

การประมาณพารามิเตอร์ของแผนการทดลองสุ่มสมบูรณ์เชิงสุ่มด้วยวิธีการเฉลี่ยค่าองค์ประกอบ
ความแปรปรวน



นางสาวกนกวรรณ ลิ้โรจนาประภา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2544

ISBN 974-03-1621-2

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

PARAMETER ESTIMATION FOR RANDOM-EFFECT COMPLETELY RANDOMIZED
DESIGN MODEL WITH VARIANCE COMPONENTS AVERAGING METHOD

Miss Kanogkan Leerojanaprapa

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Statistics

Department of Statistics

Faculty of Commerce and Accountancy

Chulalongkorn University

Academic Year 2001

ISBN 974-03-1621-2

กนกกรรณ ลีโรจนประภา : การประมาณพารามิเตอร์ของแผนการทดลองสุ่มสมบูรณ์เชิงสุ่มด้วยวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวน. (PARAMETER ESTIMATION FOR RANDOM-EFFECT COMPLETELY RANDOMIZED DESIGN MODEL WITH VARIANCE COMPONENTS AVERAGING METHOD)

อ. ที่ปรึกษา : รศ. ดร. สุพล คุรงค์วัฒนา, 170 หน้า. ISBN 974-03-1621-2.

วัตถุประสงค์ของการวิจัยในครั้งนี้ เพื่อหาวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนที่เหมาะสมจากวิธีพื้นฐานทั้ง 3 วิธี (ตัวประมาณแบบความควรจะเป็นสูงสุด, ตัวประมาณแบบกำลังสองไม่แปรเปลี่ยน และตัวประมาณแบบความควรจะเป็นสูงสุดแบบมีข้อจำกัด) โดยศึกษาวิธีการถ่วงน้ำหนัก 4 วิธีดังนี้ 1)วิธีการถ่วงน้ำหนักที่เท่ากัน 2)วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของ $\bar{y}_{..}$ 3)วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของตัวประมาณ σ^2 หรือองค์ประกอบความแปรปรวน และ4)วิธีการถ่วงน้ำหนักโดยอาศัยค่าสัมบูรณ์ต่ำสุดซึ่งเป็นการวนซ้ำโดยอาศัยเทคนิคสมการเชิงเส้น เพื่อแก้ปัญหาสมการเป้าหมายที่ทำให้ผลรวมค่าสัมบูรณ์ของความคลาดเคลื่อนมีค่าต่ำสุด ใน การศึกษาใช้ตัวแบบ $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ เมื่อ $i = 1, 2, \dots, a$ และ $j = 1, 2, \dots, n$ โดยที่ y_{ij} เป็นสัญลักษณ์แทนค่าสังเกตที่ j ซึ่งได้รับปัจจัยในระดับที่ i μ แทนค่าเฉลี่ยรวม τ_i แทนผลกระทบเชิงสุ่มของปัจจัยทดลองที่ i และมีการแจกแจงปกติที่อิสระกันซึ่งมีค่าเฉลี่ย 0 ความแปรปรวน σ^2 ε_{ij} แทนความคลาดเคลื่อนเชิงสุ่มของค่าสังเกตที่ j ปัจจัยทดลองที่ i และมีการแจกแจงปกติที่อิสระกันซึ่งมีค่าเฉลี่ย 0 ความแปรปรวน σ^2 เมื่อ a และ n แทนจำนวนระดับของปัจจัย และจำนวนค่าสังเกตในแต่ละระดับของปัจจัยตามลำดับ ในการศึกษาได้ทำการจำลองข้อมูลจากเทคนิคมอนติคาร์โล ด้วยโปรแกรม S-PLUS 2000 โดยกำหนดและสร้างพารามิเตอร์ σ^2 , μ และ σ^2 ขึ้น และใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบตัวประมาณ

จากการศึกษาพบว่าการประมาณองค์ประกอบความแปรปรวน σ^2 และ $\sigma^2 + \sigma^2$ เหมาะสมที่สุดเมื่อนำเฉพาะตัวประมาณแบบความควรจะเป็นสูงสุด และตัวประมาณแบบกำลังสองไม่แปรเปลี่ยนเท่านั้นมาใช้ในการเฉลี่ยถ่วงน้ำหนักในการประมาณ σ^2 ดังกล่าว เมื่อ k มีค่าน้อย และจำนวนระดับของปัจจัยก็มีค่าน้อยเช่นกัน พบว่า วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของตัวประมาณ σ^2 เป็นตัวประมาณที่ดีที่สุด เมื่อ k มีค่าน้อย จำนวนระดับของปัจจัยมาก และจำนวนค่าสังเกตในแต่ละระดับของปัจจัยน้อย วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของ $\bar{y}_{..}$ เป็นตัวประมาณที่ดีที่สุด ส่วนกรณี k มีค่าน้อย จำนวนระดับของปัจจัย และจำนวนค่าสังเกตในแต่ละระดับของปัจจัยมาก วิธีค่าสัมบูรณ์ต่ำสุดเป็นตัวประมาณที่ดีที่สุด เมื่อ k มาก กรณีส่วนใหญ่ตัวประมาณที่ดีที่สุดคือวิธีค่าสัมบูรณ์ต่ำสุด ยกเว้นเฉพาะกรณีเมื่อจำนวนระดับของปัจจัยน้อย ที่วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของตัวประมาณ σ^2 เป็นตัวประมาณที่ดีที่สุด ในการประมาณ $\sigma^2 + \sigma^2$ เมื่อ k มีค่าน้อย กรณีจำนวนระดับของปัจจัย และจำนวนค่าสังเกตในแต่ละระดับของปัจจัยน้อย พบว่าวิธีค่าสัมบูรณ์ต่ำสุดเป็นตัวประมาณที่ดีที่สุด เมื่อ k มีค่าน้อย จำนวนระดับของปัจจัยมาก และจำนวนค่าสังเกตในแต่ละระดับของปัจจัยน้อย วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของ $\bar{y}_{..}$ เป็นตัวประมาณที่ดีที่สุด และกรณี k มีค่าน้อย และจำนวนค่าสังเกตในแต่ละระดับของปัจจัยมาก วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของตัวประมาณองค์ประกอบความแปรปรวนเป็นตัวประมาณที่ดีที่สุด เมื่อ k มีค่ามาก การประมาณ $\sigma^2 + \sigma^2$ ให้ผลสรุปเช่นเดียวกับการประมาณ σ^2

ภาควิชา สถิติ

สาขาวิชา สถิติ

ปีการศึกษา 2544

ลายมือชื่อนิสิต กนกกรรณ ลีโรจนประภา

ลายมือชื่ออาจารย์ที่ปรึกษา

4282151926 : MAJOR STATISTICS

KEY WORD : VARIANCE COMPONENTS / COMPLETELY RANDOMIZED DESIGN / AVERAGING METHOD

KANOGKAN LEEROJANAPRAPA : PARAMETER ESTIMATION FOR RANDOM-EFFECT COMPLETELY RANDOMIZED DESIGN MODEL WITH VARIANCE COMPONENTS AVERAGING METHOD. THESIS ADVISOR : ASSOCIATE PROFESSOR. SUPOL DURONGWATANA, Ph.D. 170 pp. ISBN 974-03-1621-2.

The objective of this study is to compare the methods of averaging variance components based on Maximum Likelihood (ML), Invariance Quadratic (IQ) and Restricted Maximum Likelihood (REML) estimators. There are 4 weighted averaging approaches. The first approach takes account of equal weights. Where as the weights that based on the estimation of variance for $\bar{y}_{..}$ and the weights that based on the estimated variance of estimator σ_{τ}^2 or variance components are considered in the second and the third approach respectively. For the fourth approach, the weights of least absolute value are based on iterative procedure using linear programming technique to solve for the minimised sum of the absolute deviations. The model used in this study is $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ when $i = 1, 2, \dots, a$ and $j = 1, 2, \dots, n$ where y_{ij} is the j^{th} observation for the i^{th} level of factor, μ is the grand mean, τ_i is the i^{th} random effect of factor and is independently distributed with mean 0 and variance σ_{τ}^2 , ε_{ij} is the random error of the j^{th} observation for the i^{th} level of factor and is also independently distributed with mean 0 and variance σ_{ε}^2 , where as a and n represent the number of levels for factor and the number of replication for treatment respectively. To generate the data for this study, The Monte Carlo simulation is applied by using S-PLUS 2000 package where σ_{ε}^2 , μ and σ_{τ}^2 are parameters in the model, all of which are specified and generated then, the mean square error is determine for comparing the estimators.

The result for the estimation of variance components σ_{τ}^2 and $\sigma_{\tau}^2 + \sigma_{\varepsilon}^2$ is optimal when only ML and IQ estimation are used for weighted averaging. For the estimation of σ_{τ}^2 , when k is small and the number of levels for factor is also small, and it is found in this case that the weighting method using estimated variance of estimator σ_{τ}^2 is the best estimator. In the case of small k and the number of levels for factor is large together with