากับ 0.50

3. กำหนดตัวแปรตอบสนองแบบอันดับที่มี 3 ระดับ

4. กำหนดตัวแบบในการวิเคราะห์ข้อมูลเป็นตัวแบบลอจิตสะสม ซึ่งมีสมการดังนี้

$$\text{logit}[\Pr(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, u_i)] = \alpha_k + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_i, \quad k = 1, \ldots, c-1.$$

ซึ่งมีพารามิเตอร์ที่ต้องประมาณค่า คือ $\alpha_1, \alpha_2, \beta_1, \beta_2$ และ σ_u^2

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

5. การสร้างตัวแปรแฝงแบบต่อเนื่อง (Y_{ij}^*) ตามสมการ

$$Y_{ij}^* = -\beta_1 x_{1ij} - \beta_2 x_{2ij} - u_i + \varepsilon_{ij}$$

โดยที่กำหนดค่าพารามิเตอร์ต่าง ๆ คล้ายคลึงกับงานวิจัยของ Spiess and Hamerle (2000) ดังนี้ β_1 = 0.5, β_2 = 0.5 อิทธิพลสุ่ม (u_i) สุ่มค่าอย่างอิสระกัน จากการแจกแจงแบบปกติ มีค่าเฉลี่ย 0 ความแปรปรวน σ_u^2 เท่ากับ 0.17, 0.82 และ 2.19 ซึ่งสอดคล้องกับสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05, 0.20 และ 0.40 ส่วนความคลาดเคลื่อน(ε_{ij}) สุ่มค่าอย่างเป็นอิสระกัน จากการแจกแจงแบบลอจิสติกมาตรฐาน ที่มีค่าเฉลี่ย 0 ความแปรปรวนเท่ากับ $\pi^2/3$

6. การแปลงค่าตัวแปรแฝงแบบต่อเนื่อง (Y_{ij}^*) ให้เป็นตัวแปรตอบสนองแบบอันดับ (Y_{ij}) ที่มี 3 ระดับ โดยกำหนดค่าจุดตัด α_1 = -1 และ α_2 = 1 (จะได้ความน่าจะเป็นของแต่ละระดับของตัวแปรตอบสนองแบบอันดับ ดังนี้ π_1 = 0.42, π_2 = 0.33, π_3 = 0.25) ตามความสัมพันธ์ดังนี้

$Y_{ij} = 1$ ถ้า $Y_{ij}^* \le \alpha_1$

$Y_{ij} = 2$ ถ้า $\alpha_1 < Y_{ij}^* \le \alpha_2$

$Y_{ij} = 3$ ถ้า $Y_{ij}^* > \alpha_2$

7. กำหนดรอบของการจำลองแบบข้อมูล จำนวน 1,000 ครั้ง

4.2 ขอบเขตวิธีประมาณพารามิเตอร์

วิธีประมาณพารามิเตอร์ที่นำมาพิจารณาในงานวิจัยนี้ ประกอบด้วยวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ ตัวแบบมาร์จินัลและตัวแบบอิทธิพลสุ่ม ดังนี้

4.2.1 วิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่

วิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ ที่นำมาพิจารณาในงานวิจัยนี้ คือ Maximum Likelihood (ML) Approach

4.2.2 วิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัล

1. Generalized Estimating Equations (GEE) Approach
2. Marginal Quasi-Likelihood (MQL) Approach

4.2.3 วิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่ม

1. Adaptive Guassian Quadrature (AGQ) Approach
2. Penalized Quasi-Likelihood (PQL) Approach

4.3 ขอบเขตการประเมินประสิทธิภาพของวิธีประมาณพารามิเตอร์

ในการประเมินประสิทธิภาพของวิธีประมาณพารามิเตอร์แต่ละวิธี จะพิจารณาจากความเอนเอียง ความคลาดเคลื่อนกำลังสองเฉลี่ยและ *Coverage* จากการจำลองแบบข้อมูล *M* ครั้ง ซึ่งมีรายละเอียดดังต่อไปนี้

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

4.3.1 ความเอนเอียง

ความเอนเอียง (*Bias*) ของค่าประมาณพารามิเตอร์ มีสูตรการคำนวณดังนี้

$$Bias = \frac{\sum_{m=1}^{M} \hat{\theta}_m}{M} - \theta$$

เมื่อ $\hat{\theta}_m$ แทนค่าประมาณพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อมูลจำลองในครั้งที่ m, $m = 1, \ldots, M$ ซึ่งวิธีประมาณพารามิเตอร์ที่ให้ค่า *Bias* น้อยกว่า จะเป็นวิธีที่ให้ค่าประมาณพารามิเตอร์ที่มีความถูกต้องมากกว่า

4.3.2 ความคลาดเคลื่อนกำลังสองเฉลี่ย

ความคลาดเคลื่อนกำลังสองเฉลี่ย (*MSE*) มีสูตรการคำนวณดังนี้

$$MSE = \frac{\sum_{m=1}^{M} (\hat{\theta}_m - \theta)^2}{M}$$

โดยที่วิธีการประมาณพารามิเตอร์ที่ให้ค่า *MSE* น้อยกว่า จะเป็นวิธีที่ให้ค่าประมาณพารามิเตอร์ที่มีความถูกต้องมากกว่า

4.3.3 *Coverage*

Coverage คือ สัดส่วนของช่วงความเชื่อมั่น ซึ่งงานวิจัยนี้ศึกษาที่ระดับ 95% ที่คำนวณจากค่าประมาณพารามิเตอร์จากการจำลองแบบข้อมูลทั้งหมด M ครั้ง ที่ครอบคลุมค่าพารามิเตอร์จริง วิธีการประมาณพารามิเตอร์ของตัวแบบที่ให้ค่า *Coverage* มากกว่า จะเป็นวิธีที่ให้ค่าประมาณพารามิเตอร์ที่มีความถูกต้องมากกว่า สูตรการคำนวณช่วงความเชื่อมั่น 95 % ของค่าประมาณพารามิเตอร์ สำหรับข้อมูลจำลองแบบครั้งที่ m มีดังนี้

$$\hat{\theta}_m \pm t_{0.025,df} \times s.e.(\hat{\theta}_m)$$

เมื่อ $t_{0.025,df}$ คือเปอร์เซ็นต์ไทล์ส่วนบนที่ 97.5% ของการแจกแจงแบบ *Student t*, *df* คือองศาความเป็นอิสระของการแจกแจงแบบ *Student t* ซึ่งคำนวณจากจำนวนกลุ่มตัวอย่าง (N) ลบด้วยจำนวนพารามิเตอร์ในตัวแบบที่ต้องประมาณค่า และ $s.e.(\hat{\theta}_m)$ คือความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์สำหรับข้อมูลจำลองแบบครั้งที่ m

5. นิยามศัพท์เฉพาะ

1. ตัวแปรตอบสนองแบบอันดับ หมายถึง ตัวแปรตอบสนองที่มีค่าวัดแบ่งเป็นหลายกลุ่มแบบเรียงลำดับได้ เช่นระดับอาการของโรค (น้อย, ปานกลาง, รุนแรง)

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2. ข้อมูลที่มีความสัมพันธ์กัน หมายถึง ข้อมูลที่มีการวัดซ้ำภายในหน่วยตัวอย่างเดียวกัน หรือข้อมูลของหน่วยตัวอย่างย่อยที่ซ้อนอยู่ภายในกลุ่มเดียวกัน

6. ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นแนวทางในการเลือกใช้วิธีประมาณพารามิเตอร์ที่เหมาะสมในการวิเคราะห์ข้อมูลด้วยตัวแบบมาร์จินัล สำหรับตัวแปรตอบสนองแบบอันดับ

2. เป็นแนวทางในการเลือกใช้วิธีประมาณพารามิเตอร์ที่เหมาะสมในการวิเคราะห์ข้อมูลด้วยตัวแบบอิทธิพลสุ่ม สำหรับตัวแปรตอบสนองแบบอันดับ

3. สามารถเลือกขนาดตัวอย่าง (จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่าง) ที่เหมาะสมกับสหสัมพันธ์ของข้อมูลภายในกลุ่มและวิธีประมาณพารามิเตอร์ที่เลือกใช้ได้อย่างเหมาะสม

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

1. ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องสำหรับการเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ ประกอบด้วยทฤษฎีเกี่ยวกับตัวแบบเชิงเส้นนัยทั่วไปกรณีของตัวแบบอิทธิพลคงที่กับตัวแบบมาร์จินัลพร้อมทั้งการประมาณพารามิเตอร์ของตัวแบบ และตัวแบบแบบผสมเชิงเส้นนัยทั่วไปกรณีของตัวแบบอิทธิพลสุ่มพร้อมทั้งการประมาณพารามิเตอร์ของตัวแบบ ซึ่งมีรายละเอียดดังต่อไปนี้

1.1 ตัวแบบเชิงเส้นนัยทั่วไป

ตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Models: GLMs) เป็นกลุ่มของตัวแบบอิทธิพลคงที่ สำหรับตัวแปรตอบสนองแบบต่าง ๆ (ต่อเนื่องและไม่ต่อเนื่อง) โดยทั่วไปจะประกอบด้วยส่วนประกอบ 3 ส่วน ส่วนแรกคือตัวแปรตอบสนอง ที่มีค่าเฉลี่ย $\mu_i = E(Y_i)$ และความแปรปรวน $var(Y_i)$ ซึ่งสมมติว่ามีการแจกแจงอยู่ในวงศ์ชี้กำลัง (Agresti 2002 : 116) มีฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(y_i;\theta_i,\phi) = \exp\left\{\left[y_i\theta_i - b(\theta_i)\right] / a(\phi) + c(y_i,\phi)\right\}$$

เมื่อ θ_i คือ Natural parameter และ ϕ คือ Dispersion parameter ส่วนที่สองคือคือตัวพยากรณ์เชิงเส้น แทนด้วย η_i ซึ่งมีรูปแบบ ดังนี้

$$\eta_i = \boldsymbol{x}_i'\boldsymbol{\beta}$$

เมื่อ $\boldsymbol{x}_i$ คือเวกเตอร์ของตัวแปรอธิบายสำหรับหน่วยที่ i ด้วยอิทธิพลคงที่ $\boldsymbol{\beta}$ และส่วนที่สาม คือ link function $g(\cdot)$ เป็นฟังก์ชันที่แสดงความสัมพันธ์ระหว่างค่าเฉลี่ยของตัวแปรตอบสนอง $E(Y_i)$ กับตัวพยากรณ์เชิงเส้น η_i ดังนี้

$$g(\mu_i) = \eta_i$$

The link function $g(\mu_i) = \mu_i$ เรียกว่า Identity link ที่กำหนดตัวแบบเชิงเส้นสำหรับค่าเฉลี่ยของตัวมันเอง ซึ่งเป็น link function สำหรับการถดถอยเชิงเส้นของตัวแปรตอบสนองที่มีการแจกแจงแบบปกติ นอกจากนี้ยังมี link function สำหรับตัวแปรตอบสนองแบบต่าง ๆ ดังตารางที่ 1

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 1 ชนิดของตัวแบบเชิงเส้นนัยทั่วไป สำหรับการวิเคราะห์เชิงสถิติ

ตัวแปรตอบสนอง	ตัวแปรอธิบาย	Link function	ตัวแบบ
ปกติ	ต่อเนื่อง	Identity	Regression
ปกติ	เชิงกลุ่ม	Identity	ANOVA
ปกติ	ผสม	Identity	ANCOVA
ทวิภาค	ผสม	Logit	Logistic
ปัวซง	ผสม	Log	Loglinear
มัลติโนเมียล	ผสม	Generalized logit	Multinomial response

ที่มา : Alan Agresti, Categorical data analysis (NewYork : John Wiley & Sons, 2002), 118.

1.1.1 ตัวแบบอิทธิพลคงที่และการประมาณพารามิเตอร์ของตัวแบบ

จากข้อมูลที่ประกอบด้วยกลุ่มซึ่งเป็นอิสระกัน N กลุ่ม โดยที่กลุ่ม i ประกอบด้วยค่าสังเกต n_i ค่าสังเกต กำหนดให้ Y_{ij} แทนค่าตอบสนองที่ $j, j = 1,..., n_i$ ในกลุ่ม $i, i = 1,..., N$ ซึ่งมีค่าตกอยู่ในระดับที่ $k, k = 1, 2,..., c$ เมื่อ c แทนจำนวนระดับของตัวแปรตอบสนอง และ $\boldsymbol{x}_{ij} = (x_{1ij},...,x_{pij})'$ แทนเวกเตอร์ของตัวแปรอธิบาย p ตัว สำหรับค่าตอบสนองที่ j ในกลุ่ม i ตัวแบบที่นิยมใช้สำหรับตัวแปรตอบสนองแบบอันดับจะใช้ link function คือ logit link โดยสร้างลอจิตของความน่าจะเป็นสะสม จึงเรียกว่าตัวแบบลอจิตสะสม กรณีของตัวแบบอิทธิพลคงที่ซึ่งสมมติว่าค่าตอบสนองเป็นอิสระกัน จะอยู่ในรูป

$$\text{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij})] = \eta_{ijk} = \alpha_k + \boldsymbol{x}'_{ij}\boldsymbol{\beta}, \quad k = 1,\ldots,c-1 \tag{1}$$

โดยที่ $\alpha_k, -\infty = \alpha_0 < \alpha_1 < \alpha_2 < ... < \alpha_{c-1} < \alpha_c = \infty$ คือค่าเริ่มต้นหรือจุดตัดของระดับที่ k ของตัวแปรตอบสนอง จำนวน $c - 1$ ค่า และ $\boldsymbol{\beta}$ คือเวกเตอร์พารามิเตอร์สำหรับอิทธิพลคงที่ของตัวแปรอธิบาย p ตัว ตามที่สัมประสิทธิ์ของตัวแบบ $\boldsymbol{\beta}$ ไม่เปลี่ยนตาม k นั่นคืออิทธิพลของตัวแปรอธิบายไม่เปลี่ยนแปลงตามระดับของตัวแปรตอบสนอง McCullagh (1980) เรียกข้อสมมติของความเท่ากันของอัตราส่วนออดส์ระหว่าง $c - 1$ ลอจิตสะสมนี้ว่า “Proportional odds assumption” ภายใต้แนวคิดที่ว่าตัวแปรตอบสนองแบบอันดับมาจากการจัดกลุ่มของตัวแปรแฝงแบบต่อเนื่อง Y^* ซึ่งเป็นตัวแปรที่ไม่ได้สังเกตค่า โดยใช้จุดตัด $\alpha_k, k = 1, ..., c - 1$ ดังนี้ $Y_{ij} = k$ เมื่อ $\alpha_{k-1} < Y^*_{ij} \le \alpha_k$ โดยที่ Y^* มีความสัมพันธ์กับตัวแปรอธิบาย x ดังนี้

$$Y^*_{ij} = -\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \varepsilon_{ij}$$

เมื่อ ε_{ij} คือความคลาดเคลื่อนของตัวแปรแฝง และ ดังนั้นความน่าจะเป็นสะสมสามารถหาได้จาก

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

$$P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}) = P(Y_{ij}^* \le \alpha_k \mid \boldsymbol{x}_{ij}) = P(-\boldsymbol{x}_{ij}'\boldsymbol{\beta} + \varepsilon_{ij} \le \alpha_k) = P(\varepsilon_{ij} \le \alpha_k + \boldsymbol{x}_{ij}'\boldsymbol{\beta}) = F(\eta_{ijk})$$

สำหรับตัวแบบลอจิตสะสม กำหนดให้ ε_{ij} มีการแจกแจงแบบลอจิสติกมาตรฐาน (ค่าเฉลี่ย 0 และความแปรปรวน $\pi^2/3$) และมีฟังก์ชันการแจกแจง (F) คือ

$$F(\eta_{ijk}) = 1/[1 + exp(-\eta_{ijk})]$$

โดยที่ $F(\eta_{ij0}) = 0$ และ $F(\eta_{ijc}) = 1$ ดังนั้น $var(Y_{ij}^*) = var(\varepsilon_{ij}) = \sigma_\varepsilon^2 = \dfrac{\pi^2}{3}$ สำหรับตัวแบบโพรบิต จะขึ้นอยู่กับการแจกแจงแบบปกติมาตรฐานของตัวแปรแฝง มักจะถูกเสนอขึ้นเป็นทางเลือกของตัวแบบลอจิสติก ดังนั้นความน่าจะเป็นที่ Y_{ij} ตกอยู่ในระดับ k หาได้จาก

$$\pi_k(\boldsymbol{x}_{ij}) = P(Y_{ij} = k \mid \boldsymbol{x}_{ij}) = F(\eta_{ijk}) - F(\eta_{ij,k-1})$$

การประมาณพารามิเตอร์ของตัวแบบอิทธิพลคงที่นี้ใช้วิธีภาวะน่าจะเป็นสูงสุด สำหรับตัวแปรตอบสนองแบบอันดับ ฟังก์ชันภาวะน่าจะเป็น คือ

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\prod_{k=1}^{c} \pi_k(\boldsymbol{x}_{ij})^{y_{ijk}}$$

โดยที่ $y_{ijk} = 1$ ถ้า $Y_{ij} = k$ และ $y_{ijk} = 0$ ถ้า $Y_{ij} \ne k$, The log-likelihood function คือ

$$l(\boldsymbol{\beta}) = \log L(\boldsymbol{\beta}) = \sum_{i=1}^{N}\sum_{j=1}^{n_i}\sum_{k=1}^{c} y_{ijk} \log \pi_k(\boldsymbol{x}_{ij})$$

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ค่าประมาณเวกเตอร์พารามิเตอร์ $\boldsymbol{\beta}$ (รวมถึง α_k) คือผลลัพธ์ของระบบสมการภาวะน่าจะเป็น

$$u(\boldsymbol{\beta}) = \frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{N} \boldsymbol{D}_i' \boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{\pi}_i) = \boldsymbol{0}$$

เมื่อ $\boldsymbol{\pi}_i = g^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta})$ คือเวกเตอร์ของความน่าจะเป็นที่สัมพันธ์กับ $\boldsymbol{Y}_i$, $\boldsymbol{D}_i$ คือเมทริกซ์ของอนุพันธ์ ที่มีสมาชิกคือ $\dfrac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\beta}} = \dfrac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\eta}_i}\dfrac{\partial \boldsymbol{\eta}_i}{\partial \boldsymbol{\beta}} = \dfrac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\eta}_i}\boldsymbol{x}_i$ และ $\boldsymbol{V}_i$ คือเมทริกซ์ทแยงมุมที่มีสมาชิกคือ $var(Y_{ij})$

เนื่องจากระบบสมการภาวะน่าจะเป็นมีลักษณะที่ไม่เป็นเชิงเส้นในเทอมของพารามิเตอร์ ทำให้ไม่สามารถแก้ระบบสมการได้โดยวิธีการปกติ จึงต้องใช้การย้อนซ้ำเพื่อหาคำตอบซึ่งมีอยู่หลายวิธี เช่น วิธี Newton-Raphson (Agresti 2002 : 144) ซึ่งมีแนวคิดจากการประมาณ $L(\boldsymbol{\beta})$ รอบ ๆ $\boldsymbol{\beta}^{(s)}$ เมื่อ $\boldsymbol{\beta}^{(s)}$ คือเวกเตอร์ $\boldsymbol{\beta}$ ในการย้อนซ้ำรอบที่ s, $s = 1, 2, \ldots$ โดยใช้การกระจายอนุกรมเทเลอร์อันดับสอง ดังนี้

$$L(\boldsymbol{\beta}) \approx L(\boldsymbol{\beta}^{(s)}) + u^{(s)}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(s)}) + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(s)})' H^{(s)}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(s)})$$

เมื่อ $u = \dfrac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$ และ $H = Hessian\ matrix$ ที่มีสมาชิกคือ $h_{ab} = \dfrac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_a \boldsymbol{\beta}_b}$ และ

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \approx u^{(s)} + H^{(s)}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(s)}) = 0$$

แล้วจะได้ว่า

$$\boldsymbol{\beta} = \boldsymbol{\beta}^{(s)} - \left(H^{(s)}\right)^{-1} u^{(s)}$$

โดยที่ $\boldsymbol{\beta}$ คือค่าเวกเตอร์ $\boldsymbol{\beta}$ ในรอบต่อไป หรือ $\boldsymbol{\beta}^{(s+1)}$ นั่นเอง การย้อนซ้ำหาคำตอบเวกเตอร์ $\hat{\boldsymbol{\beta}}$ ที่ทำให้ $L(\boldsymbol{\beta})$ มีค่าสูงสุดมีขั้นตอนดังนี้

ขั้นตอนที่ 1 เดาตำตอบเริ่มต้น ที่คาดว่าทำให้ $L(\boldsymbol{\beta})$ มีค่าสูงสุด แทนด้วย $\boldsymbol{\beta}^{(0)}$

ขั้นตอนที่ 2 คำนวณ $u^{(0)} = \left.\dfrac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(0)}}$ และ $H^{(0)}$ ที่มีสมาชิกคือ $h_{ab}^{(0)} = \left.\dfrac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \beta_b}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(0)}}$

ขั้นตอนที่ 3 คำนวณหาคำตอบเริ่มต้นใหม่

$$\boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}^{(0)} - \left(H^{(0)}\right)^{-1} u^{(0)}$$

$$\vdots$$

$$\boldsymbol{\beta}^{(s+1)} = \boldsymbol{\beta}^{(s)} - \left(H^{(s)}\right)^{-1} u^{(s)}$$

ขั้นตอนที่ 4 ทำการย้อนซ้ำไปเรื่อย ๆ จนกว่า $|\boldsymbol{\beta}^{(s+1)} - \boldsymbol{\beta}^{(s)}|$ มีค่าน้อยกว่าค่าที่กำหนด เช่น 0.0001 จะได้ว่า $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}^{(s+1)}$ และ $cov(\hat{\boldsymbol{\beta}}) = J^{-1}$ เมื่อ $J = E\left(-\dfrac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \beta_b}\right) = E(-H)$ เรียกว่า Information matrix จะได้ว่าวิธีการย้อนซ้ำอีกวิธีหนึ่งคือ Fisher – scoring จะคำนวณหาคำตอบเริ่มต้นใหม่ ในขั้นที่ 3 จากสมการ

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

$$\boldsymbol{\beta}^{(s+1)} = \boldsymbol{\beta}^{(s)} + \left(J^{(s)}\right)^{-1} u^{(s)}$$

การวิเคราะห์ข้อมูลที่ค่าตอบสนองมีความสัมพันธ์กัน โดยใช้ตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Models: GLMs) ภายใต้ข้อสมมติว่าค่าสังเกตแต่ละค่าเป็นอิสระกันนั้น อาจจะให้ผลสรุปที่ผิดพลาดได้ (Agresti 2002 : 491) ทางเลือกของปัญหานี้จึงมีกลุ่มของตัวแบบส่วนใหญ่ที่ใช้ในการวิเคราะห์ข้อมูลที่ค่าตอบสนองมีความสัมพันธ์กันอยู่ 2 ตัวแบบ คือตัวแบบมาร์จินัล หรือ Population-averaged model และตัวแบบอิทธิพลสุ่ม หรือ Clustered-specific model (หรือบางทีเรียกว่า Subject-specific model) ตัวแบบอิทธิพลสุ่มจะอ้างอิงการแจกแจงแบบมีเงื่อนไขในแต่ละกลุ่ม ส่วนตัวแบบมาร์จินัลจะอ้างอิงการแจกแจงแบบมาร์จินัลหรือเฉลี่ยทุก ๆ กลุ่มในประชากร การเลือกใช้ตัวแบบใดก็ขึ้นอยู่กับว่าผู้วิจัยต้องการตีความหมายของพารามิเตอร์ในระดับกลุ่มหรือระดับประชากร

1.1.2 ตัวแบบมาร์จินัลและการประมาณพารามิเตอร์ของตัวแบบ

ตัวแบบมาร์จินัลจะอ้างอิงการแจกแจงแบบมาร์จินัลหรือเฉลี่ยทุก ๆ กลุ่มในประชากรจึงเป็นการตีความหมายพารามิเตอร์ในระดับประชากร ซึ่งใช้กันอย่างแพร่หลายในการวิจัยด้านการแพทย์ที่เกี่ยวกับโรคระบาด ซึ่งสนใจความถี่รวมของการเกิดโรคในประชากร (Agresti,

2002) ในการกำหนดตัวแบบลอจิตสะสมให้เป็นตัวแบบมาร์จินัลจะมีรูปแบบของสมการเหมือนกับตัวแบบลอจิตสะสม สำหรับอิทธิพลคงที่ ดังสมการ (1) แต่จะพิจารณาความสัมพันธ์ระหว่างค่าตอบสนองในกลุ่มไว้ในเมทริกซ์ความแปรปรวนร่วมของตัวแปรตอบสนอง จึงทำให้การประมาณพารามิเตอร์ของตัวแบบ คำนวณยากขึ้น ซึ่งวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลที่นิยมใช้มีดังนี้

1.1.2.1 Generalized Estimating Equations Approach

Generalized Estimating Equations (GEE) เสนอโดย Liang & Zeger (1986) ซึ่งเป็นวิธีประมาณพารามิเตอร์ที่ใช้กันอย่างแพร่หลายวิธีหนึ่งสำหรับตัวแบบมาร์จินัลเมื่อค่าตอบสนองมีความสัมพันธ์กัน โดยวิธีนี้จะหลีกเลี่ยงข้อสมมติเกี่ยวกับการแจกแจงและโครงสร้างสหสัมพันธ์ของค่าตอบสนองในกลุ่มซึ่งเป็นที่ต้องการสำหรับฟังก์ชันภาวะน่าจะเป็น โดยตัวประมาณที่ได้ไม่ใช่ตัวประมาณภาวะน่าจะเป็นสูงสุด แต่เป็นตัวประมาณที่คงเส้นคงวาและมีการแจกแจงแบบปกติเมื่อใกล้อนันต์ ถึงแม้ว่าจะกำหนดโครงสร้างสหสัมพันธ์ของข้อมูลในกลุ่มไม่ถูกต้อง (Liang & Zeger, 1986) เริ่มต้นจะทำการแปลงตัวแปรตอบสนองแบบอันดับเป็นตัวแปรชี้วัดแบบทวิภาค $c-1$ ตัวแปร โดยให้ $y_{ijk}=1$ ถ้าค่าสังเกตที่ $\underline{i}$ ในกลุ่มที่ $\underline{i}$ มีค่าตอบสนองเป็น k , $k = 1, \ldots, c-1$, ระบบสมการสำหรับการประมาณ $\boldsymbol{\beta}$ (รวมถึงจุดตัด α_k) เรียกว่า The Generalized Estimating Equationss อยู่ในรูป

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

$$u(\boldsymbol{\beta}) = \sum_{i=1}^{N} \boldsymbol{D}_i' \boldsymbol{V}_i^{-1} [\boldsymbol{y}_i - \pi_i] = \boldsymbol{0}$$

โดยที่ $\boldsymbol{y}_i = (\boldsymbol{y}_{i1}', \ldots, \boldsymbol{y}_{in_i}')$ คือเวกเตอร์ของค่าตอบสนองสำหรับกลุ่มที่ i, $\boldsymbol{\pi}_i$ คือเวกเตอร์ของความน่าจะเป็นที่สัมพันธ์กับ $\boldsymbol{y}_i$ และ $\boldsymbol{D}_i = \partial \pi_i / \partial \boldsymbol{\beta}$ และ $\boldsymbol{V}_i = \boldsymbol{A}_i^{1/2} \boldsymbol{R}_i(\alpha) \boldsymbol{A}_i^{1/2}$ คือ Working covariance matrix เมื่อ $\boldsymbol{A}_i$ คือเมทริกซ์ทแยงมุมที่มีสมาชิกแนวทแยงมุมคือ $var(\boldsymbol{y}_i)$ และ และ $\boldsymbol{R}_i(\alpha) = \boldsymbol{C}(\alpha) \otimes \boldsymbol{S}$ เรียกว่า Working correlation matrix เมื่อ $\boldsymbol{C}(\alpha)$ คือเมทริกซ์ขนาด $n_i \times n_i$ ของสหสัมพันธ์ระหว่างค่าตอบสนองในกลุ่ม และ $\boldsymbol{S}$ คือเมทริกซ์ขนาด $(c-1) \times (c-1)$ ของสหสัมพันธ์ระหว่างตัวแปรชี้วัดแบบทวิภาค จำนวน $c-1$ ตัวแปร ภายในแต่ละค่าสังเกต ดังนี้

$$\boldsymbol{S} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1(c-1)} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2(c-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{(c-1)1} & \rho_{(c-1)2} & \cdots & \rho_{(c-1)(c-1)} \end{bmatrix}$$

โดยที่สำหรับตัวแบบลอจิตสะสม คาดว่าสหสัมพันธ์ระหว่างตัวแปรชี้วัดแบบทวิภาค y_{ijk} และ $y_{ijk'}$ นิยามโดย $\rho_{kk'} = \rho_{k'k} = [exp(\alpha_k - \alpha_{k'})]^{1/2}$ เมื่อ $k < k'$ (Kenward et al., 1994) โครงสร้างสหสัมพันธ์ระหว่างค่าตอบสนองในกลุ่ม $\boldsymbol{C}(\alpha)$ ที่ใช้กันอย่างกว้างขวาง มีดังนี้

1. Independence:

$$\boldsymbol{C}(\alpha)=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & 1\end{bmatrix}$$

2. Exchangeable:

$$\boldsymbol{C}(\alpha)=\begin{bmatrix}1 & \alpha & \cdots & \alpha\\ \alpha & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \alpha\\ \alpha & \cdots & \alpha & 1\end{bmatrix}$$

3. Autoregressive (AR) เช่น First- order Autoregressive (AR1):

$$\boldsymbol{C}(\alpha)=\begin{bmatrix}1 & \alpha^{d_{12}} & \cdots & \alpha^{d_{1n_i}}\\ \alpha^{d_{21}} & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \alpha^{d_{(n_i-1),n_i}}\\ \alpha^{d_{n_i1}} & \cdots & \alpha^{d_{n_i,(n_i-1)}} & 1\end{bmatrix}$$

เมื่อ $d_{jj'}=d_{j'j}=|j-j'|$

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

4. M-dependent เช่น 1-dependent :

$$\boldsymbol{C}(\alpha)=\begin{bmatrix}1 & \alpha & \cdots & 0\\ \alpha & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \alpha\\ 0 & \cdots & \alpha & 1\end{bmatrix}$$

5. Unstructed : ไม่มีข้อกำหนดเกี่ยวกับโครงสร้างสหสัมพันธ์

โดยที่โครงสร้างสหสัมพันธ์ไม่ส่งผลต่อค่าประมาณพารามิเตอร์มากนัก (Parson et al. 2009)

Lipsitz et al. (1994) แนะนำให้ใช้วิธี Fisher - scoring ในการหาผลลัพธ์ของระบบสมการภาวะน่าจะเป็นข้างต้นสำหรับ $\boldsymbol{\beta}$ โดยใช้วิธีโมเมนต์ในการปรับปรุงการประมาณสหสัมพันธ์ $\rho_{kk'}$ ในแต่ละรอบของการย้อนซ้ำ และจะได้ว่าเมทริกซ์ความแปรปรวนร่วมของ $\hat{\boldsymbol{\beta}}$ คือ

$$cov(\hat{\boldsymbol{\beta}})=\left[\sum_{i=1}^{N}\boldsymbol{D}_i'\boldsymbol{V}_i^{-1}\boldsymbol{D}_i\right]^{-1}\left[\sum_{i=1}^{N}\boldsymbol{D}_i'\boldsymbol{V}_i^{-1}cov(\boldsymbol{y}_i)\boldsymbol{V}_i^{-1}\boldsymbol{D}_i\right]\left[\sum_{i=1}^{N}\boldsymbol{D}_i'\boldsymbol{V}_i^{-1}\boldsymbol{D}_i\right]^{-1}$$

ประมาณขึ้นโดยการแทน $\hat{\pi}_i$ จากตัวแบบที่เหมาะสม และแทน $cov(\boldsymbol{y}_i)$ ด้วยเมทริกซ์ความแปรปรวนร่วมของ $\boldsymbol{y}_i$

1.1.2.2 Marginal Quasi-Likelihood Approach

Marginal Quasi-Likelihood (MQL) เสนอโดย Goldstein (1991) ในการประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัล เนื่องจากจุดสำคัญของตัวแบบสำหรับข้อมูลมีความสัมพันธ์กันคือโครงสร้างของตัวแบบอยู่ในรูปค่าเฉลี่ยแบบมีเงื่อนไขของตัวแปรตอบสนอง

$$\mu_{ij}^{u} = E(Y_{ij} \mid \boldsymbol{u}_i) = g^{-1}(\boldsymbol{x}_{ij}'\boldsymbol{\beta} + \boldsymbol{z}_{ij}'\boldsymbol{u}_{\mathrm{i}})$$

เมื่อแยกอิทธิพลสุ่มในแต่ละกลุ่ม นั่นหมายความว่า $\boldsymbol{\beta}$ แทนอิทธิพลของตัวแปรอธิบายที่แต่ละระดับของกลุ่ม ในการประมาณอิทธิพลของตัวแปรอธิบายโดยเฉลี่ยในประชากร จึงมีความเหมาะสมมากกว่า การกำหนดตัวแบบเชิงเส้นนัยทั่วไปในเทอมของค่าเฉลี่ยแบบมาร์จินัล เป็นดังนี้

$$\boldsymbol{\mu}_i = E(\boldsymbol{Y}_i) = g^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta})$$

การเขียนตัวแบบในรูป $\boldsymbol{Y}_i = \boldsymbol{\mu}_i^u + \boldsymbol{\varepsilon}_i$ โดยที่ $var(\boldsymbol{\varepsilon}_i) = \boldsymbol{a}_i v(\boldsymbol{\mu}_i^u)$ และ $\boldsymbol{u} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_u)$ เมื่อ $\boldsymbol{a}_i$ คือเวกเตอร์ค่าคงที่ และ $v(\boldsymbol{\mu}_i^u)$ คือฟังก์ชันความแปรปรวนในเทอมของค่าเฉลี่ยแบบมีเงื่อนไข จะได้ว่า $\boldsymbol{Y}_i \approx h(\boldsymbol{x}_i'\boldsymbol{\beta}) + h'(\boldsymbol{x}_i'\boldsymbol{\beta})\boldsymbol{z}_i'\boldsymbol{u}_i + \boldsymbol{\varepsilon}_i$ (Goldstein 1991) เมื่อ $h = g^{-1}$ นิยาม $\boldsymbol{V}_0$ และ $\boldsymbol{D}$ เป็นเมทริกซ์ทแยงมุมที่มีสมาชิกแนวทแยงมุมคือ $\boldsymbol{a}_i v(\boldsymbol{\mu}_i^u)$ และ ตามลำดับ $\partial\boldsymbol{\mu}_i / \partial\boldsymbol{\beta}$ จะได้ว่า

$$var(\boldsymbol{Y}) = \boldsymbol{V}_0 + \boldsymbol{D}^{-1}\boldsymbol{z}\,\boldsymbol{\Sigma}_u\boldsymbol{z}\,\boldsymbol{D}^{-1}$$

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตัวแบบมาร์จินัลของรูปแบบนี้ ตรวจสอบโดย Zeger et al. (1988) ซึ่งได้แสดงว่าค่าเฉลี่ยแบบมาร์จินัลสำหรับข้อมูลมีความสัมพันธ์กันซึ่งอิทธิพลสุ่มมีการแจกแจงแบบปกติมักจะกำหนดในรูป

$$\mu_{ij}^{u} = E(Y_{ij} \mid \boldsymbol{u}_i) = h(\boldsymbol{x}_{ij}'\boldsymbol{\beta} + \boldsymbol{z}_{ij}'\boldsymbol{u}_{\mathrm{i}})$$

สำหรับตัวแปรตอบสนองแบบอันดับจะใช้ logit link function โดยที่

$$E(\boldsymbol{Y}_i) \approx \frac{1}{1 + exp(-c_i\boldsymbol{x}_i'\boldsymbol{\beta})}$$

โดยที่ $c_i = | c^2\boldsymbol{\Sigma}_u\boldsymbol{z}_i\boldsymbol{z}_i' + I |^{-1/2} = (1 + c^2\boldsymbol{z}_i'\boldsymbol{\Sigma}_u\boldsymbol{z}_i)^{-1/2}$ และ $c = 16\sqrt{3}/(15\pi)$ โดยได้ประมาณความแปรปรวนของตัวแปรตอบสนองโดยใช้ Working covariance matrix ในกระบวนการประมาณแบบย้อนซ้ำ แต่อยู่บนพื้นฐานของเมทริกซ์ความแปรปรวนร่วมที่ได้มาจากสมการปกติ

การประมาณพารามิเตอร์ $\boldsymbol{\beta}$ (รวมถึงจุดตัด α_k) ในตัวแบบมาร์จินัลโดยใช้ Quasi-likelihood equation ที่เหมาะสมสำหรับตัวแปรตอบสนอง การแทนเวกเตอร์ค่าเฉลี่ยแบบมาร์จินัลด้วย $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_{n_i})'$ สมการปกติหรือ The estimating equations คือ

$$u(\boldsymbol{\beta}) = \frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\beta}} var(\boldsymbol{y})(\boldsymbol{y} - \boldsymbol{\mu}) = \boldsymbol{0}$$

ซึ่งเขียนได้ในรูป

$$\boldsymbol{X'(DV_0D + z\,\Sigma_u z')^{-1}\,D(y - \mu) = 0}$$

ด้วย $\boldsymbol{\eta} = (\eta_1, ..., \eta_{n_i})'$ แทนเวกเตอร์ของตัวพยากรณ์เชิงเส้น $\eta_i = \boldsymbol{x}_i'\boldsymbol{\beta}$ กระบวนการ Fisher-scoring นำไปสู่กระบวนการย้อนซ้ำแบบถ่วงน้ำหนักของเวกเตอร์ $\boldsymbol{Y} = \boldsymbol{\eta} + \boldsymbol{D}(\boldsymbol{y} - \boldsymbol{\mu})$ บน $\boldsymbol{x}$ ด้วยเมทริกซ์น้ำหนัก

$$\boldsymbol{V}^{-1} = (\boldsymbol{D}\boldsymbol{V}_0\boldsymbol{D} + \boldsymbol{z}\,\boldsymbol{\Sigma}_u\boldsymbol{z}')^{-1}$$

สำหรับแต่ละขั้นตอนของการย้อนซ้ำ ปัญหาก็คือความสมมูลของการประมาณ $\boldsymbol{\beta}$ ที่สัมพันธ์กับตัวแบบของตัวแปรตอบสนองที่มีการแจกแจงแบบปกติ $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{zu} + \boldsymbol{\varepsilon}$ ด้วย $\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \boldsymbol{W}^{-1})$ เมื่อ $\boldsymbol{W} = \boldsymbol{D}\boldsymbol{V}_0\boldsymbol{D}$ และ $\boldsymbol{u} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_u)$ ซึ่งสามารถประมาณได้จาก $\hat{\boldsymbol{u}} = \boldsymbol{\Sigma}_u\boldsymbol{z}\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{x}'\boldsymbol{\beta})$

1.2 ตัวแบบผสมเชิงเส้นนัยทั่วไป

ตัวแบบผสมเชิงเส้นนัยทั่วไป (Generalized Linear Mixed Models: GLMMs) เป็นตัวแบบที่ขยายจากตัวแบบเชิงเส้นนัยทั่วไป โดยการรวมอิทธิพลสุ่มในตัวพยากรณ์เชิงเส้น ค่าคาดหวังของตัวแปรตอบสนองซึ่งสัมพันธ์กับตัวพยากรณ์เชิงเส้นด้วย link function ถูกกำหนดโดย

$$\mu_{ij}^u = E(Y_{ij} \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i) = \eta_{ij} = \boldsymbol{x}_{ij}'\boldsymbol{\beta} + \boldsymbol{z}_{ij}'\boldsymbol{u}_{\mathrm{i}}$$

เมื่อ $\boldsymbol{z}_{ij}$ เป็นเวกเตอร์ขนาด $r \times 1$ ของตัวแปรอธิบายสำหรับอิทธิพลสุ่ม r ตัว และ $\boldsymbol{u}_i$ คือเวกเตอร์อิทธิพลสุ่ม r ตัว ของกลุ่ม i ซึ่งสมมติว่ามีการแจกแจงแบบปกติหลายตัวแปร ด้วยค่าเฉลี่ย $\boldsymbol{0}$ และเมทริกซ์ความแปรปรวนร่วม $\boldsymbol{\Sigma}_u$ ดังนั้นค่าเฉลี่ยแบบมีเงื่อนไข μ_{ij}^u คือ $E(Y_{ij} \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i)$ ซึ่งอยู่ในเทอมของเวกเตอร์ของอิทธิพลสุ่ม กรณีอย่างง่ายคือให้ $\boldsymbol{z}_{ij}$ เป็นเวกเตอร์ของ 1 แล้ว u_i จะเป็นอิทธิพลสุ่มตัวแปรเดียวจาก $N(0, \sigma_u^2)$ นี่คือค่าคาดหวังของการแจกแจงแบบมีเงื่อนไขของตัวแปรตอบสนองเมื่อกำหนดเงื่อนไขบนอิทธิพลสุ่ม ส่งผลให้ตัวแบบ GLMMs มักอ้างถึงตัวแบบเงื่อนไข (Conditional model) ในเชิงเปรียบเทียบกับตัวแบบมาร์จินัล (Marginal model) ซึ่งเป็นทางเลือกของตัวแบบ GLMs สำหรับข้อมูลที่มีความสัมพันธ์กัน (Agresti 2002)

1.2.1 ตัวแบบอิทธิพลสุ่มและการประมาณพารามิเตอร์ของตัวแบบ

สำหรับตัวแบบลอจิตสะสม เมื่อเพิ่มเทอมของอิทธิพลสุ่ม $\boldsymbol{u}_i$ เข้าไปในตัวพยากรณ์เชิงเส้น จะเรียกว่าตัวแบบลอจิตสะสมแบบอิทธิพลสุ่มหรือตัวแบบผสมเชิงเส้นนัยทั่วไป มีรูปแบบคือ

$$\mathrm{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i)] = \alpha_k + \boldsymbol{x}_{ij}'\boldsymbol{\beta} + \boldsymbol{z}_{ij}'\boldsymbol{u}_i, \quad k = 1, \ldots, c-1$$

โดยที่ $\boldsymbol{u}_i$ แทนเวกเตอร์อิทธิพลสุ่ม r ตัว ของกลุ่ม i ซึ่งเป็นอิสระต่อกันจากการแจกแจงแบบปกติ ที่มีค่าเฉลี่ย $\boldsymbol{0}$ และเมทริกซ์ความแปรปรวนร่วม $\boldsymbol{\Sigma}_u$ และ $\boldsymbol{z}_{ij}'$ คือเวกเตอร์ของตัวแปรอธิบายสำหรับอิทธิพลสุ่ม ตัวแบบอิทธิพลสุ่มอย่างง่าย คือ Random intercept models $\left(\text{สำหรับ } \boldsymbol{z}_{ij}' = 1\right)$ อยู่ในรูป

$$\mathrm{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, u_i)] = \alpha_k + \boldsymbol{x}_{ij}'\boldsymbol{\beta} + u_i, \quad k = 1, \ldots, c-1$$

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

โดยที่ u_i แทนอิทธิพลสุ่มของกลุ่ม i ซึ่งเป็นอิสระต่อกันจากการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และความแปรปรวน σ_u^2 ภาวะน่าจะเป็นแบบมีเงื่อนไขของกลุ่ม i คือ จากแนวคิดที่ว่าตัวแปรตอบสนองแบบอันดับมาจากการจัดกลุ่มของตัวแปรแฝงแบบต่อเนื่อง Y^* ซึ่งเป็นตัวแปรที่ไม่ได้สังเกตค่า สำหรับตัวแบบอิทธิพลสุ่ม กรณีของ Random intercept models จะได้ว่า Y^* มีความสัมพันธ์กับตัวแปรอธิบาย x ดังนี้

$$Y_{ij}^* = -\boldsymbol{x}_{ij}'\boldsymbol{\beta} - u_i + \varepsilon_{ij}$$

เมื่อ ε_{ij} คือความคลาดเคลื่อนของตัวแปรแฝง และ $Y_{ij} = k$ เมื่อ $\alpha_{k-1} < Y_{ij}^* \le \alpha_k$

ดังนั้นความน่าจะเป็นสะสมแบบมีเงื่อนไข สามารถหาได้จาก

$$\begin{aligned} P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, u_i) &= P(Y_{ij}^* \le \alpha_k \mid \boldsymbol{x}_{ij}, u_i) \\ &= P(-\boldsymbol{x}_{ij}'\boldsymbol{\beta} - u_i + \varepsilon_{ij} \le \alpha_k) \\ &= P(\varepsilon_{ij} \le \alpha_k + \boldsymbol{x}_{ij}'\boldsymbol{\beta} + u_i) \\ &= F(\eta_{ijk}^u) \end{aligned}$$

เมื่อ

$$F(\eta_{ijk}^u) = 1/[1 + exp(-\eta_{ijk}^u)]$$

โดยที่ $F(\eta_{ij0}^u) = 0$ และ $F(\eta_{ijc}^u) = 1$ ดังนั้น $var(Y_{ij}^*) = var(u_i) + var(\varepsilon_{ij}) = \sigma_\varepsilon^2 + \sigma_u^2 = \frac{\pi^2}{3} + \sigma_u^2$

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ดังนั้นความน่าจะเป็นที่ Y_{ij} ตกอยู่ในระดับ k หาได้จาก

$$\pi_k^u(\boldsymbol{x}_{ij}) = P(Y_{ij} = k \mid \boldsymbol{x}_{ij}, u_i) = F(\eta_{ijk}^u) - F(\eta_{ij,k-1}^u)$$

การประมาณพารามิเตอร์ของตัวแบบอิทธิพลสุ่มใช้วิธีภาวะน่าจะเป็นสูงสุด สำหรับตัวแปรตอบสนองแบบอันดับ ฟังก์ชันภาวะน่าจะเป็นแบบมีเงื่อนไขของกลุ่ม i คือ

$$f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta}) = \prod_{j=1}^{n_i} \prod_{k=1}^{c} \pi_k^u(\boldsymbol{x}_{ij})^{y_{ijk}}$$

โดยที่ $y_{ijk} = 1$ ถ้า $Y_{ij} = k$ และ $y_{ijk} = 0$ ถ้า $Y_{ij} \ne k$, ข้อสมมติที่ว่าค่าตอบสนองของหน่วยที่ i เป็นอิสระต่อกันเมื่อกำหนดอิทธิพลสุ่ม (จึงสามารถนำมาคูณกันได้) เรียกข้อสมมติความเป็นอิสระกันแบบมีเงื่อนไข ภาวะน่าจะเป็นแบบมาร์จินัลของ $\boldsymbol{Y}_i$ ในประชากรแสดงเป็นสูตรของอินทิกรัลของภาวะน่าจะเป็นแบบมีเงื่อนไข $f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta})$

$$L_i(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = \int f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta}) \times g(\boldsymbol{u}_i; \boldsymbol{\Sigma}_u) d\boldsymbol{u}_i$$

เมื่อ $g(\boldsymbol{u}_i; \boldsymbol{\Sigma}_u)$ แทนฟังก์ชันความหนาแน่นของอิทธิพลสุ่ม ซึ่งมักจะสมมติว่าเป็นฟังก์ชันความหนาแน่นแบบปกติหลายตัวแปร ขณะที่ $f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta})$ แทนภาวะน่าจะเป็นแบบมีเงื่อนไขของกลุ่ม i และ $L_i(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u)$ คือภาวะน่าจะเป็นแบบไม่มีเงื่อนไข (แบบมาร์จินัล) สำหรับเวกเตอร์ค่าตอบสนองของกลุ่ม i แล้ว The log-likelihood function จากกลุ่มตัวอย่าง N กลุ่ม คือ

$$l(\boldsymbol{\beta},\boldsymbol{\Sigma}_u) = \log L(\boldsymbol{\beta},\boldsymbol{\Sigma}_u) = \log \prod_{i=1}^{N} L_i(\boldsymbol{\beta},\boldsymbol{\Sigma}_u) = \sum_{i=1}^{N} \log L_i(\boldsymbol{\beta},\boldsymbol{\Sigma}_u) \qquad (2)$$

การหาค่าประมาณพารามิเตอร์ที่ทำให้ภาวะน่าจะเป็นนี้มีค่าสูงสุด จะทำให้ได้ตัวประมาณภาวะน่าจะเป็นสูงสุด (บางทีเรียกว่าตัวประมาณภาวะน่าจะเป็นแบบมาร์จินัลสูงสุด) ของเวกเตอร์พารามิเตอร์ $\boldsymbol{\beta}$ (รวมถึงจุดตัด α_k) และเมทริกซ์ความแปรปรวนร่วมของอิทธิพลสุ่ม $\boldsymbol{\Sigma}_u$ จะเห็นว่าภาวะน่าจะเป็นมีเทอมของอินทิกรัลของฟังก์ชันที่ซับซ้อน คำนวณยาก จึงต้องมีการประมาณฟังก์ชันภาวะน่าจะเป็นนี้ขึ้นมาก่อนที่จะใช้วิธีภาวะน่าจะเป็นสูงสุดในการประมาณพารามิเตอร์ (Agresti, 2002) ซึ่งมีวิธีการอยู่หลายวิธีที่พัฒนาขึ้นมา โดยวิธีที่ได้รับความนิยม มีดังนี้

1.2.1.1 Gaussian Quadrature Approach

วิธี Gaussian Quadrature (GQ) หรือเรียกอีกชื่อหนึ่งว่า Gauss-Hermite quadrature (GH) เสนอโดย Liu and Pierce (1994) เป็นวิธีประมาณอินทิกรัล โดยการหาผลรวมแบบถ่วงน้ำหนัก เมื่อกำหนดจำนวน Quadrature point ในแต่ละมิติของการอินทิเกรต ข้อสังเกตหนึ่งของวิธีการนี้ก็คือเป็นการหาผลรวมตลอดจำนวน Quadrature point ที่มีค่ามาก โดยเฉพาะเมื่ออิทธิพลสุ่มมีจำนวนมากขึ้น ในการแก้ปัญหานี้จึงมีการพัฒนาเป็นวิธี Adaptive Gaussian Quadrature โดย Pinhero & Bate (1995) ซึ่งใช้จำนวน Quadrature point ในแต่ละมิติน้อย ๆ โดยการปรับ Location และ Dispersion ของการแจกแจงที่จะอินทิเกรต วิธี Gaussian Quadrature เป็นวิธีสำหรับการประมาณอินทิกรัลของฟังก์ชัน $f(\cdot)$ คูณกับฟังก์ชันอื่นที่มีฟังก์ชันความหนาแน่นแบบปกติ โดยหาผลรวมแบบถ่วงน้ำหนักแบบจำกัดในจุดที่แน่นอน สำหรับกรณีอิทธิพลสุ่มตัวแปรเดียว การประมาณจะอยู่ในรูป

$$\int_{-\infty}^{\infty} f(\boldsymbol{u}) exp(-\boldsymbol{u}^2) d\boldsymbol{u} \approx \sum_{t=1}^{q} c_t f(d_t)$$

เมื่อ c_t คือน้ำหนัก และ d_t คือ Quadrature points ที่ทำให้เรียบ การประมาณจะดีขึ้นเมื่อ จำนวน Quadrature points (q) เพิ่มขึ้น ภาวะน่าจะเป็นที่ประมาณขึ้นสามารถใช้ในการประมาณพารามิเตอร์ที่ทำให้ภาวะน่าจะเป็นมีค่าสูงสุดได้โดยขั้นตอนวิธีมาตรฐาน เช่นวิธีนิวตัน-รัฟสัน (Newton-Raphson) จะทำให้ได้ตัวประมาณภาวะน่าจะเป็นสูงสุด $\hat{\boldsymbol{\beta}}$ และ $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}$ การหาส่วนกลับของ Information matrix จากค่าสังเกต จะให้ความคลาดเคลื่อนมาตรฐานของตัวประมาณภาวะน่าจะเป็นสูงสุด สำหรับตัวแบบที่ซับซ้อนมากขึ้น การประมาณที่เหมาะสม ในการประมาณความคลาดเคลื่อนมาตรฐาน อาจจะใช้จำนวน Quadrature points มากกว่าการประมาณ $\hat{\boldsymbol{\beta}}$ ซึ่งเราจะเพิ่มจำนวน Quadrature points ขึ้นเรื่อย ๆ จนมีการเปลี่ยนแปลงที่น้อยมากทั้งตัวประมาณสัมประสิทธิ์และความคลาดเคลื่อนมาตรฐาน ดังนี้

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จาก The log-likelihood function ของความน่าจะเป็นแบบมาร์จินัล ดังสมการ (2) ดังนั้นอาจใช้พีชคณิตเชิงเส้นเขียนตัวพยากรณ์เชิงเส้นใหม่ในรูปแบบทั่วไป ดังนี้

$$\boldsymbol{\eta}_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{u}_i$$

โดยที่ $\boldsymbol{\eta}_{ij} = (\eta_{ij1},\ldots,\eta_{ijc})'$ ในการหาเมทริกซ์ความแปรปรวนร่วม เทอมของอิทธิพลสุ่มอาจเขียนใหม่เป็น $\boldsymbol{u}_i = \boldsymbol{\Sigma}_u^{1/2} a_i$ เมื่อ $\boldsymbol{\Sigma}_u^{1/2}$ คือรากที่สองของ $\boldsymbol{\Sigma}_{\boldsymbol{u}}$ บนพื้นฐานของการแยกเชิงสเปกตรัม และ a_i เป็นเวกเตอร์สุ่มมาตรฐาน นั่นคือ $a_i \sim N(\boldsymbol{0},\boldsymbol{I})$ จึงเขียนตัวพยากรณ์เชิงเส้น $\boldsymbol{\eta}_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{u}_i$ ใหม่ในรูปพีชคณิตเชิงเส้นง่าย ๆ โดยให้ $\boldsymbol{\theta} = vec(\boldsymbol{\Sigma}_u^{1/2})$ เมื่อ vec คือการนำคอลัมน์เวกเตอร์มาเรียงต่อกันเป็นเวกเตอร์เดียว ได้เป็น

$$\boldsymbol{\eta}_{ij} = [\boldsymbol{x}'_{ij}, a'_i \otimes \boldsymbol{z}'_{ij}]\begin{bmatrix}\boldsymbol{\beta}\\ \boldsymbol{\theta}\end{bmatrix}$$

โดยมีพารามิเตอร์ที่ไม่ทราบค่า $\boldsymbol{\tau}' = (\boldsymbol{\beta}',\boldsymbol{\theta}')$ ดังนั้นการแจกแจงของอิทธิพลสุ่มจะถูกกำหนดโดย a_i และ $\boldsymbol{\tau}$ แทนที่ $\boldsymbol{u}_i$ และ $\boldsymbol{\beta}$ เราจะเขียนฟังก์ชันความหนาแน่นของค่าสังเกตเป็น $f(y_{ij} \mid \boldsymbol{\tau}, a_i)$ แทน $f(y_{ij} \mid \boldsymbol{\beta}, \boldsymbol{u}_i)$ การใช้วิธี Gauss-Hermite Quadrature จะประมาณ $l(\boldsymbol{\beta},\boldsymbol{\Sigma}_{\boldsymbol{u}})$ ด้วย

$$l^{GH}(\boldsymbol{\tau}) = \sum_{i=1}^{N} \log L_i^{GH}(\boldsymbol{\tau})$$

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

โดยที่ $L_i^{GH}(\boldsymbol{\tau}) = \sum_{t=1}^{q} \upsilon_t f(\boldsymbol{Y}_i \mid d_t;\boldsymbol{\tau})$, $d_t, t = 1,\ldots,q$ คือ Quadrature points และ $\upsilon_t, t = 1,\ldots,q$ แทนน้ำหนัก ทั้ง d_t และ υ_t นั้นสามารถเปิดได้จากตาราง (Abramowitz and Stegun 1972 : 924) The score function บนพื้นฐานของ $l^{GH}(\boldsymbol{\tau})$ อยู่ในรูป

$$s^{GH}(\boldsymbol{\tau}) = \sum_{i=1}^{N}\sum_{t=1}^{q} c_{it}^{GH}(\boldsymbol{\tau}) \frac{\partial \log f(\boldsymbol{Y}_i \mid d_t;\boldsymbol{\tau})}{\partial \boldsymbol{\tau}}$$

โดยที่

$$c_{it}^{GH}(\boldsymbol{\tau}) = \frac{\upsilon_t f(\boldsymbol{Y}_i \mid d_t;\boldsymbol{\tau})}{\sum_{t=1}^{q} \upsilon_t f(\boldsymbol{Y}_i \mid d_t;\boldsymbol{\tau})} \qquad \text{และ} \qquad \sum_{j=1}^{q} c_{it}^{GH}(\boldsymbol{\tau}) = 1$$

The score function $s^{GH}(\boldsymbol{\tau})$ อาจจะคำนวณโดยใช้วิธีของ GLMs โดยให้ $\eta_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + a'_i \otimes \boldsymbol{z}'_{ij}\boldsymbol{\theta}$ และหา

$$\frac{\partial \log f(\boldsymbol{Y}_i \mid d_t;\boldsymbol{\tau})}{\partial \boldsymbol{\beta}} = \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}\delta(\eta_{ijk})\sigma^{-1}(\eta_{ijk})\{y_{ij} - h((\eta_{ijk}))\}$$

$$\frac{\partial \log f(\boldsymbol{Y}_i \mid d_t;\boldsymbol{\tau})}{\partial \boldsymbol{\theta}} = \sum_{j=1}^{n_i} (d_t \otimes \boldsymbol{z}_{ij})\delta(\eta_{ijk})\sigma^{-1}(\eta_{ijk})\{y_{ij} - h((\eta_{ijk}))\}$$

โดยที่ $\delta(\eta_{ijk}) = \partial h(\eta_{ijk}) / \partial \eta$ และ $\sigma(\eta_{ijk}) = V(y_{ij} \mid \eta_{ijk}))$

ผลลัพธ์ของสมการ $s^{GH}(\boldsymbol{\tau}) = 0$ ได้มาโดยกระบวนการย้อนซ้ำ เช่นวิธี Newton-Raphson หรือ Fisher scoring ในการย้อนซ้ำรอบที่ *s, s=1, ...* ค่าน้ำหนัก $c_{it}^{GH}(\boldsymbol{\tau})$ จะคำนวณจากค่าประมาณในรอบที่ *s* หรือ $\hat{\boldsymbol{\tau}}^{(s)}$ และปรับปรุงค่าประมาณให้ดีขึ้นจากการแก้สมการ

$$\sum_{i=1}^{N}\sum_{t=1}^{q} c_{it}^{GH}(\hat{\boldsymbol{\tau}}^{(s)})\frac{\partial \log f(\boldsymbol{Y}_i \mid d_t;\hat{\boldsymbol{\tau}}^{(s+1)})}{\partial \boldsymbol{\tau}} = 0$$

สำหรับอิทธิพลสุ่ม $\boldsymbol{u}_i$ สามารถประมาณได้โดยใช้วิธีการของ Empirical Bayes สำหรับกรณีที่อิทธิพลสุ่มมีตัวเดียว ตัวประมาณ $\hat{\boldsymbol{u}}_i$ นิยามโดย

$$\hat{\boldsymbol{u}}_i = E(\boldsymbol{u}_i \mid \boldsymbol{Y}_i) = h_i^{-1}\int_{\boldsymbol{u}_i} \boldsymbol{u}_i f(\boldsymbol{Y}_i \mid \boldsymbol{u}_i)\mathrm{g}(\boldsymbol{u}_i)d\boldsymbol{u}_i$$

เมื่อ $f(\boldsymbol{Y}_i \mid \boldsymbol{u}_i)$ แทนความน่าจะเป็นแบบมีเงื่อนไขสำหรับกลุ่มที่ *i* และ h_i คล้ายคลึงกับความน่าจะเป็นแบบมาร์จินัลหรือค่าเฉลี่ยของการแจกแจงภายหลังแบบมาร์จินัลนั่นเอง เช่นเดียวกับความแปรปรวนของการแจกแจงภายหลังได้มาจาก

$$V(\hat{\boldsymbol{u}}_i \mid \boldsymbol{Y}_i) = h_i^{-1}\int_{\boldsymbol{u}_i} (\hat{\boldsymbol{u}}_i - \boldsymbol{u}_i)^2 f(\boldsymbol{Y}_i \mid \boldsymbol{u}_i)f(\boldsymbol{u}_i)d\boldsymbol{u}_i$$

1.2.1.2 Penalized Quasi-Likelihood Approach

Penalized Quasi-Likelihood (PQL) เสนอโดย Green (1987) ในการประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่ม เนื่องจากฟังก์ชันภาวะน่าจะเป็นได้มาจากการอินทิเกรตอิทธิพลสุ่ม $\boldsymbol{u}$ ออกจากการแจกแจงร่วมของ $\boldsymbol{y}$ และ $\boldsymbol{u}$ โดยใช้วงศ์ชี้กำลังแทนแต่ละส่วนของการแจกแจงร่วม วิธีในการประมาณฟังก์ชันคือใช้การกระจายอนุกรมเทเลอร์อันดับที่สอง รอบ ๆ จุด $\bar{\boldsymbol{u}} \approx E(\boldsymbol{u} \mid \boldsymbol{y})$ ฟังก์ชันที่ประมาณสำหรับอินทิกรัลคือกำลังสองของ $(\boldsymbol{u} - \bar{\boldsymbol{u}})$ และอยู่ในรูปค่าคงที่คูณกับฟังก์ชันความหนาแน่นแบบปกติหลายตัวแปร ดังนั้นอินทิกรัลจึงอยู่ในรูปแบบปกติ การประมาณอินทิกรัลชนิดนี้เรียกว่า Laplace approximation สำหรับวิธีนี้ (Breslow and Clayton 1993) การประมาณอินทิกรัลจะได้ฟังก์ชันการประมาณ The log-likelihood function เรียกว่า Penalized quasi-likelihood ซึ่งอยู่ในรูป

$$q(\boldsymbol{\beta}, \boldsymbol{y}) - (1/2)\bar{\boldsymbol{u}}'\Sigma_{\boldsymbol{u}}^{-1}\bar{\boldsymbol{u}}$$

โดยที่ $q(\boldsymbol{\beta}, \boldsymbol{y})$ คล้ายคลึงกับฟังก์ชัน Quasi-log-likelihood สำหรับ GLM เมื่อกำหนดเงื่อนไขบน $\boldsymbol{u} = \bar{\boldsymbol{u}}$ การคำนวณหาค่าสูงสุดจะใช้วิธีสำหรับตัวแบบผสมเชิงเส้น (Linear Mixed Models: LMMs) กรณีของตัวแปรตอบสนองที่มีการแจกแจงแบบปกติ ซึ่งกำหนดความเป็นเชิงเส้นของลอจิตว่าเป็นตัวแปรตอบสนอง และนำมาซึ่งผลการย้อนซ้ำของชุดสมการภาวะน่าจะเป็นในเทอมของ $\boldsymbol{\beta}$ และ $\boldsymbol{u}$ วิธี PQL ไม่ต้องใช้การอินทิเกรตเชิงตัวเลขจึงง่ายกว่าวิธีของตัวประมาณภาวะน่าจะเป็นสูงสุดและการคำนวณสามารถทำได้เมื่อตัวอย่างมีขนาดใหญ่หรือตัวแบบที่มีโครงสร้างของอิทธิพลสุ่มซับซ้อน แต่มีความเอนเอียงเมื่อความแปรปรวนของอิทธิพลสุ่มมากขึ้น

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2. งานวิจัยที่เกี่ยวข้อง

Liang and Zeger (1986) เสนอวิธี Generalized Estimating Equations (GEE) ในการประมาณพารามิเตอร์ของตัวแบบมาร์จินัล สำหรับข้อมูลตัวแปรตอบสนองแบบทวิภาคที่มีการวัดซ้ำในช่วงเวลาหนึ่ง ซึ่งเป็นวิธีการที่ขยายจากวิธีการของตัวแบบเชิงเส้นนัยทั่วไป ภายใต้โครงสร้างสหสัมพันธ์แบบ Independence, M-dependence และ Exchangeable และพบว่าตัวประมาณพารามิเตอร์พร้อมทั้งความแปรปรวน ด้วยวิธี GEE จะเป็นตัวประมาณที่คงเส้นคงวาและมีการแจกแจงแบบปกติเมื่อใกล้อนันต์

Green (1987) เสนอวิธี Penalized Quasi-Likelihood (PQL) ในการประมาณภาวะน่าจะเป็นของตัวแบบอิทธิพลสุ่ม โดยใช้ Laplace approximation โดยพบว่าวิธี PQL จะให้ค่าประมาณพารามิเตอร์ที่มีความเอนเอียงกรณีที่ขนาดตัวอย่างมีขนาดเล็กและเมื่อสหสัมพันธ์ของข้อมูลในกลุ่มมีค่ามาก

Goldstein (1991) เสนอวิธี Marginal Quasi-Likelihood (MQL) ในการประมาณพารามิเตอร์ของตัวแบบมาร์จินัล โดยพบว่าวิธี MQL จะให้ค่าประมาณพารามิเตอร์ที่มีความเอนเอียงกรณีที่ขนาดตัวอย่างมีขนาดเล็ก

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

Breslow and Clayton (1993) ได้กล่าวถึงการประมาณพารามิเตอร์ในตัวแบบ สำหรับตัวแปรตอบสนองแบบทวิภาค เมื่อค่าตอบสนองมีความสัมพันธ์กัน โดยใช้วิธีประมาณพารามิเตอร์ 2 วิธีคือ Penalized Quasi -Likelihood (PQL) และ Marginal Quasi-Likelihood (MQL) โดยการจำลองแบบข้อมูลภายใต้เงื่อนไข ดังนี้ จำนวนกลุ่มตัวอย่างเท่ากับ 100 ขนาดกลุ่มตัวอย่างเท่ากับ 7 และอิทธิพลสุ่ม 2 ตัวที่เป็นอิสระกันซึ่งมีความแปรปรวน (1.0, 0.0) และ (0.50, 0.25) ผลการศึกษาพบว่าวิธี PQL ให้ค่าประมาณสัมประสิทธิ์ของตัวแบบและความคลาดเคลื่อนมาตรฐานที่ต่ำกว่าค่าจริง และให้ค่าประมาณองค์ประกอบความแปรปรวนที่ต่ำกว่าค่าจริงอย่างมาก ส่วนวิธี MQL ให้ค่าประมาณความแปรปรวนของอิทธิพลสุ่มที่สูงกว่าค่าจริง

Liu and Pierce (1994) เสนอวิธี Gaussian Quadrature (GQ) หรือเรียกอีกชื่อหนึ่งว่า Gauss-Hermite quadrature (GH) ในการประมาณภาวะน่าจะเป็นของตัวแบบอิทธิพลสุ่ม ซึ่งติดเทอมของอินทิกรัลอยู่ จึงคำนวณยาก โดยวิธีการหาผลรวมแบบถ่วงน้ำหนัก โดยกำหนดจำนวน Quadrature point ในการหาผลรวม ก่อนที่จะหาค่าประมาณพารามิเตอร์ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นมีค่าสูงสุด โดยวิธีนี้เป็นวิธีที่จะมีความถูกต้องมากขึ้น เมื่อจำนวน Quadrature point มากขึ้น และก็จะใช้เวลาในการคำนวณมากขึ้นด้วย

Pinheiro and Bates (1995) ได้ปรับปรุงวิธี Gaussian Quadrature โดยการปรับ Location และ Dispersion ของการแจกแจงที่จะอินทิเกรต เพื่อให้ใช้จำนวน Quadrature point ในแต่ละมิติของการอินทิเกรตน้อย ๆ และเรียกว่า “Adaptive Gaussian Quadrature ” หรือ AGQ

Tuzt and Hennevogl (1996) ได้ทำการศึกษาการวิเคราะห์ข้อมูลตัวแปรตอบสนองแบบอันดับ เมื่อค่าตอบสนองมีความสัมพันธ์กัน โดยเปรียบเทียบวิธีการประมาณพารามิเตอร์ 3 วิธีบนพื้นฐานของ EM Algorithm ประกอบด้วย วิธี Gaussian Quadrature (GQ), Monte Carlo (MC) และวิธี EM Algorithm บนพื้นฐานของ Posterior modes ผลการศึกษา พบว่าวิธี GQ และ MC ให้ค่าประมาณที่มีค่าใกล้เคียงกันสำหรับจำนวน Quadrature points ของ GQ ประมาณ 8-10 และจำนวน Quadrature points ของ MC ประมาณ 20 และพบว่าวิธี GQ และ MC ให้ค่าประมาณที่แข็งแรงกว่าการใช้ตัวแบบอิทธิพลคงที่ และพบว่าค่าประมาณความคลาดเคลื่อนมาตรฐานของวิธี MC สูงกว่าวิธี GQ ทั้งหมด นอกจากนี้ยังมีข้อเสนอว่าขนาดของกลุ่มตัวอย่างไม่ควรน้อยเกินไป เช่น ถ้าตัวแปรตอบสนองมี 3 ระดับ ขนาดของกลุ่มตัวอย่างเท่ากับ 5 ก็เพียงพอ และถ้าขนาดกลุ่มตัวอย่างมีค่ามาก อาจจะมีปัญหาเนื่องจากมีพารามิเตอร์ที่ต้องประมาณจำนวนมาก

Spiess and Hamerle (2000) ได้ทำการจำลองแบบข้อมูลเพื่อเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบสำหรับตัวแปรตอบสนองแบบทวิภาคที่มีความสัมพันธ์กัน โดยใช้ตัวแบบโพรบิต วิธีประมาณพารามิเตอร์ที่นำมาพิจารณาประกอบด้วยวิธี Gaussian Quadrature (GQ), Generalized Estimating Equations (GEE) และวิธี MECOSA ซึ่งเป็นวิธีที่เสนอโดย Schepers et al. (1991) โดยใช้ตัวแบบโพรบิต เงื่อนไขในการจำลองแบบข้อมูล 500 ครั้ง มีดังนี้ จำนวนกลุ่มตัวอย่างเท่ากับ 50, 100, 250, 50 และ 1,000 ขนาดกลุ่มตัวอย่างเท่ากับ 5 ความแปรปรวนของอิทธิพลสุ่มเท่ากับ 0.2, 0.5 และ 0.8 ตัวแปรอธิบายจำนวน 3 ตัวแปร ซึ่งมีการแจกแจงแบบปกติ ยูนิฟอร์ม ที่มีค่าเฉลี่ยเท่ากับ 0 และเบอร์นูลลีที่มีความน่าจะเป็นเท่ากับ 0.5 ผลการศึกษาพบว่าวิธี GQ และ GEE มีแนวโน้มที่ให้ค่าประมาณสูงกว่าค่าจริง วิธี GQ มีประสิทธิภาพสูงกว่าวิธี GEE เมื่อสหสัมพันธ์ในกลุ่มมีค่ามาก แต่มีประสิทธิภาพต่ำกว่าวิธี GEE เมื่อสหสัมพันธ์ในกลุ่มมีค่าน้อยและจำนวนกลุ่มตัวอย่างมีขนาดเล็ก นอกจากนี้ยังพบว่าเมื่อจำนวนกลุ่มตัวอย่างเพิ่มขึ้น ค่าประมาณพารามิเตอร์จะมีความเอนเอียงและความคลาดเคลื่อนกำลังสองเฉลี่ยลดลง และเมื่อสหสัมพันธ์ในกลุ่มมีค่าเพิ่มขึ้น ค่าประมาณพารามิเตอร์จะมีความเอนเอียงสูงขึ้น โดยที่ในบรรดาปัจจัยต่าง ๆ เช่นจำนวนกลุ่มตัวอย่าง ขนาดของสัมประสิทธิ์ของตัวแบบ และสหสัมพันธ์ในกลุ่ม พบว่าสหสัมพันธ์ในกลุ่มมีอิทธิพลต่อความถูกต้องของค่าประมาณพารามิเตอร์มากที่สุด

Chen and Kuo (2001) ได้แสดงวิเคราะห์การวิเคราะห์ตัวแบบมัลติโนเมียลลอจิตที่รวมอิทธิพลสุ่ม โดยใช้ความสัมพันธ์ระหว่างตัวแบบมัลติโนเมียลและตัวแบบปัวซงเพื่อปรับตัวแบบ

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

มัลติโนเมียลลอจิตที่รวมอิทธิพลสุ่มโดยการเสนอตัวแบบปัวซงล็อกลิเนียร์ (Poisson log-linear models) และตัวแบบปัวซงนอนลิเนียร์ (Poisson nonlinear models) เนื่องจากภาวะน่าจะเป็นของทั้ง 2 ตัวแบบสมมูลกับภาวะน่าจะเป็นของตัวแบบมัลติโนเมียลลอจิตซึ่งยังไม่มีโปรแกรมสำเร็จรูปในการวิเคราะห์ และวิธีการของตัวแบบเหล่านี้ทำได้ง่ายด้วยคำสั่งทางสถิติที่มีอยู่ในโปรแกรม SAS ตัวแบบปัวซงล็อกลิเนียร์อิทธิพลคงที่ ใช้คำสั่ง GENMOD ตัวแบบปัวซงล็อกลิเนียร์อิทธิพลสุ่ม ใช้คำสั่ง GLIMMIX ตัวแบบปัวซงล็อกลิเนียร์จะต้องมีการประมาณค่าพารามิเตอร์รบกวนเพิ่มอีก $\sum_{i=1}^{N} n_i$ เมื่อ N คือจำนวนกลุ่มตัวอย่าง และ n_i คือจำนวนค่าสังเกตภายในกลุ่ม ดังนั้นตัวแบบปัวซงล็อกลิเนียร์จึงมีความไม่พอเพียงในการคำนวณเนื่องจากมีจำนวนพารามิเตอร์ที่ต้องประมาณค่าจำนวนมาก ส่วนตัวแบบปัวซงนอนลิเนียร์อิทธิพลคงที่และอิทธิพลสุ่ม ใช้คำสั่ง NLMIXED และใช้ชุดข้อมูลการเลือกยี่ห้อสินค้า ในการอธิบายวิธีการที่นำเสนอ ผลการศึกษาพบว่าวิธี NLMIXED เป็นที่น่าพอใจมากกว่าวิธี GLIMMIX เนื่องจากมีจำนวนพารามิเตอร์ที่ต้องประมาณค่าน้อยกว่าและใช้เวลาน้อยกว่า

Tutz and Kauermann (2003) ได้เสนอตัวแบบอิทธิพลสุ่มที่มีสัมประสิทธิ์ β_t ไม่คงที่เปลี่ยนแปลงไปตามเวลา ในการศึกษาระยะยาว โดยใช้วิธี Gaussian Quadrature (GQ) ในการหาค่าสูงสุดของภาวะน่าจะเป็น เพื่อประมาณพารามิเตอร์ของตัวแบบลอจิสติกสำหรับตัวแปรตอบสนองแบบทวิภาค โดยกล่าวว่าวิธี Gaussian Quadrature (GQ) ใช้ได้ดีเมื่ออิทธิพลสุ่มมีการแจกแจงแบบปกติ 1 ตัวแปร เนื่องจากใช้เวลาคำนวณน้อยกว่า ส่วนวิธี EM Algorithm จะใช้ได้ดีในกรณีที่มีอิทธิพลสุ่มหลายตัวแปร

Kuss and McLerran (2007) ได้แสดงการประมาณค่าพารามิเตอร์สำหรับตัวแบบมัลติโนเมียลลอจิต เมื่อค่าตอบสนองมีความสัมพันธ์กัน โดยใช้โปรแกรม SAS® โดยมีแรงจูงใจมาจากงานวิจัยของ Chen and Kuo (2001) ที่ใช้ภาวะน่าจะเป็นแบบปัวซง ซึ่งมีความถูกต้องทางสถิติแต่มีความไม่พอเพียงในการคำนวณ งานวิจัยนี้จึงเสนอวิธีการประมาณค่าพารามิเตอร์ 4 วิธี ประกอบด้วย Gaussian Quadrature (GQ) และ Penalized Quasi-Likelihood (PQL) สำหรับตัวแบบอิทธิพลสุ่ม และ Generalized Estimating Equations (GEE) และ Marginal Quasi-Likelihood (MQL) สำหรับตัวแบบมาร์จินัล และใช้ข้อมูลตัวอย่าง 2 ชุด ในการเปรียบเทียบ ผลการศึกษาพบว่าวิธี GEE และ MQL ให้ค่าประมาณที่ใกล้เคียงกันมาก ส่วนวิธี GQ และ PQL ให้ค่าประมาณที่ต่างกันบางครั้ง เนื่องมาจากความแตกต่างของการประมาณค่าองค์ประกอบของความแปรปรวน ซึ่งเป็นที่รู้กันดีว่าวิธี PQL ให้ค่าประมาณความแปรปรวนของอิทธิพลสุ่มที่ต่ำกว่าค่าจริง และวิธี GQ ให้ค่าประมาณความแปรปรวนของอิทธิพลสุ่มสูงที่สุด

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

Moineddin, Matheson and Glazier (2007) ได้ทำการจำลองแบบข้อมูลเพื่อประเมินอิทธิพลของขนาดตัวอย่างทั้งจำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องของค่าประมาณพารามิเตอร์และองค์ประกอบความแปรปรวนของตัวแบบลอจิสติกพหุระดับ สำหรับตัวแปรตอบสนองแบบทวิภาค นอกจากนี้ยังศึกษาอิทธิพลของสัดส่วนผลตอบสนองและสหสัมพันธ์ในกลุ่มด้วย โดยความถูกต้องของค่าประมาณพารามิเตอร์พิจารณาจากความเอนเอียงสัมพัทธ์และความถูกต้องของความคลาดเคลื่อนของค่าประมาณพารามิเตอร์พิจารณาจาก Coverage (%) ที่ 95% โดยใช้การแจกแจงแบบปกติเมื่อใกล้อนันต์ จากการจำลองแบบข้อมูล 1,000 ครั้ง ภายใต้เงื่อนดังนี้ จำนวนกลุ่มตัวอย่างเท่ากับ 30, 50 และ 100 ขนาดกลุ่มตัวอย่างเท่ากับ 5, 30 และ 50 ความแปรปรวนของอิทธิพลสุ่มเท่ากับ 0.13, 0.67 และ 2.00 ซึ่งสอดคล้องกับค่าสหสัมพันธ์ในกลุ่ม 0.04, 0.17 และ 0.38 ตามลำดับ ตัวแปรอธิบายมีการแจกแจงแบบปกติมาตรฐาน จำนวน 2 ตัวแปร ตัวแปรตอบสนองแบบทวิภาคที่มีสัดส่วนผลตอบสนองเท่ากับ 0.10, 0.34 และ 0.45 และประมาณพารามิเตอร์โดยใช้วิธี Adaptive Gaussian Quadrature (AGQ) ผลการศึกษาพบว่าค่าประมาณพารามิเตอร์ไม่เอนเอียงแต่องค์ประกอบความแปรปรวนมีแนวโน้มต่ำกว่าค่าจริง เมื่อจำนวนกลุ่มตัวอย่างมีขนาดเล็ก และยังพบว่าจำนวนกลุ่มตัวอย่างมีอิทธิพลต่อความถูกต้องของค่าประมาณพารามิเตอร์มากกว่าขนาดกลุ่มตัวอย่าง และถ้าสัดส่วนผลตอบสนองมีค่าน้อย ควรเพิ่มจำนวนกลุ่มตัวอย่างหรือขนาดกลุ่มตัวอย่าง และควรใช้จำนวนกลุ่มตัวอย่างเท่ากับ 100 และขนาดกลุ่มตัวอย่างเท่ากับ 50 จึงมีความถูกต้องของค่าประมาณพารามิเตอร์มากที่สุด นอกจากนี้ยังพบว่าเมื่อขนาดตัวอย่างเพิ่มขึ้นส่งผลให้อัตราการลู่เข้าหาคำตอบสูงขึ้น

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

Parsons et al. (2009) ได้ศึกษาการวิเคราะห์ข้อมูลที่มีความสัมพันธ์กันโดยใช้ตัวแบบ Proportional odds และใช้วิธี Generalized Estimating Equations (GEE) ในการประมาณพารามิเตอร์ของตัวแบบ และทำการเปรียบเทียบค่าประมาณพารามิเตอร์ เมื่อโครงสร้างสหสัมพันธ์ของข้อมูลภายในกลุ่มแตกต่างกัน 3 แบบ ประกอบด้วย Independence, Uniform และ AR (1) ผลการศึกษา พบว่าค่าประมาณพารามิเตอร์แตกต่างกันเพียงเล็กน้อย แสดงว่าค่าประมาณพารามิเตอร์จากวิธี GEE นั้นมั่นคง (Robust) ต่อการกำหนดโครงสร้างสหสัมพันธ์

บทที่ 3

วิธีดำเนินการวิจัย

งานวิจัยนี้เป็นการศึกษาเชิงประจักษ์ โดยการจำลองแบบข้อมูลเพื่อเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ และศึกษาผลกระทบของขนาดตัวอย่างและสหสัมพันธ์ของข้อมูลในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์

1. ข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่นำมาใช้ในการวิจัยเป็นข้อมูลจากการจำลองแบบ โดยมีตัวแปรแปรตอบสนองแบบอันดับที่มี 3 ระดับ ตัวแปรอธิบายแบบเบอร์นูลลี ที่มีความน่าจะเป็น เท่ากับ 0.5 ภายใต้เงื่อนไขของจำนวนกลุ่มตัวอย่าง (50, 100, 250) ขนาดกลุ่มตัวอย่าง (5, 7, 10) และสหสัมพันธ์ในกลุ่ม (0.05, 0.20, 0.40)

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2. เครื่องมือที่ใช้ในการวิจัย

โปรแกรมสำเร็จรูป SAS® version 9.1.3

3. สถิติที่ใช้ในการวิจัย

ในการประเมินประสิทธิภาพของวิธีประมาณพารามิเตอร์แต่ละวิธี จะพิจารณาจากสถิติดังต่อไปนี้

3.1 ความเอนเอียง

ความเอนเอียง (*Bias*) ของค่าประมาณพารามิเตอร์ มีสูตรการคำนวณดังนี้

$$Bias = \frac{\sum_{m=1}^{M} \hat{\theta}_m}{M} - \theta$$

เมื่อ $\hat{\theta}_m$ แทนค่าประมาณพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อมูลจำลองในครั้งที่ m, $m = 1, \ldots, M$ ซึ่งวิธีประมาณพารามิเตอร์ที่ให้ค่า *Bias* น้อยกว่า จะเป็นวิธีที่ให้ค่าประมาณพารามิเตอร์ที่มีความถูกต้องมากกว่า

3.2 ความคลาดเคลื่อนกำลังสองเฉลี่ย

ความคลาดเคลื่อนกำลังสองเฉลี่ย (*MSE*) มีสูตรการคำนวณดังนี้

$$MSE = \frac{\sum_{m=1}^{M}(\hat{\theta}_m - \theta)^2}{M}$$

โดยที่วิธีประมาณพารามิเตอร์ที่ให้ค่า *MSE* น้อยกว่า จะเป็นวิธีที่ให้ค่าประมาณพารามิเตอร์ที่มีความถูกต้องมากกว่า

3.3 *Coverage*

Coverage คือ สัดส่วนของช่วงความเชื่อมั่น ซึ่งงานวิจัยนี้ศึกษาที่ระดับ 95% ที่คำนวณจากค่าประมาณพารามิเตอร์จากการจำลองข้อมูลทั้งหมด *M* ครั้ง ที่ครอบคลุมค่าพารามิเตอร์จริง วิธีการประมาณพารามิเตอร์ของตัวแบบที่ให้ค่า *Coverage* มากกว่า จะเป็นวิธีที่ให้ค่าประมาณพารามิเตอร์ที่มีความถูกต้องมากกว่า สูตรการคำนวณช่วงความเชื่อมั่น 95 % ของค่าประมาณพารามิเตอร์ สำหรับข้อมูลจำลองแบบครั้งที่ *m* มีดังนี้

$$\hat{\theta}_m \pm t_{0.025,df} \times s.e.(\hat{\theta}_m)$$

เมื่อ $t_{0.025,df}$ คือเปอร์เซ็นต์ไทล์ส่วนบนที่ 97.5% ของการแจกแจงแบบ *Student t*, *df* คือองศาความเป็นอิสระของการแจกแจงแบบ *Student t* ซึ่งคำนวณจากจำนวนกลุ่มตัวอย่าง (*N*) ลบด้วยจำนวนพารามิเตอร์ในตัวแบบที่ต้องประมาณค่า และ $s.e.(\hat{\theta}_m)$ คือความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์สำหรับข้อมูลจำลองแบบครั้งที่ *m*

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

4. ขั้นตอนการจำลองแบบข้อมูล

งานวิจัยนี้มีขั้นตอนการจำลองแบบข้อมูล ดังนี้

1. กำหนดจำนวนกลุ่ม (N) เท่ากับ 50, 100 และ 250 ขนาดกลุ่ม (n_i) เท่ากับ 5, 7 และ 10

2. กำหนดตัวแปรอิสระ มีการแจกแจงแบบเบอร์นูลลี จำนวน 2 ตัวแปร ที่มีความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจเท่ากับ 0.50

3. กำหนดตัวแปรตอบสนองแบบมีลำดับที่มีการแจกแจงแบบมัลติโนเมียล 3 กลุ่ม ซึ่งมีความน่าจะเป็นของกลุ่ม ดังนี้

4. กำหนดตัวแบบในการวิเคราะห์ข้อมูลเป็นตัวแบบลอจิตสะสม ซึ่งมีสมการดังนี้

$$\text{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, u_i)] = \alpha_k + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_i, \quad k = 1, \ldots, c-1$$

ซึ่งมีพารามิเตอร์ที่ต้องประมาณค่า คือ $\alpha_1, \alpha_2, \beta_1, \beta_2$ และ σ_u^2

5. สร้างตัวแปรแฝงแบบต่อเนื่อง (Y_{ij}^*) ตามสมการ

$$Y_{ij}^* = -\beta_1 x_{1ij} - \beta_2 x_{2ij} - u_i + \varepsilon_{ij}$$

โดยที่กำหนดค่าต่าง ๆ ในสมการดังนี้

5.1 ค่าพารามิเตอร์ $\beta_1 = 0.5, \beta_2 = 0.5$ (ซึ่งคล้ายคลึงกับงานวิจัยของ Spiess and Hamerle 2000)

5.2 อิทธิพลสุ่ม (u_i) สุ่มค่าอย่างอิสระกัน จากการแจกแจงแบบปกติ มีค่าเฉลี่ย 0 ความแปรปรวน σ_u^2 เท่ากับ 0.17, 0.82 และ 2.19 ซึ่งสอดคล้องกับสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05, 0.20 และ 0.40

5.3 ความคลาดเคลื่อน(ε_{ij}) สุ่มค่าอย่างอิสระกัน จากการแจกแจงแบบลอจิสติกมาตรฐาน ที่มีค่าเฉลี่ย 0 ความแปรปรวนเท่ากับ $\pi^2/3$

6. แปลงค่าตัวแปรแฝงแบบต่อเนื่อง (Y_{ij}^*) ให้เป็นตัวแปรตอบสนองแบบอันดับ (Y_{ij}) ที่มี 3 ระดับ โดยกำหนดค่าจุดตัด $\alpha_1 = -1$ และ $\alpha_2 = 1$ (เพื่อให้ได้ความน่าจะเป็นของระดับของตัวแปรตอบสนองแบบอันดับมีค่าไม่น้อยเกินไป ดังนี้ $\pi_1 = 0.42$, $\pi_2 = 0.33$, $\pi_3 = 0.25$) ตามความสัมพันธ์ ดังนี้

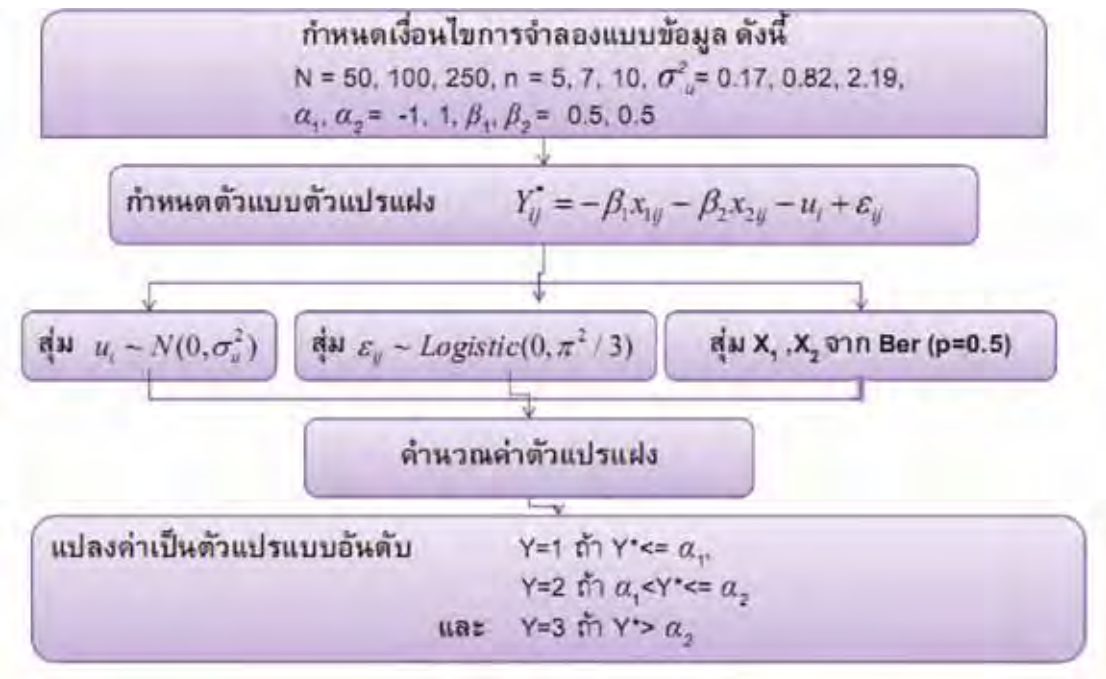
$Y_{ij} = 1$ ถ้า $Y_{ij}^* \le \alpha_1$

$Y_{ij} = 2$ ถ้า $\alpha_1 < Y_{ij}^* \le \alpha_2$

$Y_{ij} = 3$ ถ้า $Y_{ij}^* > \alpha_2$

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ดังแผนภูมิที่ 1



แผนภูมิที่ 1 ขั้นตอนการจำลองแบบข้อมูล

5. วิธีการวิเคราะห์ข้อมูล

1. จำลองแบบข้อมูลตามขั้นตอนการจำลองแบบข้อมูล ขั้นตอนที่ 1 – 6

2. สร้างตัวแบบลอจิตสะสมสำหรับอิทธิพลคงที่ และบันทึกค่าประมาณพารามิเตอร์ของตัวแบบพร้อมทั้งค่าคลาดเคลื่อนมาตรฐาน ด้วยวิธีประมาณพารามิเตอร์ของตัวแบบอิทธิพลคงที่ คือ วิธีภาวะน่าจะเป็นสูงสุด (ML) โดยใช้คำสั่ง PROC LOGISTIC

3. สร้างตัวแบบลอจิตสะสมแบบมาร์จินัล และบันทึกค่าประมาณพารามิเตอร์ของตัวแบบประกอบด้วยพร้อมทั้งค่าคลาดเคลื่อนมาตรฐาน ด้วยวิธี GEE และ MQL โดยใช้คำสั่ง PROC GENMOD และ PROC GLIMMIX ตามลำดับ

4. สร้างตัวแบบลอจิตสะสมแบบอิทธิพลสุ่ม และบันทึกค่าประมาณพารามิเตอร์ของตัวแบบประกอบด้วยพร้อมทั้งค่าคลาดเคลื่อนมาตรฐาน ด้วยวิธี AGQ และ PQL โดยใช้คำสั่ง PROC GLIMMIX และ PROC NLMIXED ตามลำดับ

5. คำนวณช่วงความเชื่อมั่น 95 % ของค่าประมาณพารามิเตอร์ต่าง ๆ ในขั้นตอนที่ 2 - 4

6. ทำซ้ำขั้นตอนที่ 1 – 4 จำนวน 1,000 ครั้ง สำหรับข้อมูลจำลองแบบแต่ละเงื่อนไข

7. คำนวณค่า Bias, MSE และ Coverage จากการวิเคราะห์ข้อมูลจำลองแบบ 1,000 ครั้ง

8. เปรียบเทียบค่า Bias, MSE และ Coverage ของค่าประมาณพารามิเตอร์จากแต่ละวิธี ดังแผนภูมิที่ 2

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

แผนภูมิที่ 2 ขั้นตอนการวิเคราะห์ข้อมูล

บทที่ 4

ผลการวิจัย

การเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ ผู้วิจัยได้ศึกษาการเปรียบเทียบวิธีประมาณพารามิเตอร์กรณีของการวิเคราะห์ด้วยตัวแบบมาร์จินัล 2 วิธี คือ Generalized Estimating Equations (GEE) และ Marginal Quasi-Likelihood (MQL) เปรียบเทียบกับวิธีภาวะน่าจะเป็นสูงสุด (ML) สำหรับตัวแบบอิทธิพลคงที่ และกรณีของการวิเคราะห์ด้วยตัวแบบอิทธิพลสุ่ม 2 วิธี คือ Adaptive Gaussian Quadrature (AGQ) และ Penalized Quasi-Likelihood (PQL) นอกจากนี้ยังศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง (N) ขนาดกลุ่มตัวอย่าง (n) และสหสัมพันธ์ในกลุ่ม (ICC) ที่มีต่อความถูกต้องของการประมาณพารามิเตอร์

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ในการเปรียบเทียบวิธีประมาณพารามิเตอร์ ผู้วิจัยทำการประเมินความถูกต้องของการประมาณพารามิเตอร์ต่าง ๆ ประกอบด้วยจุดตัด (α_1, α_2) สัมประสิทธิ์ของตัวแบบ (β_1, β_2) และองค์ประกอบความแปรปรวน (σ_u^2) โดยพิจารณาจากความเอนเอียง (Bias) ความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และสัดส่วนของช่วงความเชื่อมั่น 95% ของค่าประมาณพารามิเตอร์ที่ครอบคลุมค่าพารามิเตอร์จริง (Coverage) จากการจำลองข้อมูลตามเงื่อนไข ดังนี้ จำนวนกลุ่มตัวอย่าง เท่ากับ 50, 100 และ 250 ขนาดกลุ่มตัวอย่าง เท่ากับ 5, 7 และ 10 และสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05, 0.20 และ 0.40 เงื่อนไขละ 1,000 ครั้ง

ในการศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่าง และสหสัมพันธ์ในกลุ่ม ที่มีต่อความถูกต้องของการประมาณพารามิเตอร์ ผู้วิจัยได้ทำการศึกษาวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลและตัวแบบอิทธิพลสุ่ม โดยแยกพิจารณาทีละพารามิเตอร์ และเสนอผลการวิเคราะห์ข้อมูล ตามลำดับ ดังนี้

1. การเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัล

ตารางที่ 2 ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.05

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	GEE	MQL	ML	GEE	MQL
50	5	α_1	-1.0	-0.900	-0.900	-0.899	0.219	0.221	0.226
		α_2	1.0	0.884	0.884	0.884	0.220	0.223	0.227
		β_1	0.5	0.451	0.451	0.451	0.239	0.236	0.242
		β_2	0.5	0.452	0.452	0.451	0.239	0.236	0.242
	7	α_1	-1.0	-0.893	-0.893	-0.893	0.185	0.188	0.193
		α_2	1.0	0.879	0.879	0.880	0.185	0.190	0.193
		β_1	0.5	0.459	0.459	0.461	0.202	0.199	0.204
		β_2	0.5	0.450	0.450	0.448	0.202	0.199	0.204
	10	α_1	-1.0	-0.895	-0.895	-0.891	0.155	0.159	0.163
		α_2	1.0	0.880	0.880	0.884	0.155	0.161	0.163
		β_1	0.5	0.452	0.452	0.450	0.169	0.165	0.171
		β_2	0.5	0.462	0.462	0.459	0.169	0.167	0.171
100	5	α_1	-1.0	-0.892	-0.892	-0.891	0.154	0.157	0.159
		α_2	1.0	0.885	0.885	0.886	0.155	0.158	0.160
		β_1	0.5	0.445	0.445	0.446	0.168	0.167	0.171
		β_2	0.5	0.454	0.454	0.454	0.169	0.167	0.171
	7	α_1	-1.0	-0.894	-0.894	-0.894	0.130	0.134	0.135
		α_2	1.0	0.877	0.877	0.878	0.131	0.135	0.136
		β_1	0.5	0.457	0.457	0.456	0.142	0.141	0.144
		β_2	0.5	0.452	0.452	0.451	0.142	0.141	0.144
	10	α_1	-1.0	-0.884	-0.884	-0.884	0.109	0.113	0.115
		α_2	1.0	0.882	0.882	0.882	0.109	0.115	0.115
		β_1	0.5	0.448	0.448	0.448	0.119	0.118	0.120
		β_2	0.5	0.451	0.451	0.451	0.119	0.118	0.120

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 2 (ต่อ)

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	GEE	MQL	ML	GEE	MQL
250	5	α_1	-1.0	-0.884	-0.884	-0.884	0.097	0.099	0.100
		α_2	1.0	0.883	0.883	0.883	0.098	0.100	0.101
		β_1	0.5	0.444	0.444	0.445	0.106	0.106	0.108
		β_2	0.5	0.449	0.449	0.449	0.106	0.106	0.108
	7	α_1	-1.0	-0.889	-0.889	-0.889	0.082	0.085	0.086
		α_2	1.0	0.878	0.878	0.878	0.083	0.085	0.086
		β_1	0.5	0.451	0.451	0.452	0.090	0.090	0.091
		β_2	0.5	0.451	0.451	0.451	0.090	0.089	0.091
	10	α_1	-1.0	-0.882	-0.882	-0.876	0.069	0.072	0.073
		α_2	1.0	0.882	0.882	0.889	0.069	0.073	0.074
		β_1	0.5	0.447	0.447	0.443	0.075	0.075	0.076
		β_2	0.5	0.451	0.451	0.446	0.075	0.075	0.076

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 2 เมื่อพิจารณาค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐานที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ (วิธี ML) และตัวแบบมาร์จินัล (วิธี GEE และวิธี MQL) ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05 พบว่าวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลทั้ง 2 วิธี ให้ค่าประมาณพารามิเตอร์ที่ใกล้เคียงกัน และใกล้เคียงกับวิธี ML โดยทั้ง 3 วิธีให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริงเพียงเล็กน้อย และวิธี ML เป็นวิธีที่ให้ค่าประมาณที่มีค่าคลาดเคลื่อนมาตรฐานต่ำที่สุด

ตารางที่ 3 ค่าเอนเอียง (Bias) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่า Coverage (%) จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.05

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	GEE	MQL	ML	GEE	MQL	ML	GEE	MQL
50	5	α_1	0.100	0.100	0.101	0.060	0.060	0.060	92.40	92.00	93.90
		α_2	-0.116	-0.116	-0.116	0.063	0.063	0.063	90.40	90.00	92.30
		β_1	-0.049	-0.049	-0.049	0.058	0.058	0.057	94.60	94.10	95.30
		β_2	-0.048	-0.048	-0.049	0.063	0.063	0.063	94.10	93.10	94.30
	7	α_1	0.107	0.107	0.107	0.048	0.048	0.048	89.80	90.80	91.62
		α_2	-0.121	-0.121	-0.120	0.049	0.049	0.049	89.80	90.00	92.02
		β_1	-0.041	-0.041	-0.039	0.043	0.043	0.043	94.70	93.70	95.19
		β_2	-0.050	-0.050	-0.052	0.040	0.040	0.040	94.40	94.30	94.79
	10	α_1	0.105	0.105	0.109	0.039	0.039	0.039	87.60	88.70	90.56
		α_2	-0.120	-0.120	-0.116	0.042	0.042	0.041	85.90	87.20	89.57
		β_1	-0.048	-0.048	-0.050	0.031	0.031	0.030	93.90	92.90	93.85
		β_2	-0.038	-0.038	-0.041	0.029	0.029	0.029	95.30	94.60	95.06
100	5	α_1	0.108	0.108	0.109	0.036	0.036	0.036	88.40	88.50	89.42
		α_2	-0.115	-0.115	-0.114	0.037	0.037	0.037	88.60	89.20	90.60
		β_1	-0.055	-0.055	-0.054	0.030	0.030	0.030	94.40	94.10	95.30
		β_2	-0.046	-0.046	-0.046	0.030	0.030	0.031	94.00	93.90	94.12
	7	α_1	0.106	0.106	0.106	0.029	0.029	0.029	86.20	87.00	88.10
		α_2	-0.123	-0.123	-0.122	0.033	0.033	0.033	83.70	85.00	85.61
		β_1	-0.043	-0.043	-0.044	0.022	0.022	0.022	93.30	93.30	93.94
		β_2	-0.048	-0.048	-0.049	0.023	0.023	0.023	93.20	92.40	93.07
	10	α_1	0.116	0.116	0.116	0.027	0.027	0.027	78.50	79.60	82.40
		α_2	-0.118	-0.118	-0.118	0.028	0.028	0.028	78.40	80.90	81.80
		β_1	-0.052	-0.052	-0.052	0.017	0.017	0.017	93.10	92.60	93.80
		β_2	-0.049	-0.049	-0.049	0.016	0.016	0.016	92.50	91.90	92.70

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 3 (ต่อ)

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	GEE	MQL	ML	GEE	MQL	ML	GEE	MQL
250	5	α_1	0.116	0.116	0.116	0.024	0.024	0.024	75.70	76.80	78.10
		α_2	-0.117	-0.117	-0.117	0.024	0.024	0.024	76.30	77.60	79.00
		β_1	-0.056	-0.056	-0.055	0.014	0.014	0.014	92.00	91.60	92.30
		β_2	-0.051	-0.051	-0.051	0.014	0.014	0.014	92.80	92.40	93.40
	7	α_1	0.111	0.111	0.111	0.020	0.020	0.020	73.00	74.60	75.10
		α_2	-0.122	-0.122	-0.122	0.022	0.022	0.022	67.80	69.90	71.20
		β_1	-0.049	-0.049	-0.048	0.010	0.010	0.010	92.30	92.10	92.90
		β_2	-0.049	-0.049	-0.049	0.010	0.010	0.010	92.60	92.40	93.10
	10	α_1	0.118	0.118	0.124	0.019	0.019	0.021	57.20	60.20	60.96
		α_2	-0.118	-0.118	-0.111	0.020	0.020	0.018	59.70	63.80	69.81
		β_1	-0.053	-0.053	-0.057	0.008	0.008	0.009	89.80	89.80	90.96
		β_2	-0.049	-0.049	-0.054	0.008	0.008	0.009	89.10	88.70	90.19

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 3 เมื่อพิจาณาค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ย ที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ (วิธี ML) และตัวแบบมาร์จินัล (วิธี GEE และ วิธี MQL) ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05 พบว่าวิธีประมาณพารามิเตอร์ทั้ง 3 วิธี ให้ค่าประมาณที่มีค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกันมาก เมื่อพิจาณาสัดส่วนของช่วงความเชื่อมั่นที่ครอบคลุมค่าจริง (Coverage) พบว่าวิธี MQL มีแนวโน้มที่ให้ค่า Coverage สูงที่สุด รองลงมาคือวิธี GEE และวิธี ML ตามลำดับ และมีค่าใกล้เคียงกันมากเมื่อจำนวนกลุ่มตัวอย่างมีขนาดใหญ่

ตารางที่ 4 ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.20

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	GEE	MQL	ML	GEE	MQL
50	5	α_1	-1.0	-0.821	-0.821	-0.822	0.217	0.233	0.240
		α_2	1.0	0.810	0.810	0.810	0.218	0.236	0.241
		β_1	0.5	0.414	0.414	0.415	0.238	0.235	0.247
		β_2	0.5	0.411	0.411	0.412	0.238	0.235	0.247
	7	α_1	-1.0	-0.823	-0.823	-0.822	0.183	0.204	0.210
		α_2	1.0	0.798	0.798	0.798	0.184	0.206	0.210
		β_1	0.5	0.419	0.419	0.42	0.201	0.198	0.208
		β_2	0.5	0.416	0.416	0.414	0.201	0.198	0.208
	10	α_1	-1.0	-0.815	-0.815	-0.813	0.153	0.180	0.184
		α_2	1.0	0.805	0.805	0.806	0.154	0.183	0.184
		β_1	0.5	0.412	0.412	0.412	0.168	0.165	0.173
		β_2	0.5	0.421	0.421	0.419	0.168	0.166	0.173
100	5	α_1	-1.0	-0.814	-0.814	-0.813	0.153	0.165	0.170
		α_2	1.0	0.810	0.810	0.811	0.153	0.168	0.170
		β_1	0.5	0.405	0.405	0.404	0.168	0.167	0.174
		β_2	0.5	0.412	0.412	0.412	0.168	0.166	0.174
	7	α_1	-1.0	-0.819	-0.819	-0.819	0.129	0.145	0.148
		α_2	1.0	0.799	0.799	0.798	0.129	0.147	0.148
		β_1	0.5	0.416	0.416	0.417	0.142	0.141	0.147
		β_2	0.5	0.414	0.414	0.414	0.142	0.141	0.147
	10	α_1	-1.0	-0.804	-0.805	-0.803	0.108	0.127	0.130
		α_2	1.0	0.810	0.810	0.810	0.108	0.130	0.131
		β_1	0.5	0.406	0.406	0.406	0.118	0.117	0.122
		β_2	0.5	0.411	0.411	0.408	0.118	0.118	0.122

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 4 (ต่อ)

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	GEE	MQL	ML	GEE	MQL
250	5	α_1	-1.0	-0.808	-0.808	-0.802	0.097	0.105	0.107
		α_2	1.0	0.807	0.807	0.811	0.097	0.106	0.107
		β_1	0.5	0.406	0.406	0.398	0.106	0.106	0.110
		β_2	0.5	0.410	0.410	0.409	0.106	0.105	0.110
	7	α_1	-1.0	-0.811	-0.811	-0.812	0.082	0.092	0.094
		α_2	1.0	0.803	0.803	0.802	0.082	0.093	0.094
		β_1	0.5	0.411	0.411	0.411	0.089	0.089	0.093
		β_2	0.5	0.411	0.411	0.411	0.089	0.089	0.093
	10	α_1	-1.0	-0.805	-0.805	-0.805	0.068	0.081	0.082
		α_2	1.0	0.808	0.808	0.807	0.068	0.083	0.082
		β_1	0.5	0.407	0.407	0.408	0.075	0.075	0.077
		β_2	0.5	0.410	0.410	0.410	0.075	0.075	0.077

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 4 เมื่อพิจารณาค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐานที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ (วิธี ML) และตัวแบบมาร์จินัล (วิธี GEE และวิธี MQL) ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.20 พบว่าวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลทั้ง 2 วิธี ให้ค่าประมาณพารามิเตอร์ที่ใกล้เคียงกันและใกล้เคียงกับวิธี ML โดยทั้ง 3 วิธีให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริงเพียงเล็กน้อย และวิธี ML เป็นวิธีที่ให้ค่าประมาณที่มีค่าคลาดเคลื่อนมาตรฐานต่ำที่สุด

ตารางที่ 5 ค่าเอนเอียง (Bias) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่า Coverage (%) จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.20

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	GEE	MQL	ML	GEE	MQL	ML	GEE	MQL
50	5	α_1	0.179	0.179	0.178	0.087	0.087	0.085	85.00	87.20	90.15
		α_2	-0.190	-0.190	-0.190	0.094	0.094	0.093	82.40	85.60	87.34
		β_1	-0.086	-0.086	-0.085	0.063	0.063	0.059	93.60	92.00	94.87
		β_2	-0.089	-0.089	-0.088	0.069	0.069	0.066	92.40	91.10	93.97
	7	α_1	0.177	0.177	0.178	0.074	0.074	0.072	80.90	85.20	88.89
		α_2	-0.202	-0.202	-0.202	0.082	0.082	0.081	78.60	84.20	85.99
		β_1	-0.081	-0.081	-0.080	0.045	0.045	0.042	93.60	92.80	95.30
		β_2	-0.084	-0.084	-0.086	0.046	0.046	0.044	92.90	92.50	94.79
	10	α_1	0.185	0.185	0.187	0.069	0.069	0.067	72.10	79.80	83.94
		α_2	-0.195	-0.195	-0.194	0.072	0.072	0.070	70.30	80.20	83.23
		β_1	-0.088	-0.088	-0.088	0.035	0.035	0.032	91.30	90.60	93.37
		β_2	-0.079	-0.079	-0.081	0.033	0.033	0.031	92.70	92.40	94.38
100	5	α_1	0.186	0.186	0.187	0.062	0.062	0.061	74.00	77.90	81.30
		α_2	-0.190	-0.190	-0.189	0.065	0.065	0.064	73.10	78.50	80.40
		β_1	-0.095	-0.095	-0.096	0.038	0.038	0.036	91.80	91.30	94.00
		β_2	-0.088	-0.088	-0.088	0.036	0.036	0.035	91.60	90.30	93.40
	7	α_1	0.181	0.181	0.181	0.054	0.054	0.053	68.40	75.50	78.54
		α_2	-0.201	-0.201	-0.202	0.061	0.061	0.061	62.60	70.30	72.82
		β_1	-0.084	-0.084	-0.083	0.027	0.027	0.025	90.20	90.40	93.18
		β_2	-0.086	-0.086	-0.086	0.028	0.028	0.026	90.20	89.00	92.78
	10	α_1	0.196	0.195	0.197	0.056	0.056	0.055	52.80	63.00	67.02
		α_2	-0.190	-0.190	-0.190	0.054	0.054	0.052	54.70	67.40	69.58
		β_1	-0.094	-0.094	-0.094	0.023	0.023	0.021	87.00	87.30	90.50
		β_2	-0.089	-0.089	-0.092	0.022	0.022	0.020	88.50	87.10	90.93

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 5 (ต่อ)

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	GEE	MQL	ML	GEE	MQL	ML	GEE	MQL
250	5	α_1	0.192	0.192	0.198	0.048	0.048	0.050	48.40	53.90	54.29
		α_2	-0.193	-0.193	-0.189	0.050	0.050	0.048	47.90	54.40	57.62
		β_1	-0.094	-0.094	-0.102	0.021	0.021	0.021	85.40	84.60	86.83
		β_2	-0.090	-0.090	-0.091	0.019	0.019	0.019	86.40	85.70	89.68
	7	α_1	0.189	0.189	0.188	0.044	0.044	0.044	36.50	45.70	48.62
		α_2	-0.197	-0.197	-0.198	0.047	0.047	0.047	35.30	43.50	44.00
		β_1	-0.089	-0.089	-0.089	0.015	0.015	0.015	83.80	83.10	86.05
		β_2	-0.089	-0.089	-0.089	0.016	0.016	0.015	82.60	82.60	86.15
	10	α_1	0.195	0.195	0.195	0.045	0.045	0.044	22.80	32.80	34.50
		α_2	-0.192	-0.192	-0.193	0.044	0.044	0.044	25.20	36.50	36.70
		β_1	-0.093	-0.093	-0.092	0.014	0.014	0.013	76.10	76.40	80.00
		β_2	-0.090	-0.090	-0.090	0.014	0.014	0.013	76.80	77.10	79.30

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 5 เมื่อพิจาณาค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ย ที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ (วิธี ML) และตัวแบบมาร์จินัล (วิธี GEE และ วิธี MQL) ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.20 พบว่าวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลทั้ง 2 วิธี ให้ค่าประมาณที่มีค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกัน โดยวิธี MQL มีแนวโน้มที่ให้ค่าประมาณพารามิเตอร์ที่มีค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธี GEE และ ML เมื่อพิจาณาสัดส่วนของช่วงความเชื่อมั่นที่ครอบคลุมค่าจริง (Coverage) พบว่าวิธี MQL ให้ค่า Coverage ที่สูงที่สุด รองลงมาคือวิธี GEE และวิธี ML ตามลำดับ

ตารางที่ 6 ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และความคลาดเคลื่อนมาตรฐาน จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.40

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	GEE	MQL	ML	GEE	MQL
50	5	α_1	-1.0	-0.714	-0.714	-0.714	0.215	0.250	0.258
		α_2	1.0	0.698	0.698	0.697	0.216	0.253	0.259
		β_1	0.5	0.356	0.356	0.357	0.238	0.234	0.251
		β_2	0.5	0.354	0.354	0.356	0.238	0.235	0.251
	7	α_1	-1.0	-0.712	-0.712	-0.715	0.182	0.225	0.231
		α_2	1.0	0.686	0.686	0.684	0.182	0.228	0.231
		β_1	0.5	0.365	0.365	0.368	0.200	0.198	0.210
		β_2	0.5	0.352	0.352	0.352	0.200	0.198	0.210
	10	α_1	-1.0	-0.703	-0.703	-0.703	0.152	0.204	0.208
		α_2	1.0	0.697	0.697	0.700	0.152	0.208	0.208
		β_1	0.5	0.359	0.359	0.359	0.167	0.164	0.174
		β_2	0.5	0.362	0.362	0.360	0.167	0.166	0.174
100	5	α_1	-1.0	-0.705	-0.705	-0.705	0.152	0.177	0.182
		α_2	1.0	0.698	0.698	0.698	0.152	0.179	0.183
		β_1	0.5	0.348	0.348	0.349	0.167	0.166	0.177
		β_2	0.5	0.356	0.356	0.356	0.167	0.166	0.177
	7	α_1	-1.0	-0.707	-0.707	-0.708	0.128	0.160	0.163
		α_2	1.0	0.688	0.688	0.686	0.128	0.162	0.163
		β_1	0.5	0.361	0.361	0.362	0.141	0.140	0.148
		β_2	0.5	0.352	0.352	0.353	0.141	0.141	0.148
	10	α_1	-1.0	-0.691	-0.691	-0.691	0.107	0.145	0.147
		α_2	1.0	0.701	0.701	0.701	0.107	0.148	0.147
		β_1	0.5	0.350	0.350	0.351	0.118	0.117	0.123
		β_2	0.5	0.354	0.354	0.353	0.118	0.118	0.123

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 6 (ต่อ)

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	GEE	MQL	ML	GEE	MQL
250	5	α_1	-1.0	-0.697	-0.697	-0.697	0.096	0.112	0.115
		α_2	1.0	0.696	0.696	0.696	0.096	0.113	0.115
		β_1	0.5	0.349	0.349	0.350	0.105	0.105	0.112
		β_2	0.5	0.353	0.353	0.352	0.105	0.105	0.112
	7	α_1	-1.0	-0.699	-0.699	-0.699	0.081	0.101	0.103
		α_2	1.0	0.692	0.692	0.692	0.081	0.103	0.103
		β_1	0.5	0.352	0.352	0.353	0.089	0.089	0.094
		β_2	0.5	0.352	0.352	0.353	0.089	0.089	0.094
	10	α_1	-1.0	-0.692	-0.692	-0.685	0.068	0.092	0.094
		α_2	1.0	0.696	0.696	0.695	0.068	0.094	0.094
		β_1	0.5	0.350	0.350	0.350	0.075	0.074	0.077
		β_2	0.5	0.353	0.353	0.349	0.075	0.075	0.077

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 6 เมื่อพิจารณาค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐานที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ (วิธี ML) และตัวแบบมาร์จินัล (วิธี GEE และวิธี MQL) ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.40 พบว่าวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลทั้ง 2 วิธี ให้ค่าประมาณพารามิเตอร์ที่ใกล้เคียงกัน และใกล้เคียงกับวิธี ML โดยทั้ง 3 วิธีให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริงค่อนข้างมาก และวิธี ML เป็นวิธีที่ให้ค่าประมาณที่มีค่าคลาดเคลื่อนมาตรฐานต่ำที่สุด

ตารางที่ 7 ค่าเอนเอียง (Bias) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่า Coverage (%) จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบมาร์จินัล เมื่อกำหนด ICC = 0.40

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	GEE	MQL	ML	GEE	MQL	ML	GEE	MQL
50	5	α_1	0.286	0.286	0.286	0.145	0.145	0.139	69.70	77.80	82.80
		α_2	-0.302	-0.302	-0.303	0.160	0.160	0.153	65.90	75.20	80.77
		β_1	-0.144	-0.144	-0.143	0.077	0.077	0.066	90.60	89.50	95.41
		β_2	-0.146	-0.146	-0.144	0.082	0.082	0.070	89.20	88.90	94.66
	7	α_1	0.288	0.288	0.285	0.135	0.135	0.128	60.90	74.30	79.01
		α_2	-0.314	-0.314	-0.316	0.152	0.152	0.149	56.00	70.70	74.29
		β_1	-0.136	-0.135	-0.132	0.058	0.058	0.047	89.90	89.00	94.65
		β_2	-0.148	-0.148	-0.148	0.060	0.060	0.053	89.40	89.20	92.65
	10	α_1	0.297	0.297	0.297	0.136	0.136	0.132	47.80	67.70	71.27
		α_2	-0.303	-0.303	-0.300	0.137	0.137	0.132	50.10	67.50	70.54
		β_1	-0.141	-0.141	-0.141	0.049	0.049	0.041	85.60	84.20	91.60
		β_2	-0.138	-0.138	-0.140	0.047	0.047	0.041	86.70	85.20	90.66
100	5	α_1	0.295	0.295	0.295	0.118	0.118	0.116	49.70	62.50	65.60
		α_2	-0.302	-0.302	-0.302	0.124	0.124	0.122	46.70	59.00	62.50
		β_1	-0.152	-0.152	-0.151	0.052	0.052	0.046	84.40	83.80	90.20
		β_2	-0.144	-0.144	-0.144	0.050	0.050	0.044	86.30	85.20	90.40
	7	α_1	0.293	0.293	0.292	0.112	0.112	0.109	39.60	54.30	59.20
		α_2	-0.312	-0.312	-0.314	0.123	0.123	0.122	34.20	50.30	52.00
		β_1	-0.139	-0.139	-0.138	0.039	0.039	0.035	83.30	82.50	88.20
		β_2	-0.148	-0.148	-0.147	0.042	0.042	0.037	81.00	81.50	88.30
	10	α_1	0.309	0.309	0.309	0.118	0.118	0.116	26.10	42.30	44.90
		α_2	-0.299	-0.299	-0.299	0.112	0.112	0.110	27.10	47.60	48.60
		β_1	-0.150	-0.150	-0.149	0.037	0.037	0.033	74.70	74.90	81.90
		β_2	-0.146	-0.146	-0.147	0.035	0.035	0.032	75.40	75.30	81.10

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 7 (ต่อ)

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	GEE	MQL	ML	GEE	MQL	ML	GEE	MQL
250	5	α_1	0.303	0.303	0.303	0.105	0.105	0.103	15.10	22.50	22.50
		α_2	-0.304	-0.304	-0.304	0.106	0.106	0.105	16.60	24.60	23.92
		β_1	-0.151	-0.151	-0.150	0.034	0.034	0.031	68.90	68.90	75.78
		β_2	-0.147	-0.147	-0.148	0.033	0.033	0.031	70.90	71.20	77.19
	7	α_1	0.301	0.301	0.301	0.100	0.100	0.100	7.80	14.40	15.30
		α_2	-0.308	-0.308	-0.308	0.105	0.105	0.104	6.10	14.90	14.27
		β_1	-0.148	-0.148	-0.147	0.029	0.029	0.027	63.30	63.20	68.89
		β_2	-0.148	-0.148	-0.147	0.030	0.030	0.028	61.40	60.80	66.74
	10	α_1	0.308	0.308	0.315	0.103	0.103	0.107	2.50	8.60	7.07
		α_2	-0.304	-0.304	-0.305	0.101	0.101	0.102	3.30	10.10	10.33
		β_1	-0.150	-0.150	-0.150	0.028	0.028	0.027	48.90	49.30	50.54
		β_2	-0.147	-0.147	-0.151	0.027	0.027	0.027	48.40	48.80	49.32

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 7 เมื่อพิจาณาค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลคงที่ (วิธี ML) และตัวแบบมาร์จินัล (วิธี GEE และ วิธี MQL) ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.40 พบว่าวิธีประมาณพารามิเตอร์สำหรับตัวแบบมาร์จินัลทั้ง 2 วิธี ให้ค่าประมาณพารามิเตอร์ที่มีค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยค่อนข้างสูง โดยวิธี MQL มีแนวโน้มที่ให้ค่าประมาณที่มีค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธี GEE และ ML เมื่อพิจาณาสัดส่วนของช่วงความเชื่อมั่นที่ครอบคลุมค่าจริง (Coverage) พบว่าวิธี MQL ให้ค่า Coverage สูงที่สุด รองลงมาคือวิธี GEE และวิธี ML ตามลำดับ

2. การเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่ม

ตารางที่ 8 ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.05

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	AGQ	PQL	ML	AGQ	PQL
50	5	α_1	-1.0	-0.900	-0.934	-0.923	0.219	0.234	0.230
		α_2	1.0	0.884	0.910	0.907	0.220	0.234	0.231
		β_1	0.5	0.451	0.467	0.459	0.239	0.247	0.244
		β_2	0.5	0.452	0.471	0.459	0.239	0.247	0.244
		σ_u^2	0.17		0.169	0.174		0.219	0.189
	7	α_1	-1.0	-0.892	-0.924	-0.914	0.154	0.198	0.196
		α_2	1.0	0.885	0.905	0.901	0.155	0.198	0.196
		β_1	0.5	0.445	0.474	0.469	0.168	0.208	0.206
		β_2	0.5	0.454	0.467	0.455	0.169	0.208	0.206
		σ_u^2	0.17		0.149	0.155		0.153	0.141
	10	α_1	-1.0	-0.884	-0.924	-0.919	0.097	0.167	0.166
		α_2	1.0	0.883	0.907	0.900	0.098	0.168	0.166
		β_1	0.5	0.444	0.467	0.459	0.106	0.173	0.172
		β_2	0.5	0.449	0.477	0.471	0.106	0.173	0.172
		σ_u^2	0.17		0.146	0.150		0.111	0.107
100	5	α_1	-1.0	-0.893	-0.921	-0.913	0.185	0.164	0.161
		α_2	1.0	0.879	0.912	0.906	0.185	0.164	0.162
		β_1	0.5	0.459	0.459	0.452	0.202	0.173	0.172
		β_2	0.5	0.450	0.467	0.462	0.202	0.173	0.172
		σ_u^2	0.17		0.148	0.147		0.145	0.127
	7	α_1	-1.0	-0.894	-0.924	-0.915	0.130	0.139	0.137
		α_2	1.0	0.877	0.902	0.896	0.131	0.139	0.137
		β_1	0.5	0.457	0.472	0.466	0.142	0.146	0.145
		β_2	0.5	0.452	0.467	0.457	0.142	0.146	0.145
		σ_u^2	0.17		0.144	0.143		0.104	0.094
	10	α_1	-1.0	-0.889	-0.912	-0.903	0.082	0.118	0.116
		α_2	1.0	0.878	0.909	0.900	0.083	0.118	0.117
		β_1	0.5	0.451	0.462	0.455	0.090	0.122	0.121
		β_2	0.5	0.451	0.464	0.458	0.090	0.122	0.121
		σ_u^2	0.17		0.144	0.143		0.076	0.070

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 8 (ต่อ)

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ความคลาดเคลื่อนมาตรฐาน		
				ML	AGQ	PQL	ML	AGQ	PQL
		α_1	-1.0	-0.895	-0.912	-0.907	0.153	0.103	0.101
		α_2	1.0	0.880	0.909	0.890	0.154	0.104	0.101
	5	β_1	0.5	0.452	0.458	0.455	0.168	0.109	0.108
		β_2	0.5	0.462	0.463	0.455	0.168	0.109	0.108
		σ_u^2	0.17		0.143	0.136		0.088	0.075
		α_1	-1.0	-0.884	-0.917	-0.907	0.108	0.088	0.086
		α_2	1.0	0.882	0.905	0.894	0.108	0.088	0.087
250	7	β_1	0.5	0.448	0.465	0.457	0.118	0.092	0.091
		β_2	0.5	0.451	0.465	0.457	0.118	0.092	0.091
		σ_u^2	0.17		0.143	0.138		0.064	0.057
		α_1	-1.0	-0.882	-0.910	-0.902	0.068	0.075	0.074
		α_2	1.0	0.882	0.909	0.905	0.068	0.075	0.074
	10	β_1	0.5	0.447	0.461	0.452	0.075	0.077	0.077
		β_2	0.5	0.451	0.464	0.459	0.075	0.077	0.077
		σ_u^2	0.17		0.143	0.138		0.048	0.046

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 8 เมื่อพิจาณาค่าประมาณพารามิเตอร์และความคลาดเคลื่อนมาตรฐานที่ได้จากวิธีประมาณพารามิเตอร์แต่ละวิธีสำหรับตัวแบบอิทธิพลสุ่ม ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05 พบว่าวิธีประมาณพารามิเตอร์ทั้ง 2 วิธี ให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริงเพียงเล็กน้อย สำหรับพารามิเตอร์จุดตัดและสัมประสิทธิ์ของตัวแบบ โดยวิธี AGQ ให้ค่าประมาณที่ใกล้เคียงค่าจริงมากกว่าวิธี PQL และมีค่าความคลาดเคลื่อนมาตรฐานสูงกว่าวิธี PQL เล็กน้อย แต่สำหรับการประมาณพารามิเตอร์องค์ประกอบความแปรปรวน พบว่าวิธี AGQ ให้ค่าประมาณที่ต่ำกว่าค่าจริงเพียงเล็กน้อย ขณะที่วิธี PQL ให้ค่าประมาณที่ต่ำกว่าค่าจริงเกือบทุกกรณี ยกเว้นกรณีที่จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่างมีขนาดเล็ก วิธี PQL จะให้ค่าประมาณที่สูงกว่าค่าจริงเล็กน้อย

ตารางที่ 9 ค่าเอนเอียง (Bias) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่า Coverage (%) จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC=0.05

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	AGQ	PQL	ML	AGQ	PQL	ML	AGQ	PQL
50	5	α_1	0.100	0.066	0.077	0.060	0.050	0.059	92.40	96.00	94.37
		α_2	-0.116	-0.090	-0.093	0.063	0.055	0.060	90.40	94.50	93.47
		β_1	-0.049	-0.033	-0.041	0.058	0.051	0.059	94.60	96.50	95.38
		β_2	-0.048	-0.029	-0.041	0.063	0.056	0.065	94.10	96.40	94.17
		σ_u^2		-0.001	0.004		0.032	0.027		74.60	80.60
	7	α_1	0.108	0.076	0.086	0.036	0.040	0.045	88.40	95.10	93.69
		α_2	-0.115	-0.095	-0.099	0.037	0.042	0.045	88.60	94.20	93.18
		β_1	-0.055	-0.026	-0.031	0.030	0.041	0.044	94.40	96.10	94.60
		β_2	-0.046	-0.033	-0.045	0.030	0.036	0.041	94.00	96.30	94.60
		σ_u^2		-0.021	-0.015		0.017	0.016		82.60	85.20
	10	α_1	0.116	0.076	0.081	0.024	0.034	0.035	75.70	93.10	92.75
		α_2	-0.117	-0.093	-0.100	0.024	0.037	0.038	76.30	91.90	90.73
		β_1	-0.056	-0.033	-0.041	0.014	0.029	0.030	92.00	95.20	93.94
		β_2	-0.051	-0.023	-0.029	0.014	0.028	0.029	92.80	96.20	95.12
		σ_u^2		-0.024	-0.020		0.012	0.011		88.80	73.40
100	5	α_1	0.107	0.079	0.087	0.048	0.029	0.032	89.80	93.40	91.67
		α_2	-0.121	-0.088	-0.094	0.049	0.031	0.034	89.80	94.50	92.62
		β_1	-0.041	-0.041	-0.048	0.043	0.028	0.030	94.70	95.80	94.94
		β_2	-0.050	-0.033	-0.038	0.040	0.028	0.031	94.40	95.90	94.30
		σ_u^2		-0.022	-0.023		0.016	0.014		83.40	82.70
	7	α_1	0.106	0.076	0.085	0.029	0.024	0.027	86.20	91.60	89.80
		α_2	-0.123	-0.098	-0.104	0.033	0.028	0.029	83.70	89.80	88.53
		β_1	-0.043	-0.028	-0.034	0.022	0.022	0.022	93.30	94.30	93.77
		β_2	-0.048	-0.033	-0.043	0.023	0.022	0.023	93.20	94.80	93.91
		σ_u^2		-0.026	-0.027		0.010	0.009		89.80	61.30
	10	α_1	0.111	0.088	0.097	0.020	0.022	0.024	73.00	87.90	86.17
		α_2	-0.122	-0.091	-0.100	0.022	0.023	0.025	67.80	86.70	85.40
		β_1	-0.049	-0.038	-0.045	0.010	0.016	0.017	92.30	94.60	94.12
		β_2	-0.049	-0.036	-0.042	0.010	0.016	0.016	92.60	94.00	93.36
		σ_u^2		-0.026	-0.027		0.006	0.006		90.70	82.30

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 9 (ต่อ)

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	AGQ	PQL	ML	AGQ	PQL	ML	AGQ	PQL
		α_1	0.185	0.088	0.093	0.069	0.018	0.020	72.10	86.40	81.89
		α_2	-0.195	-0.091	-0.110	0.072	0.019	0.023	70.30	85.20	80.00
	5	β_1	-0.088	-0.042	-0.045	0.035	0.013	0.014	91.30	93.50	91.89
		β_2	-0.079	-0.037	-0.045	0.033	0.013	0.013	92.70	94.30	94.59
		σ_u^2		-0.027	-0.034		0.008	0.007		90.00	29.10
		α_1	0.196	0.083	0.093	0.056	0.015	0.016	52.80	84.10	79.74
		α_2	-0.190	-0.095	-0.106	0.054	0.017	0.018	54.70	80.80	77.01
250	7	β_1	-0.094	-0.035	-0.043	0.023	0.009	0.010	87.00	94.00	92.82
		β_2	-0.089	-0.035	-0.043	0.022	0.010	0.010	88.50	94.30	94.40
		σ_u^2		-0.027	-0.032		0.005	0.004		91.30	58.30
		α_1	0.195	0.090	0.098	0.045	0.014	0.015	22.80	76.80	74.18
		α_2	-0.192	-0.091	-0.095	0.044	0.014	0.015	25.20	75.80	75.18
	10	β_1	-0.093	-0.039	-0.048	0.014	0.007	0.008	76.10	92.40	91.68
		β_2	-0.090	-0.036	-0.041	0.014	0.007	0.007	76.80	92.90	93.97
		σ_u^2		-0.027	-0.032		0.003	0.003		88.70	69.60

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 9 เมื่อพิจาณาค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่ม ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05 พบว่าวิธี AGQ ให้ค่าประมาณที่มีค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยสำหรับทุกพารามิเตอร์ต่ำกว่าวิธี PQL ทุกกรณี เมื่อพิจาณาสัดส่วนของช่วงความเชื่อมั่นที่ครอบคลุมค่าจริง (Coverage) พบว่าทั้ง 2 วิธี ให้ค่า Coverage สำหรับพารามิเตอร์จุดตัดและสัมประสิทธิ์ของตัวแบบค่อนข้างสูง แต่ให้ค่า Coverage สำหรับพารามิเตอร์องค์ประกอบความแปรปรวนค่อนข้างต่ำ โดยเฉพาะวิธี PQL ให้ค่า Coverage สำหรับพารามิเตอร์องค์ประกอบความแปรปรวนต่ำมาก (29.10%) เมื่อจำนวนกลุ่มตัวอย่างเท่ากับ 250 และขนาดกลุ่มตัวอย่างเท่ากับ 5 โดยที่วิธี AGQ ให้ค่า Coverage สูงกว่าวิธี PQL ทุกกรณี

ตารางที่ 10 ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐาน จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.20

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	AGQ	PQL	ML	AGQ	PQL
50	5	α_1	-1.0	-0.821	-0.938	-0.909	0.217	0.266	0.255
		α_2	1.0	0.810	0.919	0.894	0.218	0.266	0.255
		β_1	0.5	0.414	0.469	0.446	0.238	0.262	0.254
		β_2	0.5	0.411	0.468	0.442	0.238	0.262	0.254
		σ_u^2	0.82		0.706	0.639		0.350	0.289
	7	α_1	-1.0	-0.814	-0.937	-0.912	0.153	0.232	0.223
		α_2	1.0	0.810	0.909	0.887	0.153	0.231	0.223
		β_1	0.5	0.405	0.475	0.456	0.168	0.219	0.213
		β_2	0.5	0.412	0.469	0.446	0.168	0.219	0.213
		σ_u^2	0.82		0.697	0.647		0.282	0.245
	10	α_1	-1.0	-0.808	-0.928	-0.911	0.097	0.203	0.197
		α_2	1.0	0.807	0.918	0.900	0.097	0.203	0.197
		β_1	0.5	0.406	0.466	0.449	0.106	0.181	0.177
		β_2	0.5	0.410	0.474	0.457	0.106	0.181	0.177
		σ_u^2	0.82		0.694	0.658		0.237	0.216
100	5	α_1	-1.0	-0.823	-0.928	-0.899	0.183	0.188	0.179
		α_2	1.0	0.798	0.922	0.893	0.184	0.188	0.179
		β_1	0.5	0.419	0.459	0.435	0.201	0.184	0.179
		β_2	0.5	0.416	0.467	0.442	0.201	0.184	0.179
		σ_u^2	0.82		0.698	0.614		0.244	0.198
	7	α_1	-1.0	-0.819	-0.934	-0.910	0.129	0.163	0.157
		α_2	1.0	0.799	0.910	0.887	0.129	0.163	0.157
		β_1	0.5	0.416	0.473	0.452	0.142	0.154	0.150
		β_2	0.5	0.414	0.468	0.448	0.142	0.154	0.150
		σ_u^2	0.82		0.692	0.628		0.198	0.171
	10	α_1	-1.0	-0.811	-0.919	-0.901	0.082	0.143	0.139
		α_2	1.0	0.803	0.923	0.904	0.082	0.143	0.139
		β_1	0.5	0.411	0.462	0.445	0.089	0.128	0.125
		β_2	0.5	0.411	0.465	0.449	0.089	0.128	0.125
		σ_u^2	0.82		0.693	0.645		0.167	0.151

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 10 (ต่อ)

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ความคลาดเคลื่อนมาตรฐาน		
				ML	AGQ	PQL	ML	AGQ	PQL
		α_1	-1.0	-0.815	-0.923	-0.895	0.153	0.118	0.113
		α_2	1.0	0.805	0.919	0.889	0.154	0.118	0.113
	5	β_1	0.5	0.412	0.461	0.436	0.168	0.116	0.113
		β_2	0.5	0.421	0.465	0.443	0.168	0.116	0.113
		σ_u^2	0.82		0.694	0.601		0.153	0.126
		α_1	-1.0	-0.804	-0.926	-0.905	0.108	0.103	0.101
		α_2	1.0	0.810	0.915	0.901	0.108	0.103	0.101
250	7	β_1	0.5	0.406	0.467	0.444	0.118	0.097	0.095
		β_2	0.5	0.411	0.466	0.450	0.118	0.097	0.095
		σ_u^2	0.82		0.694	0.621		0.125	0.115
		α_1	-1.0	-0.805	-0.920	-0.907	0.068	0.091	0.086
		α_2	1.0	0.808	0.921	0.884	0.068	0.091	0.086
	10	β_1	0.5	0.407	0.462	0.444	0.075	0.081	0.079
		β_2	0.5	0.410	0.464	0.449	0.075	0.081	0.079
		σ_u^2	0.82		0.696	0.640		0.106	0.088

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 10 เมื่อพิจาณาค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐานที่ได้จากวิธีประมาณพารามิเตอร์แต่ละวิธีสำหรับตัวแบบอิทธิพลสุ่ม ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.20 พบว่าวิธีประมาณพารามิเตอร์ทั้ง 2 วิธี ให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริงเพียงเล็กน้อย โดยวิธี AGQ ให้ค่าประมาณที่ใกล้เคียงค่าจริงมากกว่าวิธี PQL และมีค่าคลาดเคลื่อนมาตรฐานสูงกว่าวิธี PQL เล็กน้อย

ตารางที่ 11 ค่าเอนเอียง (Bias) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่า Coverage (%) จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.20

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	AGQ	PQL	ML	AGQ	PQL	ML	AGQ	PQL
50	5	α_1	0.179	0.062	0.091	0.087	0.074	0.070	85.00	95.30	95.01
		α_2	-0.190	-0.081	-0.106	0.094	0.077	0.073	82.40	94.00	93.57
		β_1	-0.086	-0.031	-0.054	0.063	0.068	0.062	93.60	95.20	95.01
		β_2	-0.089	-0.032	-0.058	0.069	0.076	0.069	92.40	94.20	94.01
		σ_u^2		-0.114	-0.181		0.137	0.116		87.80	72.30
	7	α_1	0.186	0.063	0.088	0.062	0.056	0.056	74.00	94.70	93.61
		α_2	-0.190	-0.091	-0.113	0.065	0.059	0.059	73.10	95.70	94.56
		β_1	-0.095	-0.025	-0.044	0.038	0.047	0.044	91.80	96.10	95.60
		β_2	-0.088	-0.031	-0.054	0.036	0.047	0.045	91.60	94.70	94.24
		σ_u^2		-0.123	-0.173		0.095	0.090		86.10	75.50
	10	α_1	0.192	0.072	0.089	0.048	0.046	0.047	48.40	94.80	93.67
		α_2	-0.193	-0.082	-0.100	0.050	0.048	0.049	47.90	92.70	92.14
		β_1	-0.094	-0.034	-0.051	0.021	0.031	0.031	85.40	96.10	95.10
		β_2	-0.090	-0.026	-0.043	0.019	0.032	0.031	86.40	96.00	95.00
		σ_u^2		-0.126	-0.162		0.073	0.073		85.10	78.70
100	5	α_1	0.177	0.072	0.101	0.074	0.039	0.041	80.90	93.50	91.05
		α_2	-0.202	-0.078	-0.107	0.082	0.041	0.043	78.60	92.70	91.05
		β_1	-0.081	-0.041	-0.065	0.045	0.035	0.034	93.60	95.40	94.77
		β_2	-0.084	-0.033	-0.058	0.046	0.036	0.035	92.90	94.50	93.56
		σ_u^2		-0.122	-0.206		0.074	0.082		87.40	73.70
	7	α_1	0.181	0.066	0.090	0.054	0.031	0.032	68.40	92.80	91.39
		α_2	-0.201	-0.090	-0.113	0.061	0.034	0.036	62.60	91.50	89.59
		β_1	-0.084	-0.027	-0.048	0.027	0.024	0.024	90.20	95.50	95.10
		β_2	-0.086	-0.032	-0.052	0.028	0.025	0.024	90.20	93.90	93.19
		σ_u^2		-0.128	-0.192		0.053	0.064		84.00	71.60
	10	α_1	0.189	0.081	0.099	0.044	0.027	0.029	36.50	90.90	89.59
		α_2	-0.197	-0.077	-0.096	0.047	0.027	0.029	35.30	91.90	90.09
		β_1	-0.089	-0.038	-0.055	0.015	0.017	0.017	83.80	94.50	93.43
		β_2	-0.089	-0.035	-0.051	0.016	0.017	0.017	82.60	94.60	94.03
		σ_u^2		-0.127	-0.175		0.042	0.052		83.60	72.50

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 11 (ต่อ)

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	AGQ	PQL	ML	AGQ	PQL	ML	AGQ	PQL
		α_1	0.185	0.077	0.105	0.069	0.020	0.023	72.10	89.30	84.64
		α_2	-0.195	-0.081	-0.111	0.072	0.022	0.026	70.30	87.90	82.54
	5	β_1	-0.088	-0.039	-0.064	0.035	0.016	0.017	91.30	92.90	91.27
		β_2	-0.079	-0.035	-0.057	0.033	0.015	0.016	92.70	93.60	92.38
		σ_u^2		-0.126	-0.219		0.040	0.064		82.40	53.20
		α_1	0.196	0.074	0.095	0.056	0.016	0.019	52.80	89.20	84.50
		α_2	-0.190	-0.085	-0.099	0.054	0.018	0.020	54.70	86.70	82.57
50	7	β_1	-0.094	-0.033	-0.056	0.023	0.010	0.012	87.00	93.10	91.04
		β_2	-0.089	-0.034	-0.050	0.022	0.011	0.011	88.50	93.90	90.80
		σ_u^2		-0.126	-0.199		0.031	0.050		79.10	38.00
		α_1	0.195	0.080	0.093	0.045	0.014	0.017	22.80	85.60	79.22
		α_2	-0.192	-0.079	-0.116	0.044	0.015	0.022	25.20	85.10	70.59
	10	β_1	-0.093	-0.038	-0.056	0.014	0.008	0.008	76.10	94.20	92.94
		β_2	-0.090	-0.036	-0.051	0.014	0.008	0.010	76.80	93.30	87.06
		σ_u^2		-0.124	-0.180		0.026	0.041		74.80	3.00

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 11 เมื่อพิจาณาค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่ม ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.20 พบว่าวิธี AGQ ให้ค่าประมาณที่มีค่าเอนเอียงต่ำกว่าวิธี PQL ค่อนข้างมาก แต่เมื่อพิจารณาค่าคลาดเคลื่อนกำลังสองเฉลี่ยทั้ง 2 วิธี มีค่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกัน โดยวิธี AGQ มีค่าต่ำกว่าวิธี PQL เกือบทุกกรณี ยกเว้นกรณีที่จำนวนกลุ่ม เท่ากับ 50 และขนาดกลุ่ม เท่ากับ 5 และ 7 เมื่อพิจาณาสัดส่วนของช่วงความเชื่อมั่นที่ครอบคลุมค่าจริง (Coverage) พบว่าทั้ง 2 วิธี ให้ค่า Coverage สำหรับพารามิเตอร์จุดตัดและสัมประสิทธิ์ของตัวแบบค่อนข้างสูง แต่ให้ค่า Coverage สำหรับพารามิเตอร์องค์ประกอบความแปรปรวนค่อนข้างต่ำ โดยเฉพาะวิธี PQL ให้ค่า Coverage สำหรับพารามิเตอร์องค์ประกอบความแปรปรวนต่ำมาก เมื่อจำนวนกลุ่มตัวอย่าง เท่ากับ 250 โดยที่วิธี AGQ ให้ค่า Coverage สูงกว่าวิธี PQL ทุกกรณี

ตารางที่ 12 ค่าเฉลี่ยของค่าประมาณพารามิเตอร์และความคลาดเคลื่อนมาตรฐาน จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.40

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ค่าคลาดเคลื่อนมาตรฐาน		
				ML	AGQ	PQL	ML	AGQ	PQL
50	5	α_1	-1.0	-0.714	-0.961	-0.920	0.215	0.324	0.299
		α_2	1.0	0.698	0.931	0.881	0.216	0.323	0.299
		β_1	0.5	0.356	0.472	0.431	0.238	0.283	0.269
		β_2	0.5	0.354	0.477	0.435	0.238	0.283	0.269
		σ_u^2	2.19		1.939	1.556		0.702	0.524
	7	α_1	-1.0	-0.705	-0.957	-0.928	0.152	0.290	0.273
		α_2	1.0	0.698	0.918	0.880	0.152	0.290	0.273
		β_1	0.5	0.348	0.485	0.453	0.167	0.234	0.225
		β_2	0.5	0.356	0.468	0.437	0.167	0.234	0.225
		σ_u^2	2.19		1.922	1.631		0.605	0.492
	10	α_1	-1.0	-0.697	-0.939	-0.920	0.096	0.263	0.251
		α_2	1.0	0.696	0.937	0.908	0.096	0.263	0.251
		β_1	0.5	0.349	0.475	0.451	0.105	0.192	0.187
		β_2	0.5	0.353	0.474	0.449	0.105	0.192	0.187
		σ_u^2	2.19		1.901	1.690		0.534	0.462
100	5	α_1	-1.0	-0.712	-0.947	-0.904	0.182	0.228	0.210
		α_2	1.0	0.686	0.936	0.884	0.182	0.228	0.210
		β_1	0.5	0.365	0.463	0.422	0.200	0.199	0.189
		β_2	0.5	0.352	0.473	0.432	0.200	0.199	0.189
		σ_u^2	2.19		1.908	1.506		0.488	0.363
	7	α_1	-1.0	-0.707	-0.952	-0.921	0.128	0.204	0.192
		α_2	1.0	0.688	0.921	0.880	0.128	0.204	0.191
		β_1	0.5	0.361	0.481	0.448	0.141	0.165	0.159
		β_2	0.5	0.352	0.469	0.437	0.141	0.165	0.159
		σ_u^2	2.19		1.901	1.590		0.422	0.339
	10	α_1	-1.0	-0.699	-0.928	-0.911	0.081	0.186	0.176
		α_2	1.0	0.692	0.941	0.907	0.081	0.186	0.176
		β_1	0.5	0.352	0.466	0.442	0.089	0.135	0.132
		β_2	0.5	0.352	0.467	0.446	0.089	0.135	0.132
		σ_u^2	2.19		1.891	1.658		0.375	0.319

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 12 (ต่อ)

N	n	พารามิเตอร์	ค่าจริง	ค่าประมาณ			ความคลาดเคลื่อนมาตรฐาน		
				ML	AGQ	PQL	ML	AGQ	PQL
		α_1	-1.0	-0.703	-0.939	-0.893	0.152	0.144	0.132
		α_2	1.0	0.697	0.935	0.882	0.152	0.144	0.132
	5	β_1	0.5	0.359	0.467	0.423	0.167	0.125	0.119
		β_2	0.5	0.362	0.469	0.425	0.167	0.125	0.119
		σ_u^2	2.19		1.902	1.484		0.307	0.222
		α_1	-1.0	-0.691	-0.943	-0.911	0.107	0.129	0.120
		α_2	1.0	0.701	0.929	0.888	0.107	0.129	0.120
250	7	β_1	0.5	0.350	0.471	0.438	0.118	0.104	0.100
		β_2	0.5	0.354	0.470	0.437	0.118	0.104	0.100
		σ_u^2	2.19		1.899	1.574		0.266	0.210
		α_1	-1.0	-0.692	-0.931	-0.911	0.068	0.117	0.111
		α_2	1.0	0.696	0.935	0.904	0.068	0.117	0.111
	10	β_1	0.5	0.350	0.467	0.443	0.075	0.086	0.083
		β_2	0.5	0.353	0.469	0.444	0.075	0.086	0.083
		σ_u^2	2.19		1.895	1.647		0.237	0.200

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 12 เมื่อพิจาณาค่าประมาณพารามิเตอร์และค่าคลาดเคลื่อนมาตรฐานที่ได้จากวิธีประมาณพารามิเตอร์แต่ละวิธีสำหรับตัวแบบอิทธิพลสุ่ม ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.40 พบว่าวิธีประมาณพารามิเตอร์ทั้ง 2 วิธี ให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริงเพียงเล็กน้อย โดยวิธี AGQ ให้ค่าประมาณที่ใกล้เคียงค่าจริงมากกว่าวิธี PQL และมีค่าค่าคลาดเคลื่อนมาตรฐานสูงกว่าวิธี PQL เล็กน้อย

ตารางที่ 13 ค่าเอนเอียง (Bias) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่า Coverage (%) จากการจำลองแบบข้อมูล 1,000 ครั้ง สำหรับตัวแบบอิทธิพลสุ่ม เมื่อกำหนด ICC = 0.40

N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	AGQ	PQL	ML	AGQ	PQL	ML	AGQ	PQL
50	5	α_1	0.286	0.039	0.080	0.145	0.106	0.092	69.70	95.80	95.45
		α_2	-0.302	-0.069	-0.119	0.160	0.111	0.101	65.90	95.10	94.64
		β_1	-0.144	-0.028	-0.069	0.077	0.079	0.069	90.60	95.10	95.45
		β_2	-0.146	-0.023	-0.065	0.082	0.088	0.077	89.20	95.00	93.72
		σ_u^2		-0.251	-0.634		0.559	0.645		85.80	68.10
	7	α_1	0.295	0.043	0.072	0.118	0.086	0.077	49.70	95.30	94.89
		α_2	-0.302	-0.082	-0.120	0.124	0.089	0.085	46.70	94.30	93.87
		β_1	-0.152	-0.015	-0.047	0.052	0.054	0.049	84.40	96.20	96.32
		β_2	-0.144	-0.032	-0.063	0.050	0.054	0.050	86.30	95.70	94.79
		σ_u^2		-0.268	-0.559		0.424	0.519		85.10	69.90
	10	α_1	0.303	0.061	0.080	0.105	0.077	0.071	15.10	94.30	93.81
		α_2	-0.304	-0.063	-0.092	0.106	0.078	0.074	16.60	93.80	93.51
		β_1	-0.151	-0.025	-0.049	0.034	0.036	0.034	68.90	96.60	95.84
		β_2	-0.147	-0.026	-0.051	0.033	0.038	0.036	70.90	95.10	94.42
		σ_u^2		-0.289	-0.500		0.368	0.442		84.00	71.50
100	5	α_1	0.288	0.053	0.096	0.135	0.052	0.049	60.90	94.70	92.56
		α_2	-0.314	-0.064	-0.116	0.152	0.057	0.057	56.00	93.50	92.44
		β_1	-0.136	-0.037	-0.078	0.058	0.040	0.038	89.90	95.50	95.11
		β_2	-0.148	-0.027	-0.068	0.060	0.041	0.038	89.40	94.00	93.89
		σ_u^2		-0.282	-0.684		0.317	0.582		84.60	49.10
	7	α_1	0.293	0.048	0.079	0.112	0.044	0.042	39.60	94.50	92.75
		α_2	-0.312	-0.079	-0.120	0.123	0.047	0.050	34.20	93.10	90.26
		β_1	-0.139	-0.019	-0.052	0.039	0.027	0.026	83.30	95.50	94.45
		β_2	-0.148	-0.031	-0.063	0.042	0.028	0.028	81.00	94.10	93.66
		σ_u^2		-0.289	-0.600		0.249	0.456		84.40	53.10
	10	α_1	0.301	0.072	0.089	0.100	0.041	0.039	7.80	92.50	91.43
		α_2	-0.308	-0.059	-0.093	0.105	0.039	0.040	6.10	93.60	91.43
		β_1	-0.148	-0.034	-0.058	0.029	0.019	0.019	63.30	94.40	94.17
		β_2	-0.148	-0.033	-0.054	0.030	0.020	0.019	61.40	93.90	93.94
		σ_u^2		-0.299	-0.532		0.227	0.374		80.30	46.00

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ตารางที่ 13 (ต่อ)

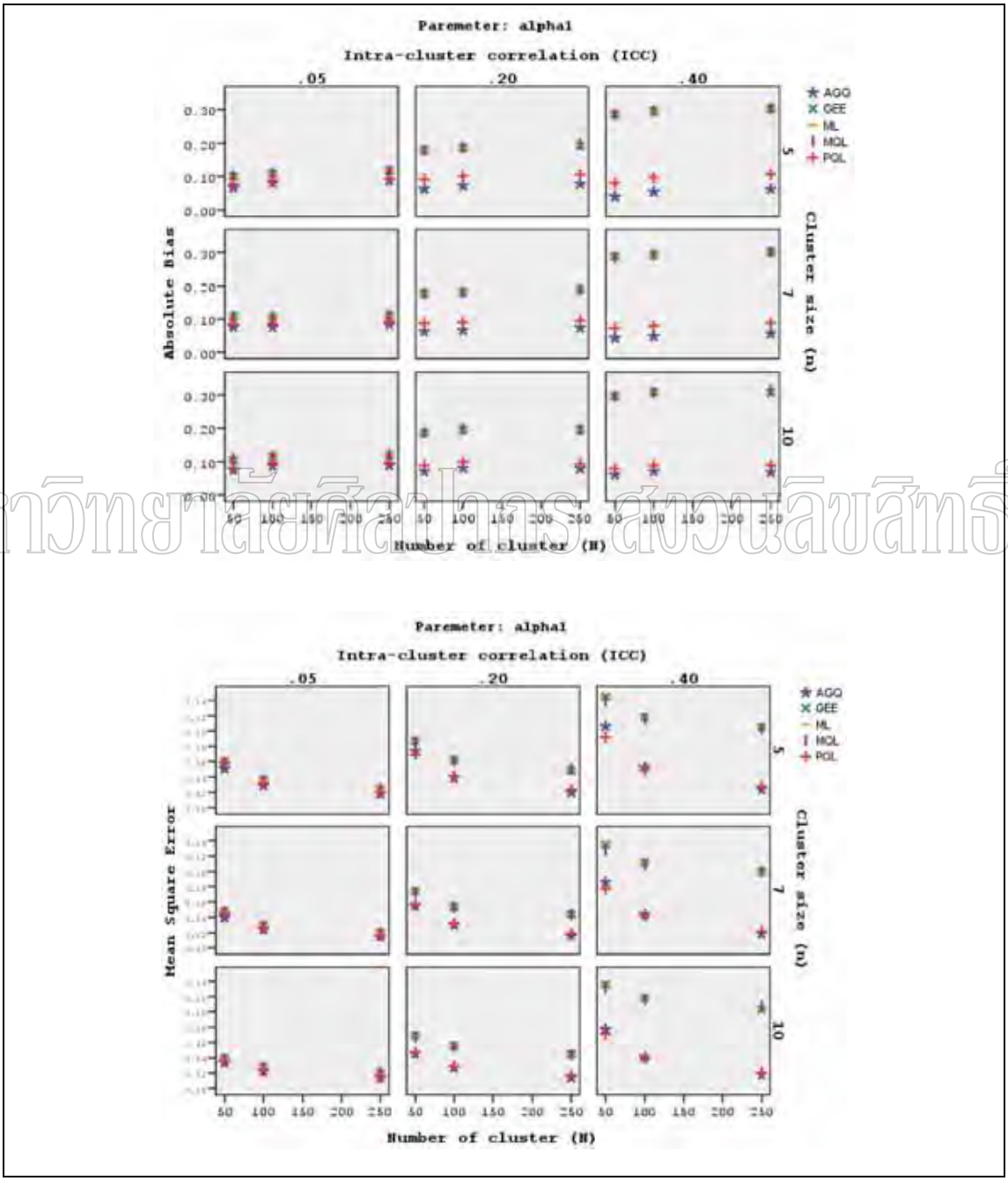
N	n	พารามิเตอร์	Bias			MSE			Coverage (%)		
			ML	AGQ	PQL	ML	AGQ	PQL	ML	AGQ	PQL
		α_1	0.297	0.061	0.107	0.136	0.024	0.028	47.80	92.90	87.89
		α_2	-0.303	-0.065	-0.118	0.137	0.026	0.032	50.10	91.50	85.59
	5	β_1	-0.141	-0.033	-0.077	0.049	0.016	0.018	85.60	94.40	91.19
		β_2	-0.138	-0.031	-0.075	0.047	0.017	0.019	86.70	93.50	92.29
		σ_u^2		-0.288	-0.706		0.179	0.544		79.50	15.20
		α_1	0.309	0.057	0.089	0.118	0.019	0.022	26.10	93.00	88.80
		α_2	-0.299	-0.071	-0.112	0.112	0.022	0.027	27.10	91.60	86.00
50	7	β_1	-0.150	-0.029	-0.062	0.037	0.011	0.013	74.70	94.20	91.20
		β_2	-0.146	-0.030	-0.063	0.035	0.012	0.013	75.40	94.70	91.40
		σ_u^2		-0.291	-0.616		0.148	0.415		77.10	18.00
		α_1	0.308	0.069	0.089	0.103	0.018	0.020	2.50	91.20	86.90
		α_2	-0.304	-0.065	-0.096	0.101	0.019	0.022	3.30	89.70	85.70
	10	β_1	-0.150	-0.033	-0.057	0.028	0.008	0.010	48.90	93.80	90.70
		β_2	-0.147	-0.031	-0.056	0.027	0.008	0.010	48.40	94.40	91.40
		σ_u^2		-0.295	-0.543		0.142	0.331		70.60	23.90

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากตารางที่ 13 เมื่อพิจาณาค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จากวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่ม ภายใต้จำนวนกลุ่มตัวอย่าง (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) เมื่อกำหนดสหสัมพันธ์ในกลุ่ม เท่ากับ 0.40 พบว่าวิธี AGQ ให้ค่าประมาณที่มีค่าเอนเอียงต่ำกว่าวิธี PQL ค่อนข้างมาก แต่เมื่อพิจารณาค่าคลาดเคลื่อนกำลังสองเฉลี่ยทั้ง 2 วิธี มีค่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกัน โดยวิธี AGQ มีค่าสูงกว่า สำหรับจำนวนกลุ่มตัวอย่างเท่ากับ 50 และ 100 แต่มีค่าต่ำกว่า สำหรับจำนวนกลุ่มตัวอย่าง เท่ากับ 250 และเมื่อพิจาณาสัดส่วนของช่วงความเชื่อมั่นที่ครอบคลุมค่าจริง (Coverage) พบว่าทั้ง 2 วิธี ให้ค่า Coverage สำหรับพารามิเตอร์จุดตัดและสัมประสิทธิ์ของตัวแบบค่อนข้างสูง แต่ให้ค่า Coverage สำหรับพารามิเตอร์องค์ประกอบความแปรปรวนค่อนข้างต่ำ โดยเฉพาะวิธี PQL ให้ค่า Coverage สำหรับพารามิเตอร์องค์ประกอบความแปรปรวนต่ำมาก เมื่อจำนวนกลุ่มตัวอย่าง เท่ากับ 250 โดยที่วิธี AGQ ให้ค่า Coverage สูงกว่าวิธี PQL ทุกกรณี

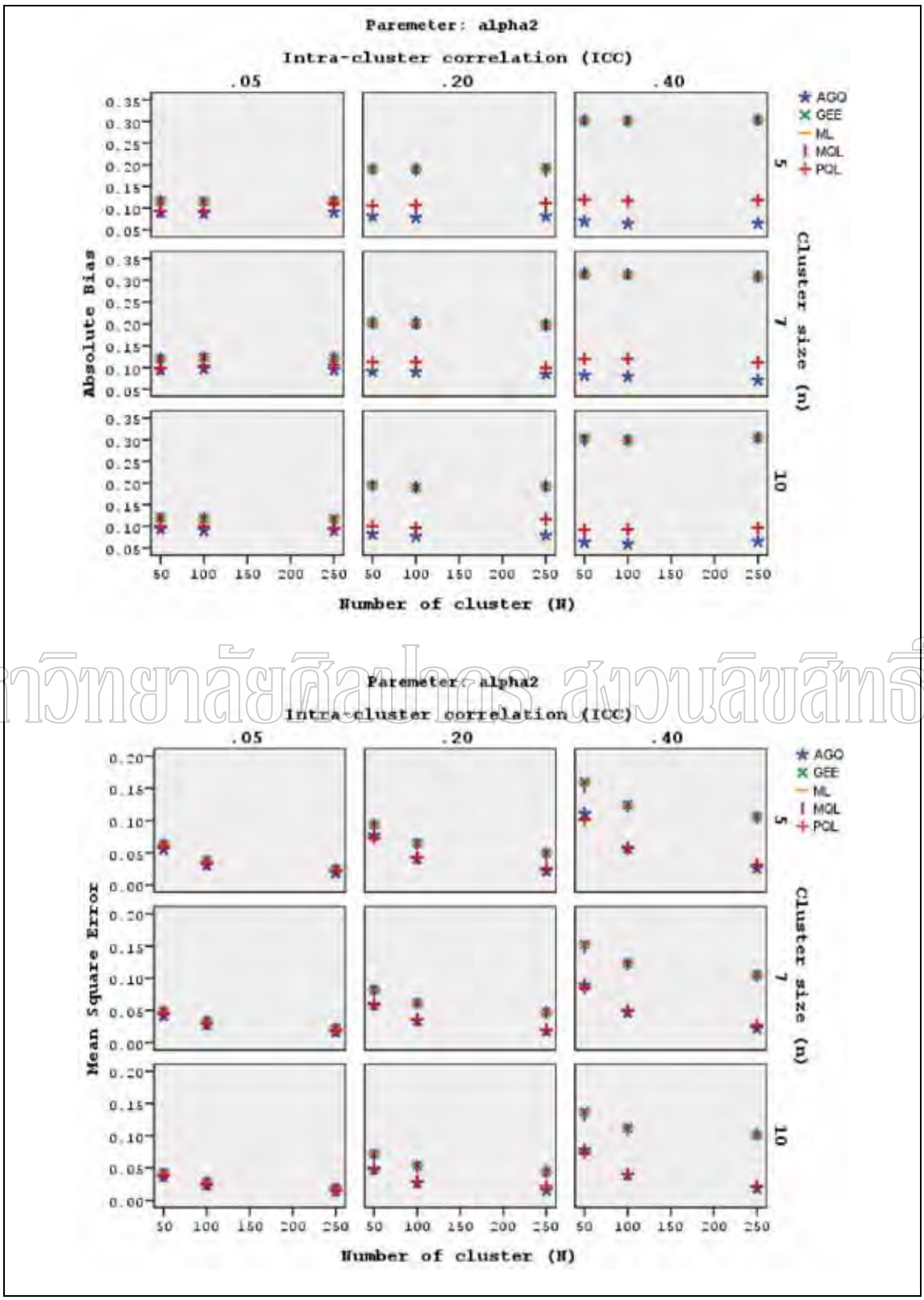
3. การศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่าง และสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์

3.1 การศึกษาผลกระทบของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องของการประมาณพารามิเตอร์



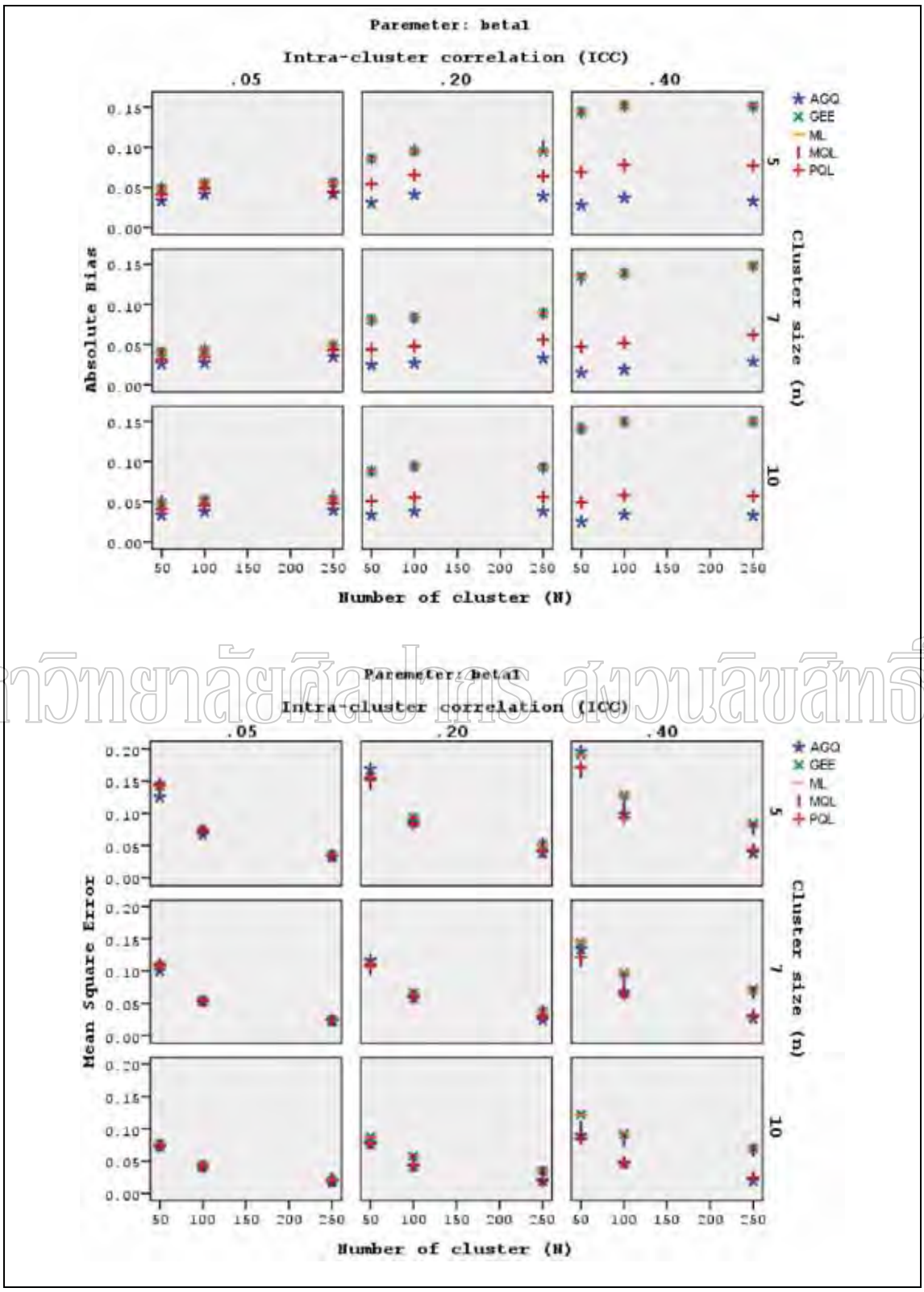
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 1 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_1



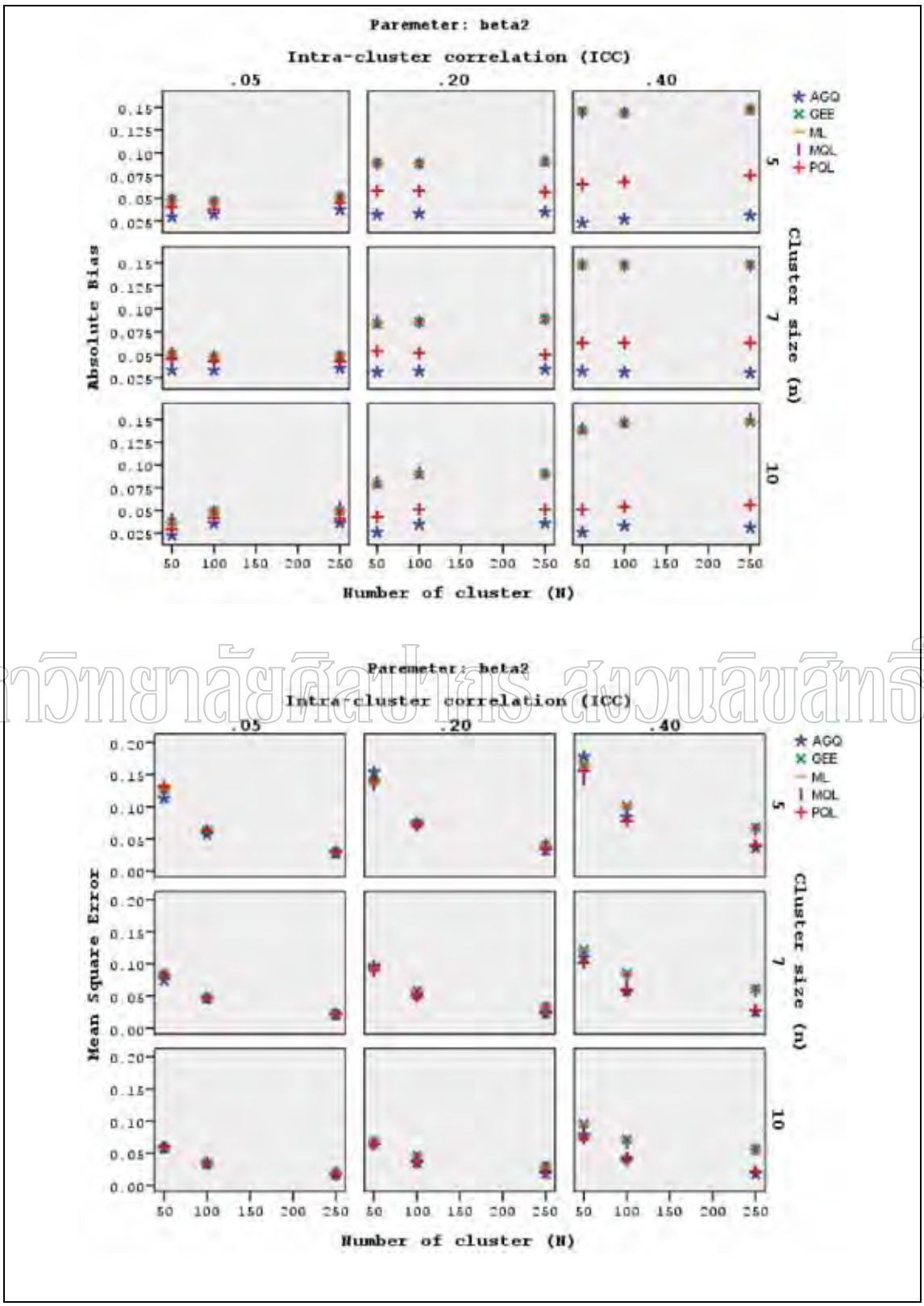
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 2 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_2



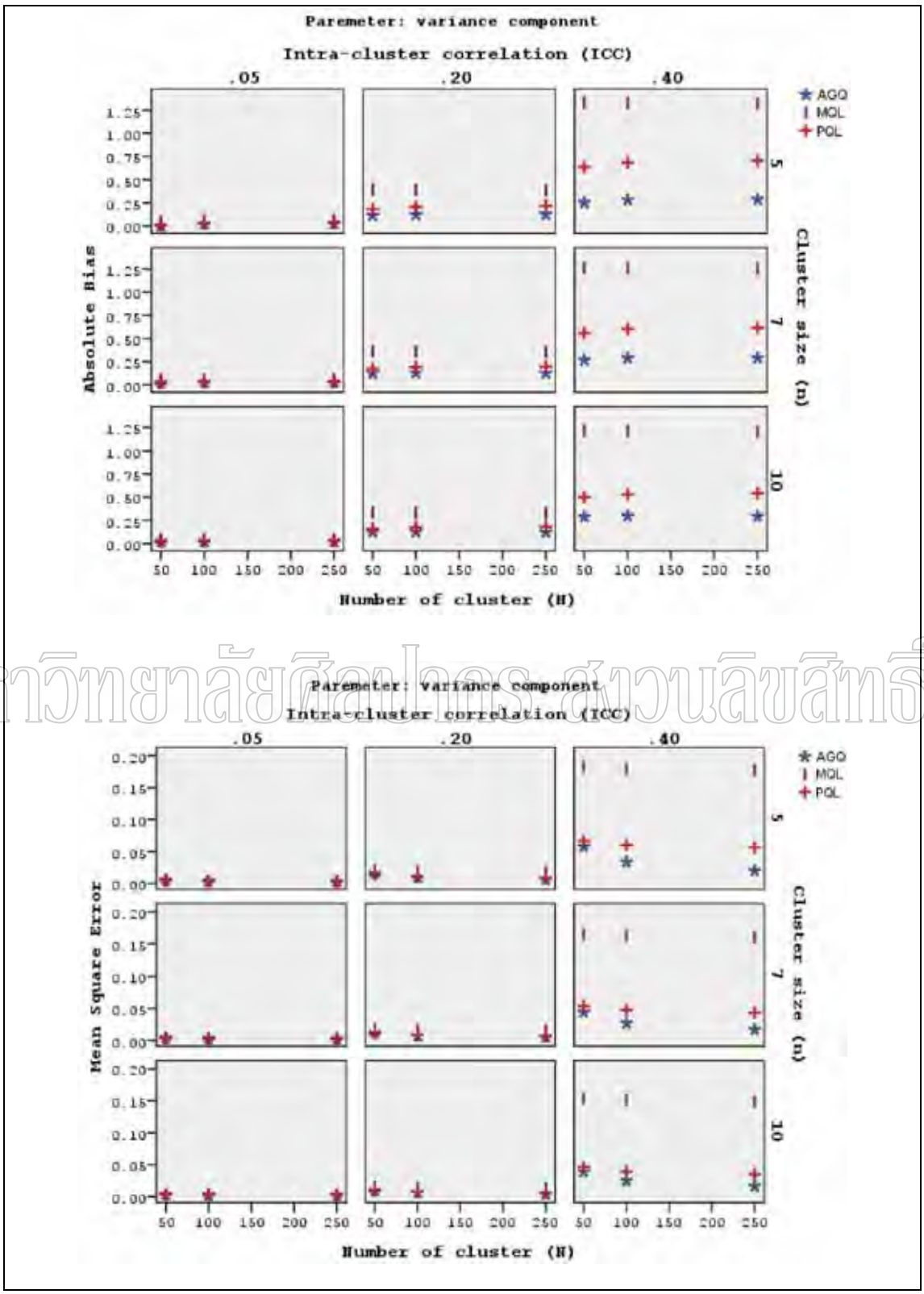
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 3 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_1



มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

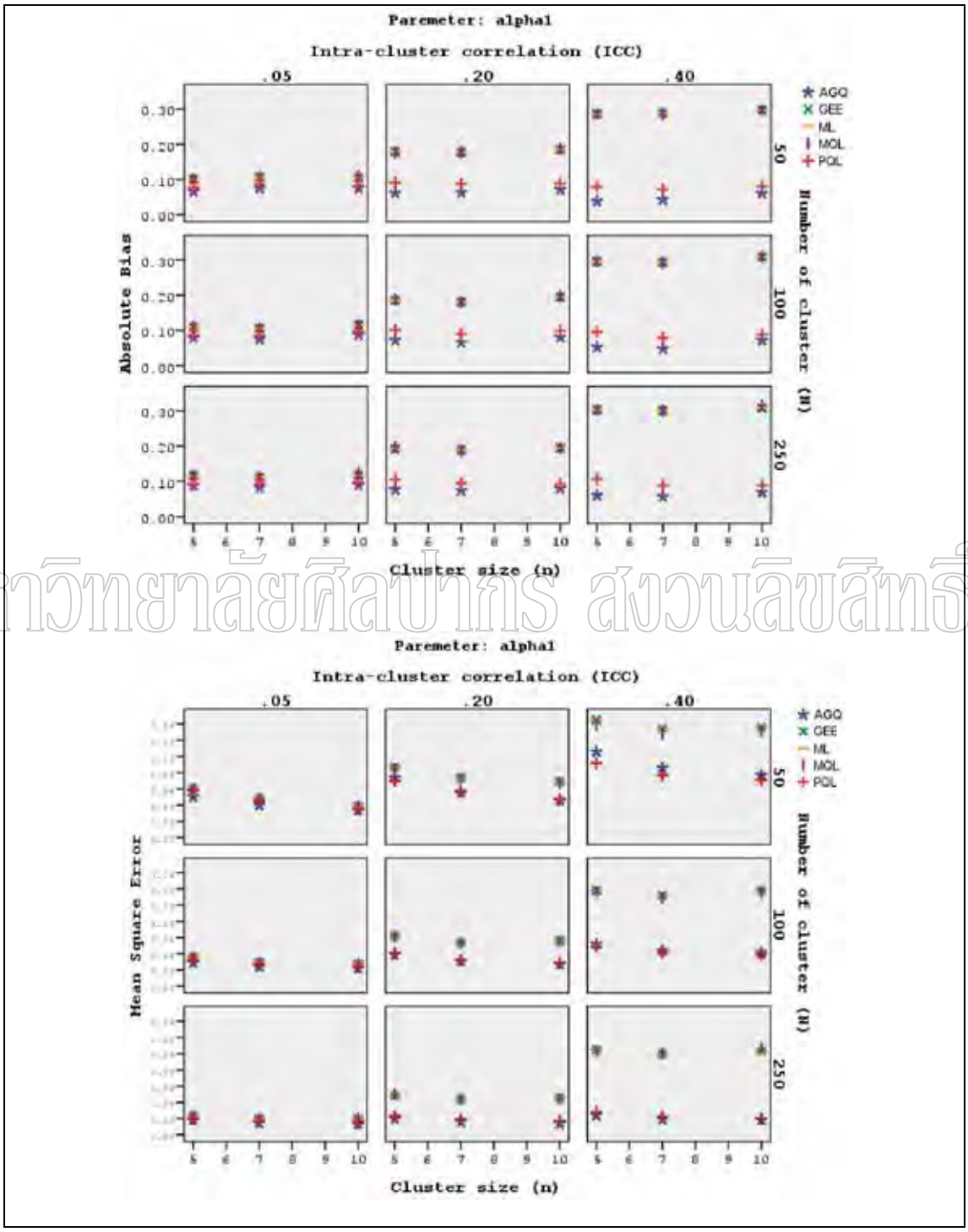
ภาพที่ 4 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_2



มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

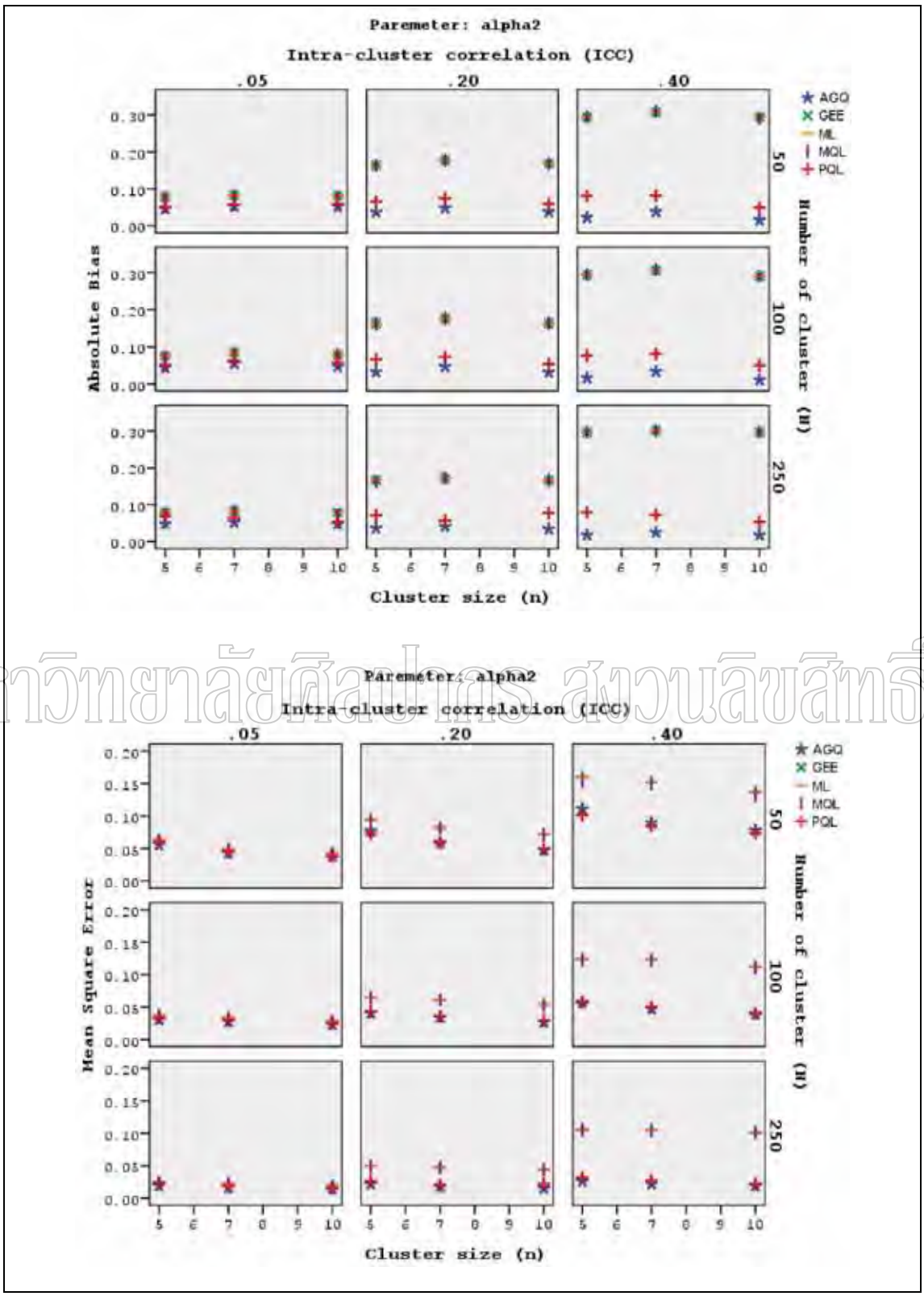
ภาพที่ 5 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ σ_u^2

3.2 การศึกษาผลกระทบของขนาดกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์



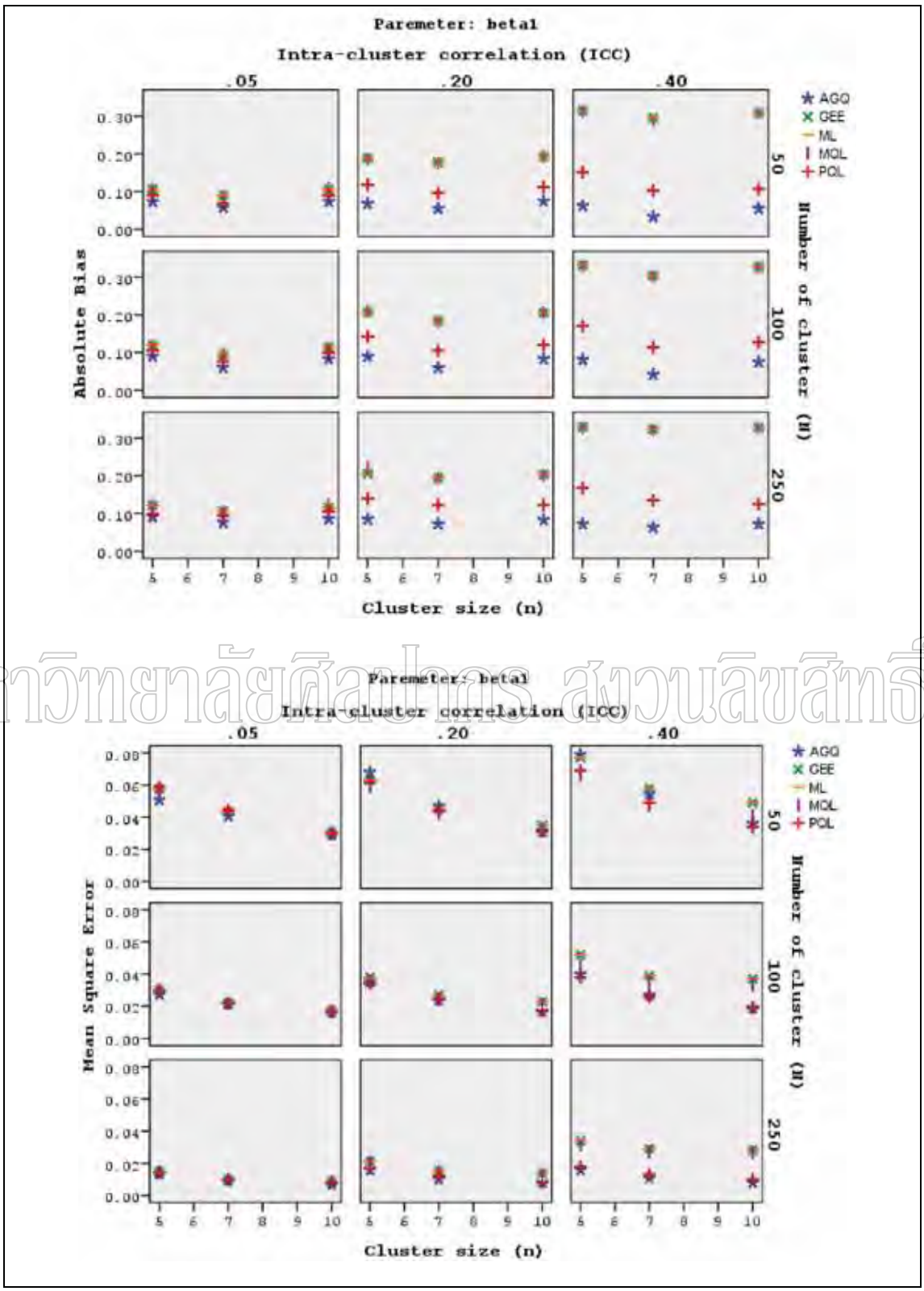
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 6 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_1



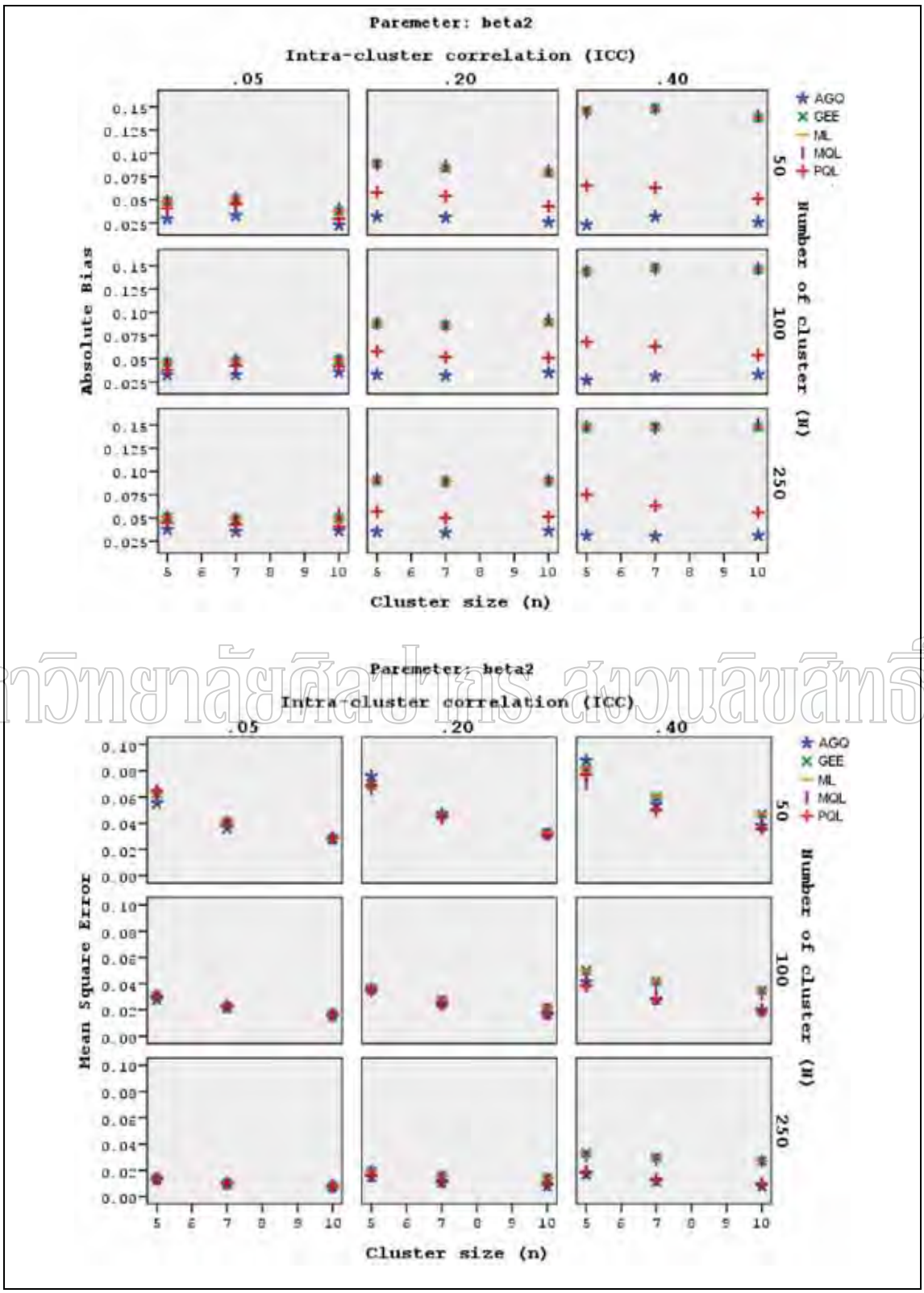
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 7 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_2



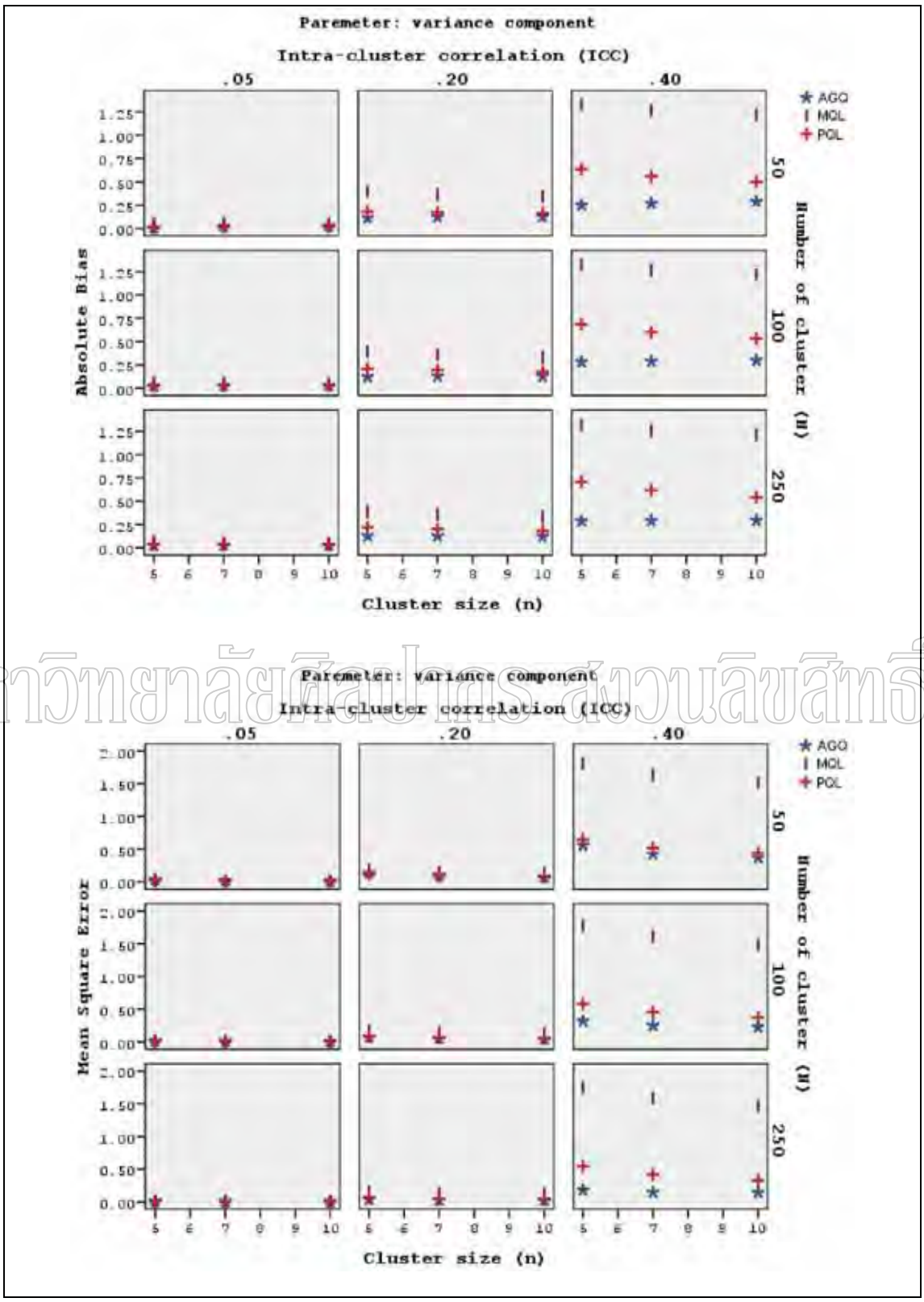
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 8 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_1



มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

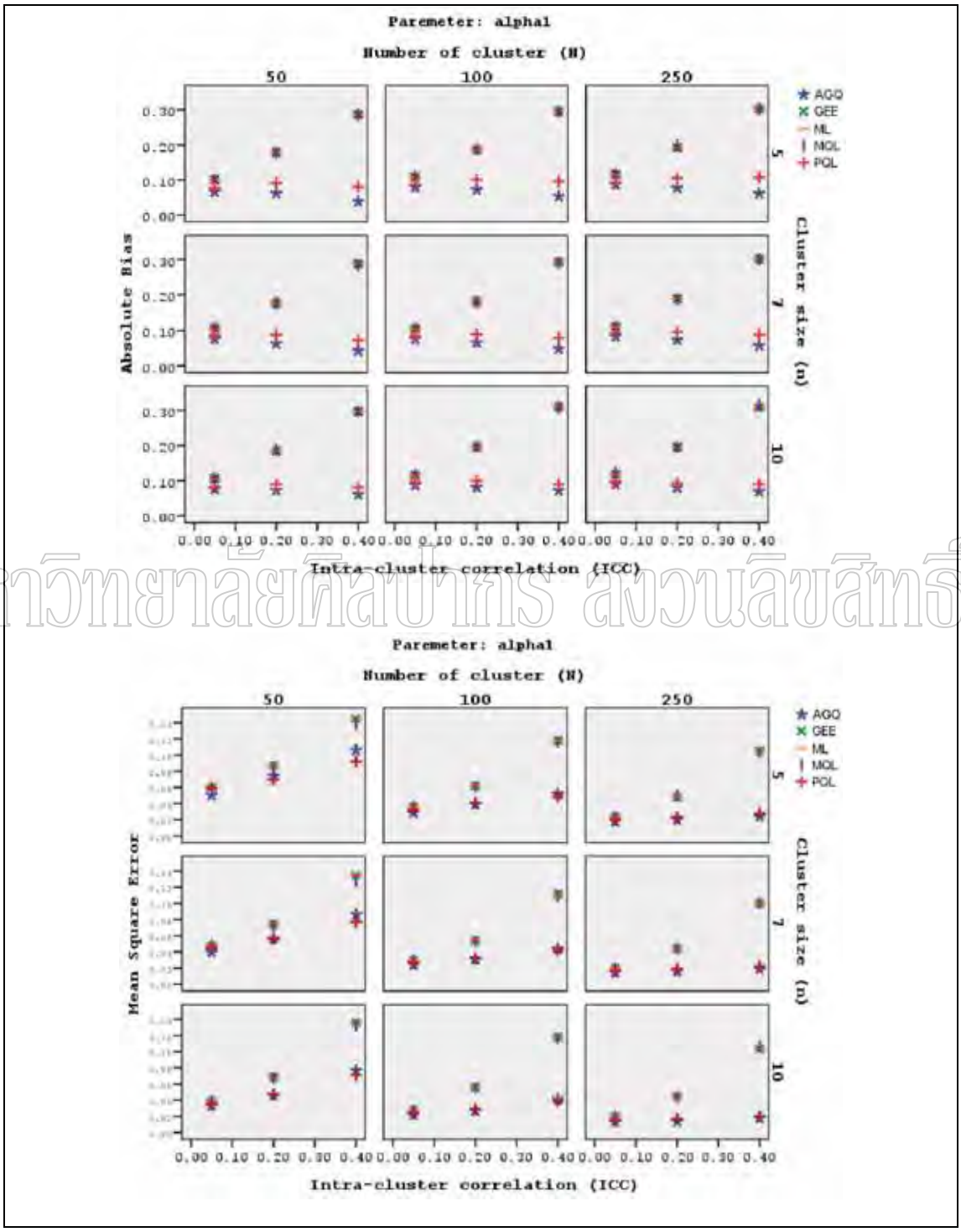
ภาพที่ 9 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_2



มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

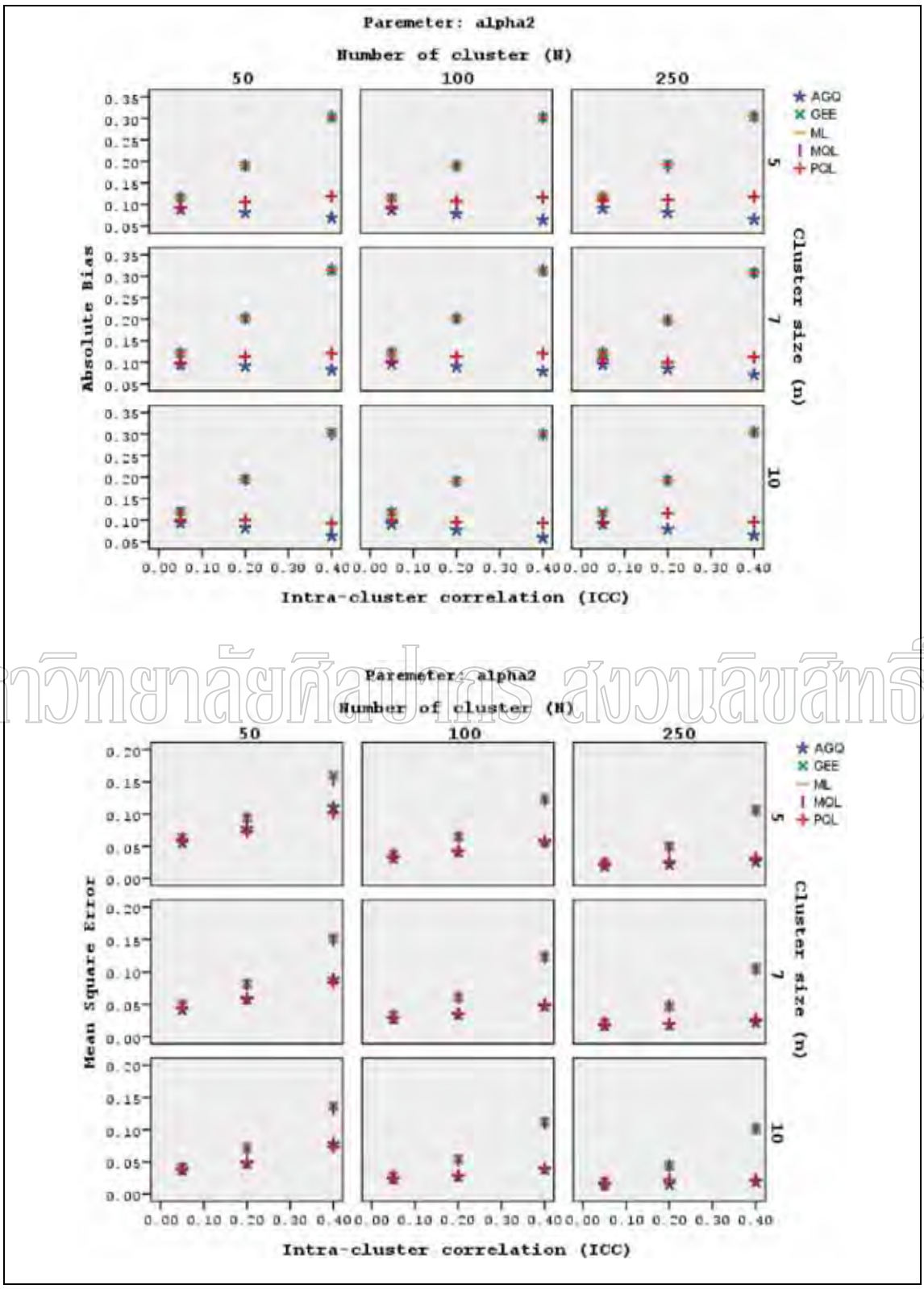
ภาพที่ 10 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของขนาดกลุ่มตัวอย่างที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ σ_u^2

3.3 การศึกษาผลกระทบของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์



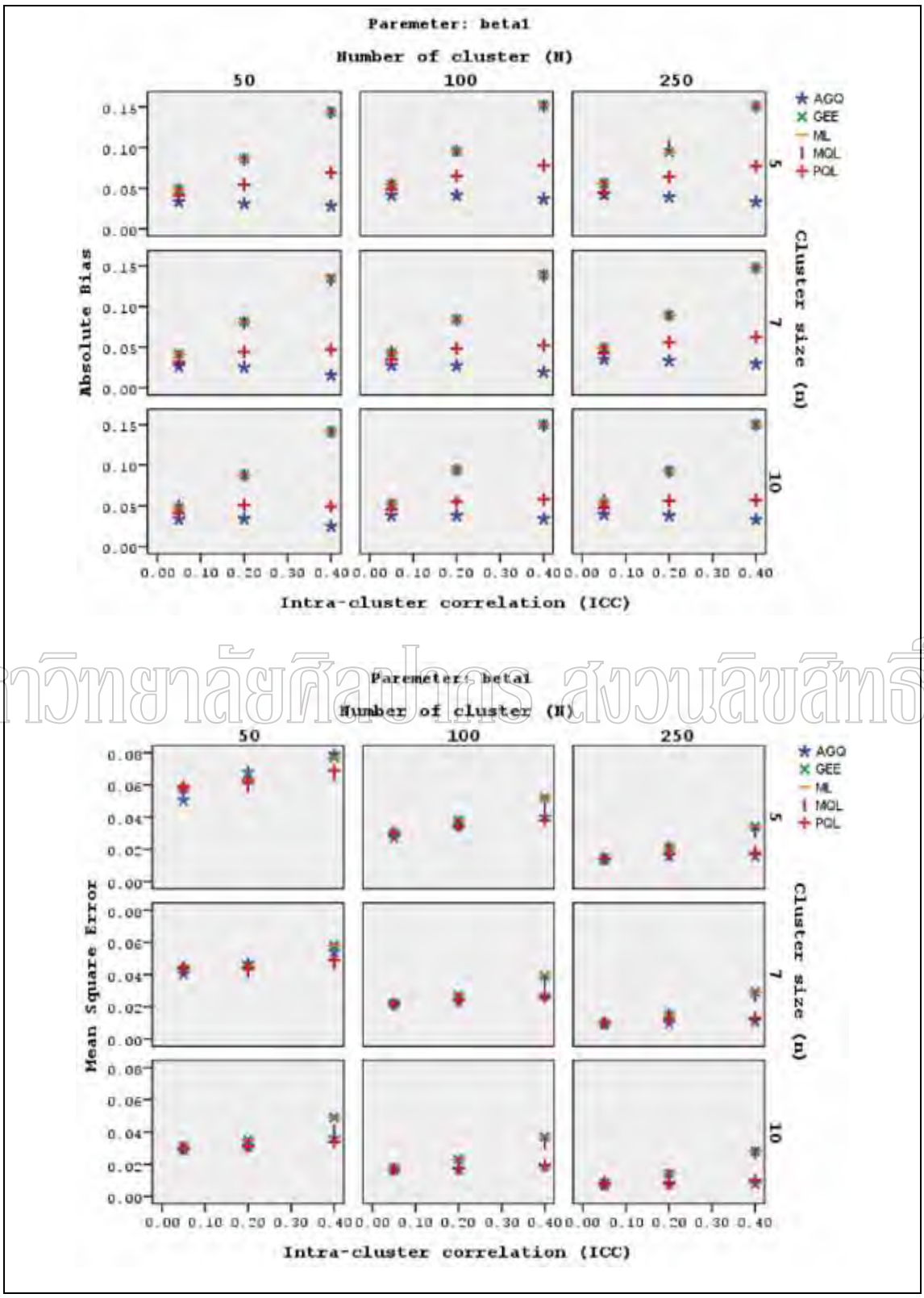
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 10 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_1



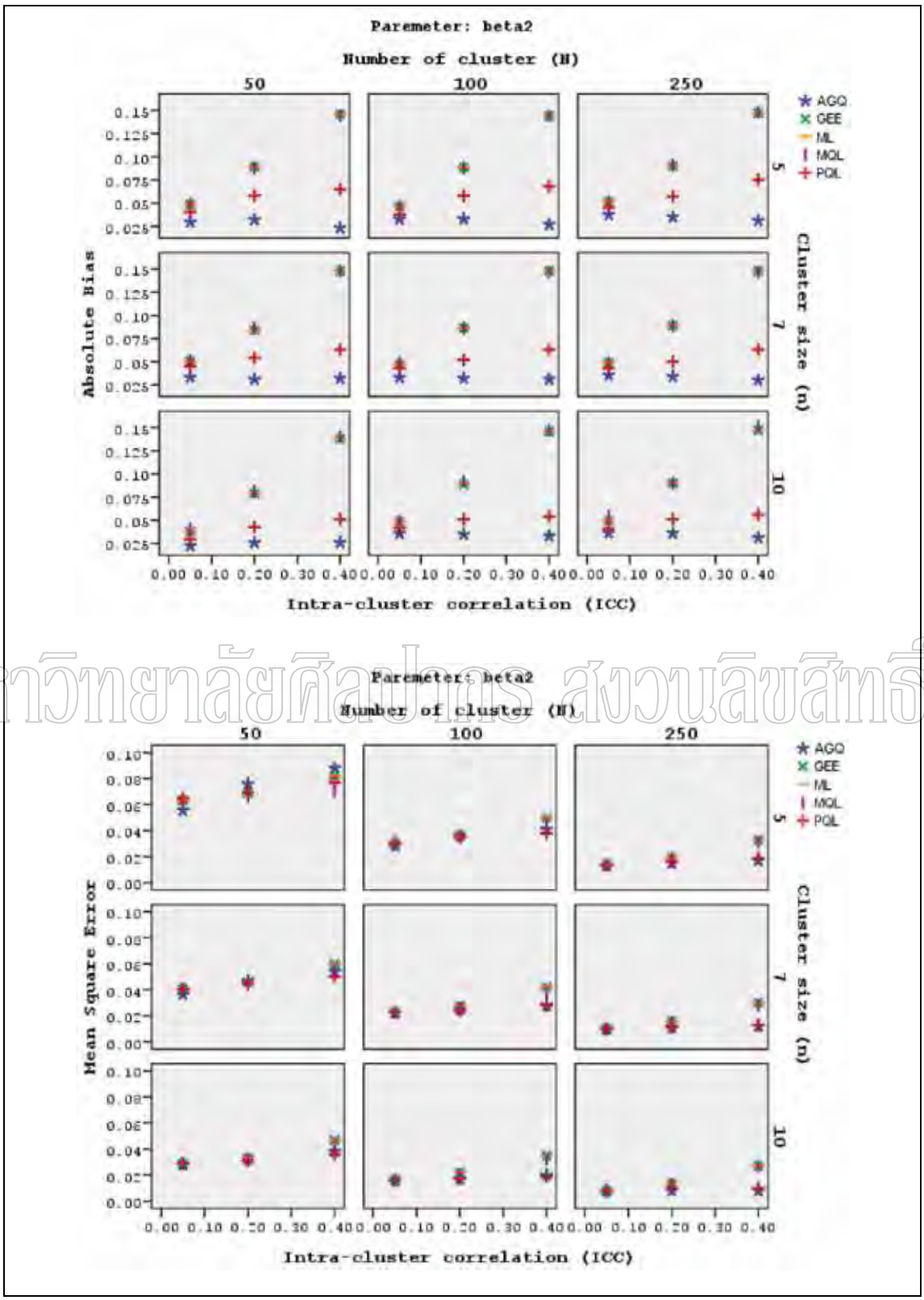
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 12 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ α_2



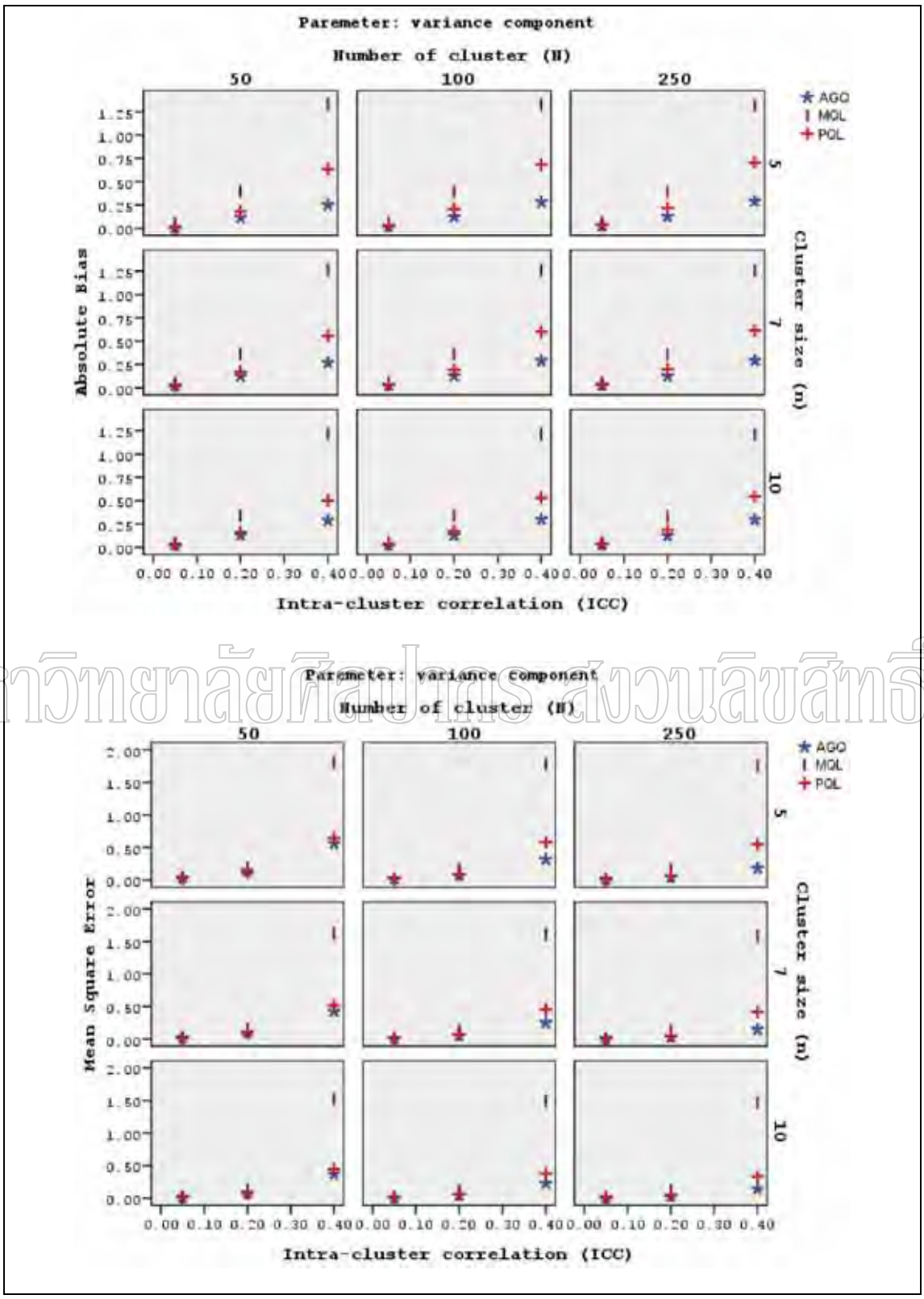
มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 13 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_1



มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 14 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ β_2



มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาพที่ 15 ค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยในการศึกษาผลกระทบของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องในการประมาณพารามิเตอร์ พารามิเตอร์ σ_u^2

จากภาพที่ 1 – 5 การศึกษาผลกระทบของจำนวนกลุ่มตัวอย่างที่มีต่อความถูกต้องของการประมาณพารามิเตอร์ ภายใต้สหสัมพันธ์ในกลุ่ม (0.2, 0.4, 0.6) และขนาดกลุ่ม (5, 7, 10) สำหรับการประมาณพารามิเตอร์จุดตัด (α_1, α_2) และสัมประสิทธิ์ของตัวแปรอธิบาย (β_1, β_2) พบว่าจำนวนกลุ่มที่เพิ่มขึ้นไม่ส่งผลกระทบต่อค่าความเอนเอียงของวิธีประมาณพารามิเตอร์ทั้ง 5 วิธี แต่ส่งผลให้ค่า MSE ลดลง โดยที่อัตราการลดลงของ MSE กรณีที่สหสัมพันธ์ในกลุ่มมีขนาดใหญ่ (ICC=0.4) มากกว่ากรณีที่สหสัมพันธ์ในกลุ่มมีขนาดกลาง (ICC=0.2) และขนาดเล็ก (ICC=0.05) สำหรับการประมาณพารามิเตอร์องค์ประกอบความแปรปรวน (σ_u^2) พบว่า จำนวนกลุ่มที่เพิ่มขึ้นไม่ส่งผลกระทบต่อค่าความเอนเอียงของวิธี MQL AGQ และ PQL และไม่ส่งผลกระทบต่อค่า MSE กรณีที่สหสัมพันธ์ในกลุ่มมีขนาดเล็ก แต่ส่งผลให้ค่า MSE ลดลงเพียงเล็กน้อย กรณีที่สหสัมพันธ์ในกลุ่มมีขนาดกลางและขนาดใหญ่

จากภาพที่ 6 – 10 การศึกษาผลกระทบของขนาดกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์ ภายใต้สหสัมพันธ์ในกลุ่ม (0.2, 0.4, 0.6) และจำนวนกลุ่ม (50, 100, 250) สำหรับการประมาณพารามิเตอร์จุดตัด (α_1, α_2) และสัมประสิทธิ์ของตัวแปรอธิบาย (β_1, β_2) พบว่าขนาดกลุ่มที่เพิ่มขึ้นไม่ส่งผลกระทบต่อค่าความเอนเอียงของวิธีประมาณพารามิเตอร์ทั้ง 5 วิธี แต่ส่งผลให้ค่า MSE ลดลงเพียงเล็กน้อย โดยที่อัตราการลดลงของ MSE กรณีที่สหสัมพันธ์ในกลุ่มมีขนาดใหญ่ (ICC=0.4) มากกว่ากรณีที่สหสัมพันธ์ในกลุ่มมีขนาดกลาง (ICC=0.2) และขนาดเล็ก (ICC=0.05) สำหรับการประมาณพารามิเตอร์องค์ประกอบความแปรปรวน (σ_u^2) พบว่าขนาดกลุ่มที่เพิ่มขึ้นไม่ส่งผลกระทบต่อค่าความเอนเอียงและค่า MSE ของวิธี MQL, AGQ และ PQL กรณีที่สหสัมพันธ์ในกลุ่มมีขนาดเล็กและไม่ส่งผลกระทบต่อค่าความเอนเอียงและค่า MSE ของวิธี AGQ ทุกกรณี แต่ส่งผลให้ค่าความเอนเอียงและค่า MSE ของวิธี MQL และ PQL ลดลงเล็กน้อย กรณีที่สหสัมพันธ์ในกลุ่มมีขนาดกลางและขนาดใหญ่

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

จากภาพที่ 11 – 15 การศึกษาผลกระทบของสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์ในตัวแบบ ภายใต้จำนวนกลุ่ม (50, 100, 250) และขนาดกลุ่ม (5, 7, 10) สำหรับการประมาณพารามิเตอร์จุดตัด (α_1, α_2) และสัมประสิทธิ์ของตัวแปรอธิบาย (β_1, β_2) พบว่าสหสัมพันธ์ในกลุ่มที่เพิ่มขึ้นส่งผลให้ค่าความเอนเอียงและค่า MSE เพิ่มขึ้นอย่างมากสำหรับวิธี ML GEE และ MQL สำหรับทุกเงื่อนไข แต่ส่งผลให้ค่าความเอนเอียงและค่า MSE เพิ่มขึ้นเพียงเล็กน้อย สำหรับวิธี AGQ และ PQL สำหรับการประมาณองค์ประกอบความแปรปรวน (σ_u^2) พบว่าสหสัมพันธ์ในกลุ่มที่เพิ่มขึ้นส่งผลให้ค่าความเอนเอียงและค่า MSE ของวิธี MQL AGQ และ PQL เพิ่มขึ้น โดยที่อัตราการเพิ่มขึ้นสำหรับค่าความเอนเอียงและค่า MSE ของวิธี MQL มากที่สุด รองลงมาคือวิธี PQL และ AGQ ตามลำดับ

บทที่ 5

สรุป อภิปรายผลและข้อเสนอแนะ

การเปรียบเทียบวิธีประมาณพารามิเตอร์ของตัวแบบผสมเชิงเส้นนัยทั่วไป สำหรับตัวแปรตอบสนองแบบอันดับ กรณีของการวิเคราะห์ด้วยตัวแบบมาร์จินัล ประกอบด้วยวิธี Generalized Estimating Equations (GEE) และ Marginal Quasi-Likelihood (MQL) เปรียบเทียบกับวิธี Maximum Likelihood (ML) สำหรับตัวแบบอิทธิพลคงที่ และกรณีของตัวแบบอิทธิพลสุ่ม ประกอบด้วยวิธี Adaptive Gaussian Quadrature (AGQ) และ Penalized Quasi-Likelihood (PQL) โดยพิจารณาจากความเอนเอียง ความคลาดเคลื่อนกำลังสองเฉลี่ยและสัดส่วนของช่วงความเชื่อมั่น 95% ของค่าประมาณพารามิเตอร์ที่ครอบคลุมค่าจริง จากการจำลองแบบข้อมูล จำนวน 1,000 ครั้ง ภายใต้เงื่อนไขของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่างและสหสัมพันธ์ในกลุ่ม นอกจากนี้ยังศึกษาถึงผลกระทบของปัจจัยดังกล่าวที่มีต่อความถูกต้องของการประมาณพารามิเตอร์อีกด้วย จากผลการวิจัย ผู้วิจัยสามารถสรุปผล อภิปรายผล พร้อมทั้งนำเสนอข้อเสนอแนะของงานวิจัย ดังนี้

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

1. สรุปผลการวิจัย

จากผลการวิจัย สามารถสรุปผลแบ่งเป็น 2 ตอน ดังนี้

1.1 การเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบผสมเชิงเส้นนัยทั่วไป

สำหรับตัวแบบมาร์จินัล วิธี GEE และ MQL ให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริง ซึ่งสอดคล้องกับงานวิจัยของ Kuss and McLerran (2007) โดยมีความถูกต้องในการประมาณพารามิเตอร์ที่ใกล้เคียงกัน เมื่อสหสัมพันธ์ในกลุ่มมีขนาดเล็ก วิธี MQL มีแนวโน้มของความถูกต้องสูงกว่าวิธี GEE เล็กน้อย แต่เมื่อสหสัมพันธ์ในกลุ่มมีขนาดกลางและขนาดใหญ่ พบว่าความถูกต้องในการประมาณพารามิเตอร์ของวิธี MQL สูงกว่าวิธี GEE อย่างชัดเจน กรณีที่จำนวนกลุ่มตัวอย่างมีขนาดเล็กและขนาดกลาง แต่มีความถูกต้องใกล้เคียงกันกรณีที่จำนวนกลุ่มมีขนาดใหญ่ และเมื่อเปรียบเทียบวิธีสำหรับตัวแบบมาร์จินัลทั้งสองนี้กับวิธี ML สำหรับตัวแบบอิทธิพลคงที่พบว่าเมื่อสหสัมพันธ์ในกลุ่มมีขนาดเล็ก วิธี ML มีความถูกต้องใกล้เคียงกับวิธี GEE แต่น้อยกว่าวิธี MQL เล็กน้อย แต่เมื่อสหสัมพันธ์ในกลุ่มมีขนาดกลางและขนาดใหญ่ พบว่าวิธี ML มีความถูกต้องน้อยกว่าวิธี GEE และ MQL มากขึ้น

สำหรับตัวแบบอิทธิพลสุ่ม วิธี AGQ ให้ค่าประมาณพารามิเตอร์ที่มีแนวโน้มต่ำกว่าค่าจริงเล็กน้อย ส่วนวิธี PQL ให้ค่าประมาณพารามิเตอร์จุดตัดและสัมประสิทธิ์ของตัวแบบ ที่มีแนวโน้มต่ำกว่าค่าจริงเล็กน้อย ให้ค่าประมาณที่มีแนวโน้มต่ำกว่าค่าจริงค่อนข้างมากสำหรับพารามิเตอร์องค์ประกอบความแปรปรวน เมื่อสหสัมพันธ์ในกลุ่มมีขนาดใหญ่ เมื่อพิจารณาความถูกต้องในการประมาณพารามิเตอร์ พบว่าวิธี AGQ มีความถูกต้องในการประมาณพารามิเตอร์มากกว่าวิธี PQL และมีความแตกต่างชัดเจนขึ้นเมื่อจำนวนกลุ่มตัวอย่างมีขนาดใหญ่

เมื่อพิจารณาวิธีการประมาณพารามิเตอร์ที่นำมาศึกษาทั้ง 5 วิธี สำหรับตัวแบบอิทธิพลคงที่ ตัวแบบมาร์จินัลและตัวแบบอิทธิพลสุ่ม พบว่าวิธี AGQ เป็นวิธีที่ให้ค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับค่าจริงมากที่สุด และมีความถูกต้องมากที่สุด สำหรับทุกกรณี

1.2 การศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มและสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์

ในการศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง พบว่าเมื่อจำนวนกลุ่มตัวอย่างเพิ่มขึ้นส่งผลให้ความถูกต้องในการประมาณพารามิเตอร์ สำหรับตัวแบบอิทธิพลคงที่ (วิธี ML) ตัวแบบมาร์จินัล (วิธีGEE และวิธี MQL) และตัวแบบอิทธิพลสุ่ม (วิธีAGQ และวิธี PQL) สูงขึ้น ซึ่งสอดคล้องกับงานวิจัยของ Moineddin, Matheson and Glazier (2007) โดยเฉพาะวิธี AGQ และ PQL

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ในการศึกษาผลกระทบของขนาดกลุ่มตัวอย่าง พบว่าเมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นส่งผลให้ความถูกต้องในการประมาณพารามิเตอร์แต่ละวิธีสูงขึ้นเพียงเล็กน้อยเท่านั้นเมื่อเทียบกับผลกระทบจากจำนวนกลุ่มตัวอย่าง

ในการศึกษาผลกระทบของสหสัมพันธ์ในกลุ่ม พบว่าเมื่อสหสัมพันธ์ในกลุ่มเพิ่มขึ้นส่งผลให้ความถูกต้องในการประมาณพารามิเตอร์แต่ละวิธีลดลง ซึ่งสอดคล้องกับงานวิจัยของ Moineddin, Matheson and Glazier (2007) โดยลดลงอย่างมากสำหรับวิธี ML, GEE และ MQL ขณะที่วิธี AGQ และ PQL ลดลงเพียงเล็กน้อย

2. อภิปรายผลการวิจัย

จากผลการวิจัย สามารถอภิปรายผลแบ่งเป็น 2 ตอน ดังนี้

2.1 การเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบผสมเชิงเส้นนัยทั่วไป

จากผลการวิจัยที่พบว่า วิธี ML สำหรับตัวแบบอิทธิพลคงที่ มีความถูกต้องในการประมาณพารามิเตอร์ที่ใกล้เคียงกับวิธี GEE และ MQL สำหรับตัวแบบมาร์จินัล เมื่อสหสัมพันธ์ในกลุ่มมีขนาดเล็ก แต่มีความถูกต้องในการประมาณพารามิเตอร์น้อยกว่าวิธี GEE และ MQL เมื่อสหสัมพันธ์ในกลุ่มมีขนาดกลางและขนาดใหญ่ ทั้งนี้อาจเป็นเพราะว่าวิธี ML ใช้ในการประมาณ

พารามิเตอร์ของตัวแบบภายใต้ข้อสมมติว่าค่าสังเกตเป็นอิสระต่อกัน ส่วนวิธี GEE และ MQL พิจารณาความสัมพันธ์ของข้อมูลในกลุ่มไว้ในเมทริกซ์สหสัมพันธ์หรือเมทริกซ์ความแปรปรวนร่วมของข้อมูลด้วย ซึ่งในกรณีที่สหสัมพันธ์ในกลุ่มมีขนาดเล็ก จึงสังเกตเห็นได้ว่าวิธี ML มีความถูกต้องใกล้เคียงกับวิธี GEE และ MQL แต่เมื่อข้อมูลในกลุ่มมีความสัมพันธ์กัน จึงทำให้การประมาณพารามิเตอร์ของวิธี GEE และ MQL มีความถูกต้องมากกว่าวิธี ML

สำหรับตัวแบบอิทธิพลสุ่ม พบว่าวิธี AGQ และ PQL ให้ค่าประมาณพารามิเตอร์สัมประสิทธิ์ของตัวแบบที่มีแนวโน้มต่ำกว่าค่าจริงเล็กน้อย แต่ให้ค่าประมาณพารามิเตอร์องค์ประกอบความแปรปรวนที่ต่ำกว่าค่าจริงอย่างมาก สำหรับวิธี PQL เมื่อสหสัมพันธ์ในกลุ่มมีขนาดใหญ่ ซึ่งสอดคล้องกับงานวิจัยของ Breslow and Clayton (1993) เมื่อพิจารณาความถูกต้องในการประมาณพารามิเตอร์ พบว่าวิธี AGQ มีความถูกต้องในการประมาณพารามิเตอร์มากกว่าวิธี PQL และมีความแตกต่างชัดเจนขึ้นเมื่อจำนวนกลุ่มตัวอย่างมีขนาดใหญ่

เมื่อพิจารณา วิธี GEE และ MQL สำหรับตัวแบบมาร์จินัลและวิธี AGQ และ PQL สำหรับตัวแบบอิทธิพลสุ่ม ซึ่งตัวแบบทั้งสองเป็นตัวแบบที่นิยมใช้ในการวิเคราะห์มูลเมื่อข้อมูลมีสหสัมพันธ์กัน พบว่าวิธีประมาณพารามิเตอร์สำหรับตัวแบบอิทธิพลสุ่มมีความถูกต้องมากกว่าวิธีสำหรับตัวแบบมาร์จินัล ทั้งนี้อาจเป็นเพราะว่าตัวแบบอิทธิพลสุ่มพิจารณาความสัมพันธ์ของข้อมูลในกลุ่มหรือความแตกต่างกันของข้อมูลระหว่างกลุ่ม แทนด้วยอิทธิพลสุ่มของกลุ่มเข้าไปในตัวแบบในเทอมของตัวพยากรณ์เชิงเส้น จึงทำให้มีความถูกต้องสูงกว่าวิธีสำหรับตัวแบบมาร์จินัล ซึ่งไม่ได้แทนความสัมพันธ์ของข้อมูลในกลุ่มเข้าไปในตัวแบบ เพียงแต่พิจารณาไว้ในเมทริกซ์สหสัมพันธ์ของข้อมูลเท่านั้น

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2.2 การศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มและสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องของการประมาณพารามิเตอร์

จากผลการวิจัยที่พบว่าขนาดตัวอย่าง (จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่าง) ที่เพิ่มขึ้นส่งผลให้ความถูกต้องในการประมาณพารามิเตอร์แต่ละวิธีสูงขึ้น โดยเฉพาะวิธี AGQ และ PQL จะมีความถูกต้องสูงมากขึ้น เมื่อสหสัมพันธ์ในกลุ่ม มีขนาดใหญ่ขึ้น ทั้งนี้อาจเป็นเพราะว่าเมื่อขนาดตัวอย่าง (จำนวนกลุ่มตัวอย่างและขนาดกลุ่มตัวอย่าง) เพิ่มขึ้น จะทำให้การประมาณพารามิเตอร์ดีขึ้น แต่สำหรับกรณีที่ข้อมูลในกลุ่มมีความสัมพันธ์กันมากขึ้น วิธี AGQ และ PQL ซึ่งแทนความสัมพันธ์ของข้อมูลนี้เข้าไปในตัวแบบด้วยจึงมีความถูกต้องสูงขึ้นมากกว่าวิธี ML ซึ่งไม่คำนึงถึงความสัมพันธ์ของข้อมูลภายในกลุ่ม และวิธี GEE และ MQL ซึ่งไม่ได้แทนอิทธิพลสุ่มเข้าไปในตัวแบบ เพียงแต่พิจารณาไว้ในเมทริกซ์สหสัมพันธ์ของข้อมูลเท่านั้น และจากผลการวิจัยที่พบว่าสหสัมพันธ์ในกลุ่มที่เพิ่มขึ้นส่งผลให้ความถูกต้องในการประมาณพารามิเตอร์แต่ละวิธีลดลง

โดยที่ความถูกต้องในการประมาณพารามิเตอร์สำหรับวิธี ML GEE และ MQL ลดลงอย่างมากเมื่อสหสัมพันธ์ในกลุ่มเพิ่มขึ้น ขณะที่วิธี AGQ และ PQL ความถูกต้องในการประมาณพารามิเตอร์ลดลงเพียงเล็กน้อยและลดลงน้อยมากเมื่อจำนวนกลุ่มตัวอย่างมีขนาดใหญ่ ทั้งนี้อาจเป็นเพราะว่าเมื่อสหสัมพันธ์ของข้อมูลในกลุ่มมากขึ้น แต่ทำการประมาณพารามิเตอร์ด้วยวิธี ML ซึ่งไม่คำนึงถึงความสัมพันธ์ของข้อมูลภายในกลุ่ม รวมถึงวิธี GEE และ MQL ซึ่งไม่ได้แทนอิทธิพลสุ่มเข้าไปในตัวแบบ เพียงแต่พิจารณาไว้ในเมทริกซ์สหสัมพันธ์ของข้อมูลเท่านั้น ส่งผลให้การประมาณพารามิเตอร์ลดลงอย่างมาก แต่สำหรับการประมาณพารามิเตอร์ด้วยวิธี AGQ และ PQLซึ่งแทนความสัมพันธ์ของข้อมูลนี้เข้าไปในตัวแบบด้วยจึงมีความถูกต้องลดลงเพียวเล็กน้อยเท่านั้น

3. ข้อเสนอแนะของงานวิจัย

ในการเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบผสมเชิงเส้นนัยทั่วไป กรณีของการวิเคราะห์ด้วยตัวแบบมาร์จินัลและตัวแบบอิทธิพลสุ่ม ภายใต้เงื่อนไขของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่างและสหสัมพันธ์ในกลุ่ม การกำหนดความแปรปรวนของอิทธิพลสุ่มในการจำลองข้อมูล ควรจะกำหนดให้สัมพันธ์กับสหสัมพันธ์ในกลุ่ม และไม่ควรกำหนดให้ค่าสหสัมพันธ์ในกลุ่มมีค่ามากเกินไป จะทำให้ค่าประมาณพารามิเตอร์ต่ำกว่าค่าจริงอย่างมาก ส่วนการเลือกใช้วิธีประมาณพารามิเตอร์สำหรับตัวแปรตอบสนองแบบอันดับ เมื่อค่าตอบสนองมีความสัมพันธ์กัน เริ่มต้นควรพิจารณาว่าจะวิเคราะห์ข้อมูลภายใต้ตัวแบบตัวแบบมาร์จินัลหรือตัวแบบอิทธิพลสุ่มจากจุดประสงค์ในการตีความไปที่ระดับประชากรหรือระดับกลุ่ม หลังจากนั้นเลือกวิธีประมาณพารามิเตอร์ให้เหมาะสมกับตัวแบบที่เลือกใช้ เช่นตัวแบบมาร์จินัล สามารถใช้วิธีใดก็ได้ทั้ง GEE และ MQL เนื่องจากผลงานวิจัยแสดงให้เห็นว่ามีประสิทธิภาพใกล้เคียงกัน และสำหรับตัวแบบอิทธิพลสุ่มควรเลือกใช้วิธี AGQ เนื่องจากมีประสิทธิภาพมากกว่าวิธี PQL

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

4. ข้อเสนอแนะในการวิจัยครั้งต่อไป

งานวิจัยนี้ทำการเปรียบเทียบวิธีประมาณพารามิเตอร์สำหรับตัวแบบผสมเชิงเส้นนัยทั่วไป กรณีของการวิเคราะห์ด้วยตัวแบบมาร์จินัลและตัวแบบอิทธิพลสุ่ม ซึ่งเป็นกลุ่มของตัวแบบที่นิยมใช้ในการวิเคราะห์มูลเมื่อข้อมูลมีสหสัมพันธ์กัน โดยวิเคราะห์ภายใต้ตัวแบบลอจิตสะสม ซึ่งเป็นตัวแบบที่นิยมใช้สำหรับตัวแปรตอบสนองแบบอันดับ และศึกษาผลกระทบของจำนวนกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่างและสหสัมพันธ์ในกลุ่มที่มีต่อความถูกต้องของค่าประมาณพารามิเตอร์จากวิธีประมาณแต่ละวิธีสำหรับแต่ละตัวแบบ ในงานวิจัยครั้งต่อไปอาจมีการศึกษาเปรียบเทียบวิธีประมาณพารามิเตอร์โดยใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบอันดับอีกตัวแบบหนึ่งที่มีความ

นิยามใกล้เคียงกัน เช่นตัวแบบโพรบิตสะสม และยังสามารถนำมาเปรียบเทียบกับตัวแบบของงานวิจัยนี้ได้อีกด้วย นอกจากนี้อาจทำการศึกษาถึงผลกระทบจากปัจจัยอื่น ๆ ที่มีต่อความถูกต้องของค่าประมาณพารามิเตอร์ เช่นสัดส่วนผลตอบสนอง จำนวนตัวแปรอธิบาย การแจกแจงของตัวแปรอธิบายหรือโครงสร้างสหสัมพันธ์ของข้อมูล และอาจเปรียบเทียบความเหมาะสมของตัวแบบสำหรับตัวแปรตอบสนองแบบอันดับที่มีความสัมพันธ์กันได้อีกด้วย

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

บรรณานุกรม

Abramowitz, M., and I. A. Stegun. Handbook of mathematical functions with formulars, graphs, and mathematical tables. NewYork : Dover Publications, 1972.

Agresti, A. Categorical data analysis. 2nd ed. NewYork : John Wiley & Sons, 2002.

Breslow, N. E, and D. G. Clayton. "Approximate inference in generalized linear mixed model." Journal of the American Statistical Association 88 (1993) : 9-25.

Chen, Z., and L. Kuo. "A note on the estimation of the multinomial logit model with random effects." The American Statistician 55, 2 (May 2001) : 89-95.

Hartzel, J., A. Agresti, and B. Caffo. "Multinomial logit random effects models." Statistical Modelling 1 (2001) : 81-102.

Hedeker, D. "A mixed effects multinomial logistic regression model." Statistical in Medicine 22 (2003) : 1433-1446.

Hedeker, D., and R. D. Gibbons. "A random effects ordinal regression model for multilevel analysis." Biometrices 50 (December 1994) : 933-944.

Kenward, M. G., E. Lesaffre, and G. Molenberghs. "An application of maximum likelihood and generalized estimating equations to the analysis of ordinal data from a longitudinal study with case missing at random." Biometrices 50, (1994) : 945 – 954.

Kuss, O., and D. McLerran. "A note on the estimation of the multinomial logistic model with correlated response in SAS." Computer Method and Programs in Medicine 87 (2007) : 262-269.

Liang, K. Y., and S. L. Zeger. "Longitudinal data analysis using generalized linear models." Biometrika 73 (1986) : 13-22.

Lipsitz, S. R., K. Kim, and L. Zhao. "Analysis of repeated categorical data using generalized estimating equation." Statistic in Medicine 13 (1994) : 1149-1163

Liu, L. C., and D. Hedeker. "A mixed effects regression models for longitudinal multivariate ordinal data." Biometrics 62 (March 2006) : 261-268.

Liu, Q., and D. A. Pierce. "A note on Guass-Hermite quadrature." Biometrika 81 (1994) : 624-9.

McCullagh, P. "Regression model for ordinal data (with discussion)." Journal of the Royal Statistical Society 42 (1980) : 109-127.

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

McCullagh, P., and J. A. Nelder. Generalized linear models. London : Chapman &Hall, 1989.

Miller, M. E., C. S. Davis, and J. R. Landis. "The analysis of longitudinal polytomous data: generalized estimation equations and connections with weighted least squares." Biometrics 49 (1993) : 1033-1044.

Moineddin, R., F. I. Matheson, and R. H. Glazier. "A simulation study of sample size for multilevel logistic regression models." BMC Medical Research Methodology (2007) : 7-34.

Neuhaus, J. M., J. D. Kalbfleisch, and W. W. Hauck. "A comparison of cluster-specific and population- averaged approaches for analyzing correlated binary data." International Statistical Review 59 (1991) : 25-35.

Parsons, N. et al. "Repeated measures proportional odds logistic regression analysis of ordinal score data in the statistical software package R." Computational Statistics and Data Analysis 53 (2009) : 632-641.

Pinheiro, J. C., and D. M. Bate. "Approximations to the log-likelihood function in the nonlinear mixed-effects models." Journal of Computational and Graphical Statistics 4 (1995) : 12-35.

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

Prentice, R. L. "Correlated binary regression with covariates specific to each binary observation." Biometrics 44 (1988) : 1033-1084.

Schepers, A., G. Arminger, and U. Kusters. "The analysis of non-metric endogenous variables in latent variable models : the MECOSA appproach." Economics and Mathematical Systems 366 (1991) : 459-472.

Spiess, M., and A. Hamerle. "A comparison of different methods for the estimation of regression models with correlated binary responses." Computational Statistics and Data Analysis 33 (2000) : 439-455.

Stiratelli, R., N. Laird, and J. H. Ware. "Random-effects models for serial observation with binary response." Biometrics 40 (1984) : 961-971.

Tutz, G., and G. Kauermann. "Generalized linear random effects models with varying coefficients." Computational Statistics and Data Analysis 43 (2003) : 13-28.

Tutz, G., and W. Hennevogl. "Random effects in ordinal regression models." Computational Statistics and Data Analysis 22 (1996) : 537-557.

Wolfinger, R., and M. O'Connell. “Generalized linear mixed models: a pseudo-likelihood approach.” Journal of Statistical Computing and Simulation 48 (1993) : 233-243.

Zeger, S. L., and R. Karim. “Generalized linear models with random effects : a Gibbs sampling approach.” Journal of the American Statistical Association 86 (1991) : 79-86.

Zeger, S. L., K. Y. Liang, and P. S. Albert. “Models for longitudinal data : a generalized estimating equation approach.” Biometrics 44 (1988) : 1049-1060.

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาคผนวก

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

โปรแกรมคอมพิวเตอร์ที่ใช้ในงานวิจัย

งานวิจัยนี้ทำการจำลองแบบข้อมูลและประมวลผลข้อมูลโดยใช้โปรแกรมสำเร็จรูป SAS® version 9.1.3 ในการจำลองแบบข้อมูลซึ่งประกอบด้วยตัวแปรอธิบาย ตัวแปรตอบสนอง และคำนวณค่าประมาณพารามิเตอร์ด้วยวิธีต่าง ๆ ที่นำเสนอในงานวิจัย โดยแสดงโปรแกรมสำหรับการจำลองแบบข้อมูลและการประมวลผลข้อมูล ตามลำดับ ดังนี้

1. โปรแกรมสำหรับการจำลองแบบข้อมูล

การจำลองแบบข้อมูลตัวแปรตอบสนองแบบอันดับ ตัวแปรอธิบาย (สำหรับอิทธิพลคงที่) ที่มีการแจกแจงแบบเบอร์นูลลี ที่มีความน่าจะเป็น เท่ากับ 0.5 ตัวแปรอธิบาย (สำหรับอิทธิพลสุ่ม) จำนวน 1 ตัวแปร ที่มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวน เท่ากับ σ_u^2 ภายใต้จำนวนกลุ่มตัวอย่าง เท่ากับ 50 100 และ 250 ขนาดกลุ่มตัวอย่าง เท่ากับ 5, 7 และ 10 และสหสัมพันธ์ในกลุ่ม เท่ากับ 0.05, 0.20 และ 0.40 จำนวน 1,000 ครั้ง โดยการเขียนโปรแกรมแมคโคร ซึ่งความหมายของสัญลักษณ์ต่าง ๆ ในตัวโปรแกรม และการเขียนโปรแกรม มีรายละเอียดดังนี้

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

nrep	หมายถึง	จำนวนครั้งของการจำลองแบบข้อมูล
N	หมายถึง	จำนวนกลุ่มตัวอย่าง
ni	หมายถึง	ขนาดกลุ่มตัวอย่าง
s2u	หมายถึง	ความแปรปรวนของอิทธิพลสุ่ม (σ_u^2)
a1	หมายถึง	สัมประสิทธิ์จุดตัดที่ 1 (α_1)
a2	หมายถึง	สัมประสิทธิ์จุดตัดที่ 2 (α_2)
b1	หมายถึง	สัมประสิทธิ์ของตัวแปรอธิบาย x_1 (β_1)
b2	หมายถึง	สัมประสิทธิ์ของตัวแปรอธิบาย x_2 (β_2)

1.1 การเขียนโปรแกรมสำหรับการจำลองแบบข้อมูล

```
%macro case(nrep=,N=,ni=,s2u=,a1=,a2=,b1=,b2=);
option ps=60 ls=80 nodate nonumber;
data sim;
nrep=&nrep;
N=&N;
ni=&ni;
a1=&a1;
a2=&a2;
b1=&b1;
b2=&b2;
s2u=&s2u;
s_u=sqrt(&s2u);
seed=12345;
pi=3.141592654;
s_e= sqrt(pi**2/3);
do rep= 1 to nrep;
        do subject=1 to N;
        u=0+sqrt(s2u)*rannor(seed);
            do repeat=1 to ni;
                e=0+s_e*rannor(seed);
                x1=ranbin(seed,1,0.5);
                x2=ranbin(seed,1,0.5);
                yc=-b1*x1-b2*x2-u+e;
                if (yc<=a1) then y=1;
                else if (yc>a1 & yc<=a2)then y=2;
                else y=3;
                output;
            end;
        end;
end;
keep rep subject x1 x2 u y;
run;
```

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2. โปรแกรมสำหรับการประมวลผลข้อมูล

2.1 การสร้างไฟล์สำหรับเก็บผลลัพธ์

```
filename logfile 'd:\out.log';
filename lisfile 'd:\out.lis';
proc printto log=logfile print= lisfile new;
run;
```

2.2 การเขียนโปรแกรมในการประมวลผลข้อมูลสำหรับวิธี ML

```
proc logistic data=sim; by rep;
        model y=x1 x2/link=clogit;
ods output parameterestimates=totfix fitstatistics=fitfix;
run;
proc sort data=fitfix;by criterion;
run;
proc univariate data=fitfix noprint ;
        by criterion;
        var interceptandcovariates;
output out=m_fitfix mean=fitfix;
run;
data totfix;set totfix;
        parmfix=variable;
        estfix=estimate;
        sefix=stderr;
        lclfix=estfix-1.96*sefix;
        uclfix=estfix+1.96*sefix;
keep parmfix classval0 estfix sefix lclfix uclfix;
run;
proc sort data=totfix;by parmfix classval0;run;
data fix;set totfix;
        a1=&a1; a2=&a2; b1=&b1; b2=&b2;
        if parmfix='Intercept' & classval0='1' then do;
            biasfix=estfix-a1;
            msefix=biasfix**2;
        if(lclfix<=a1<=uclfix)then covfix=100;else covfix=0;
        output;
        end;
        if parmfix='Intercept' & classval0='2' then do;
            biasfix=estfix-a2;
            msefix=biasfix**2;
        if(lclfix<=a2<=uclfix)then covfix=100;else covfix=0;
        output;
        end;
        if parmfix='x1' then do;
            biasfix=estfix-b1;
            msefix=biasfix**2;
        if(lclfix<=b1<=uclfix)then covfix=100;else covfix=0;
        output;
        end;
        if parmfix='x2' then do;
            biasfix=estfix-b2;
            msefix=biasfix**2;
        if(lclfix<=b2<=uclfix)then covfix=100;else covfix=0;
        output;
        end;
keep parmfix classval0 estfix sefix biasfix msefix covfix;
run;
proc univariate data=fix noprint ;by parmfix classval0;
        var estfix sefix biasfix msefix covfix;
output out=m_fix mean=m_estfix m_sefix m_biasfix m_msefix m_covfix;
run;
```

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2.3 การเขียนโปรแกรมในการประมวลผลข้อมูลสำหรับวิธี GEE

```
proc genmod data=sim;by rep;
        class subject ;
        model y=x1 x2/dist=mult link=cumlogit;
        repeated subject=subject/type=ind;
ods output geeemppest=totgee modelfit=fitgee;
run;
data fitgee;set fitgee;
        neg2logL=-2*value;
keep criterion neg2logL;
run;
proc univariate data=fitgee noprint;
        var neg2logL;
output out=m_fitgee mean=fitgee;
run;
data totgee;set totgee;
        parmgee=parm;estgee=estimate;segee=stderr;
        lclgee=lowercl;uclgee=uppercl;
keep parmgee estgee segee lclgee uclgee;
run;
proc sort data=totgee;by parmgee;run;
data gee;set totgee;
        a1=&a1; a2=&a2; b1=&b1; b2=&b2;
        if parmgee='Intercept1' then do;
           biasgee=estgee-a1;
           msegee=biasgee**2;
        if(lclgee<=a1<=uclgee)then covgee=100;else covgee=0;
        output;
        end;
        if parmgee='Intercept2' then do;
           biasgee=estgee-a2;
           msegee=biasgee**2;
        if(lclgee<=a2<=uclgee)then covgee=100;else covgee=0;
        output;
        end;
        if parmgee='x1' then do;
           biasgee=estgee-b1;
           msegee=biasgee**2;
        if(lclgee<=b1<=uclgee)then covgee=100;else covgee=0;
        output;
        end;
        if parmgee='x2' then do;
           biasgee=estgee-b2;
           msegee=biasgee**2;
        if(lclgee<=b2<=uclgee)then covgee=100;else covgee=0;
        output;
        end;
keep parmgee level estgee segee biasgee msegee covgee;
run;
proc univariate data=gee noprint ;by parmgee;
        var  estgee segee biasgee msegee covgee;
output out=m_gee mean= m_estgee m_segee m_biasgee m_msegee m_covgee;
run;
```

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2.4 การเขียนโปรแกรมในการประมวลผลข้อมูลสำหรับวิธี MQL

```
proc glimmix data=sim method=mmpl;by rep;
        class subject;
        model y=x1 x2/s dist=mult link=clogit;
        random intercept/subject=subject;
ods output parameterestimates=totmql covparms=sumql
        fitstatistics=fitmql;
run;
proc univariate data=fitmql noprint;
        var value;
output out=m_fitmql mean=fitmql;
run;
data sumql;set sumql;
        parmmql=translate(subject,'x3','subject');
        estmql=estimate;semql=stderr;
        lclmql=estmql-1.96*semql;uclmql=estmql+1.96*semql;
keep rep parmmql estmql semql lclmql uclmql;
run;
data totmql;set totmql;
        level=y;parmmql=effect;estmql=estimate;semql=stderr;
        lclmql=estmql-tinv(0.975,df)*semql;
        uclmql=estmql+tinv(0.975,df)*semql;
keep rep level parmmql estmql semql lclmql uclmql;
run;
data totmql;merge totmql sumql;by rep parmmql;run;
proc sort data=totmql;by parmmql level;run;
data mql;set totmql;a1=&a1; a2=&a2; b1=&b1; b2=&b2; s2u=&s2u;
        if parmmql='Intercept' & level='1' then do;
            biasmql=estmql-a1;msemql= biasmql**2;
        if(lclmql<=a1<=uclmql)then covmql=100;else covmql=0;output;
        end;
        if parmmql='Intercept'& level='2' then do;
            biasmql=estmql-a2;
            msemql=biasmql**2;
        if(lclmql<=a2<=uclmql)then covmql=100;else covmql=0;output;
        end;
        if parmmql='x1' then do;
            biasmql=estmql-b1;
            msemql=biasmql**2;
        if(lclmql<=b1<=uclmql)then covmql=100;else covmql=0;output;
        end;
        if parmmql='x2' then do;
        biasmql=estmql-b2;
        msemql=biasmql**2;
        if(lclmql<=b2<=uclmql)then covmql=100;else covmql=0;output;
        end;
        if parmmql='x3' then do;
        biasmql=estmql-s2u;
        msemql=biasmql**2;
        if(lclmql<=s2u<=uclmql)then covmql=100;else covmql=0;output;
        end;
keep parmmql level estmql semql biasmql msemql covmql;
run;
proc univariate data=mql noprint ;by parmmql level;
        var estmql semql biasmql msemql covmql;
output out=m_mql mean=m_estmql m_semql m_biasmql m_msemql m_covmql;
run;
```

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2.5 การเขียนโปรแกรมในการประมวลผลข้อมูลสำหรับวิธี AGQ

```
proc nlmixed data=sim;by rep;
        parms a1=&a1 a2=&a2 b1=&b1 b2=&b2 s2u=&s2u;
        eta=b1*x1+b2*x2+u;
        if (y=1) then pi=1/(1+exp(-a1-eta));
        else if (y=2) then pi=1/(1+exp(-a2-eta))-1/(1+exp(-a1-eta));
        else if (y=3) then pi=1-1/(1+exp(-a2-eta));
        if (pi>1E-8) then ll=log(pi); else ll=-1E100;
        model y~general(ll);
        random u~normal(0,s2u) subject=subject;
ods output parameterestimates=totgq fitstatistics=fitgq;
run;
proc sort data=fitgq;by descr;run;
proc univariate data=fitgq noprint ;by descr;
        var value;
output out=m_fitgq mean=fitgq;
run;
data totgq;set totgq;
        parmgq=parameter;
        estgq=estimate;
        segq=standarderror;
        lclgq=lower;uclgq=upper;
keep rep parmgq estgq segq lclgq uclgq;
run;
proc sort data=totgq;by parmgq;run;
data gq;set totgq;
        a1=&a1; a2=&a2; b1=&b1; b2=&b2; s2u=&s2u;
        if parmgq='a1' then do;
            biasgq=estgq-a1;
            msegq=biasgq**2;
        if(lclgq<=a1<=uclgq)then covgq=100;else covgq=0;output;
        end;
        if parmgq='a2' then do;
            biasgq=estgq-a2;
            msegq=biasgq**2;
        if(lclgq<=a2<=uclgq)then covgq=100;else covgq=0;output;
        end;
        if parmgq='b1' then do;
            biasgq=estgq-b1;msegq=biasgq**2;
        if(lclgq<=b1<=uclgq)then covgq=100;else covgq=0;output;
        end;
        if parmgq='b2' then do;
            biasgq=estgq-b2;
            msegq=biasgq**2;
        if(lclgq<=b2<=uclgq)then covgq=100;else covgq=0;output;
        end;
        if parmgq='s2u' then do;
            biasgq=estgq-s2u;
            msegq=biasgq**2;
        if(lclgq<=s2u<=uclgq)then covgq=100;else covgq=0;output;
        end;
keep parmgq estgq segq biasgq msegq covgq;
run;
proc univariate data=gq noprint ;by parmgq;
        var  estgq segq biasgq msegq covgq;
output out=m_gq mean= m_estgq m_segq m_biasgq m_msegq m_covgq;
run;
```

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2.7 การเขียนโปรแกรมในการประมวลผลข้อมูลสำหรับวิธี PQL

```
proc glimmix data=sim method=rspl;by rep;
        class subject;
        model y=x1 x2/s dist =mult link=clogit;
        random intercept/subject=subject;
ods output parameterestimates=totpql covparms=supql
        fitstatistics=fitpql;
run;
proc univariate data=fitpql noprint;
        var value;output out=m_fitpql mean=fitpql;
run;
data supql;set supql;
        parmpql=translate(subject,'x3','subject');
        estpql=estimate;sepql=stderr;
        lclpql=estpql-1.96*sepql;uclpql=estpql+1.96*sepql;
keep rep parmpql estpql sepql lclpql uclpql;
run;
data totpql;set totpql;
        level=y;parmpql=effect;estpql=estimate;sepql=stderr;
        lclpql=estpql-tinv(0.975,df)*sepql;
        uclpql=estpql+tinv(0.975,df)*sepql;
keep rep level parmpql estpql sepql lclpql uclpql;
run;
data totpql;merge totpql supql;by rep parmpql;run;
proc sort data=totpql;by parmpql level;run;
data pql;set totpql;
        a1=&a1; a2=&a2; b1=&b1; b2=&b2; s2u=&s2u;
        if parmpql='Intercept'& level='1' then do;
            biaspql=estpql-a1;msepql= biaspql**2;
        if(lclpql<=a1<=uclpql)then covpql=100;else covpql=0;output;
        end;
        if parmpql='Intercept' & level='2' then do;
            biaspql=estpql-a2;msepql=biaspql**2;
        if(lclpql<=a2<=uclpql)then covpql=100;else covpql=0;output;
        end;
        if parmpql='x1' then do;
            biaspql=estpql-b1;msepql=biaspql**2;
        if(lclpql<=b1<=uclpql)then covpql=100;else covpql=0;output;
        end;
        if parmpql='x2' then do;
            biaspql=estpql-b2;msepql=biaspql**2;
        if(lclpql<=b2<=uclpql)then covpql=100;else covpql=0;output;
        end;
        if parmpql='x3' then do;
            biaspql=estpql-s2u;msepql=biaspql**2;
        if(lclmpl<=s2u<=uclpql)then covpql=100;else covpql=0;output;
        end;
keep level parmpql estpql sepql biaspql msepql covpql;
run;
proc univariate data=pql noprint ;by parmpql level;
        var estpql sepql biaspql msepql covpql;
output out=m_pql mean=m_estpql m_sepql m_biaspql m_msepql m_covpql;
run;
```

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

2.8 การเขียนโปรแกรมในการเก็บผลลัพธ์ที่ได้จากการประมวลผลทั้ง 5 วิธี

```
proc printto log=logfile print='d:\50-5-0.17.txt' new;run;
proc print data=m_fix;run;
proc print data=m_gee;run;
proc print data=m_mql;run;
proc print data=m_gq;run;
proc print data=m_pql;run;
proc printto log=logfile print='d:\fit50-5-0.17.txt' new;run;
proc print data=m_fitfix;run;
proc print data=m_fitgee;run;
proc print data=m_fitmql;run;
proc print data=m_fitgq;run;
proc print data=m_fitpql;run;
proc datasets;
        delete sim totfix totgee totmql totgq totpql sumql supql fix
            gee mql gq pql;
run;
%mend case;
```

2.9 การเรียกใช้แมคโครภายใต้เงื่อนไขต่าง ๆ

2.9.1 กรณีของจำนวนกลุ่มตัวอย่าง เท่ากับ 50

```
%case(nrep=1000,N=50,ni=5,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=5,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=5,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=7,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=7,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=7,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=10,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=10,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=50,ni=10,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
```

2.9.2 กรณีของจำนวนกลุ่มตัวอย่าง เท่ากับ 100

```
%case(nrep=1000,N=100,ni=5,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=5,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=5,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=7,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=7,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=7,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=10,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=10,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=100,ni=10,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
```

2.9.3 กรณีของจำนวนกลุ่มตัวอย่าง เท่ากับ 250

```
%case(nrep=1000,N=250,ni=5,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=5,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=5,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=7,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=7,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=7,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=10,s2u=0.17,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=10,s2u=0.82,a1=-1,a2=1,b1=0.5,b2=0.5);
%case(nrep=1000,N=250,ni=10,s2u=2.19,a1=-1,a2=1,b1=0.5,b2=0.5);
```

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ประวัติผู้วิจัย

ชื่อ-สกุล นางสาวชญาณัฐ โพธิ์นอก

ที่อยู่ 95 หมู่ 17 ตำบลหนองแวง กิ่งอำเภอเทพารักษ์ จังหวัดนครราชสีมา 30210

ประวัติการศึกษา

พ.ศ. 2549 สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต สาขาสถิติ เกียรตินิยมอันดับ 1 มหาวิทยาลัยศิลปากร

พ.ศ. 2550 ศึกษาต่อระดับปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ มหาวิทยาลัยศิลปากร

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์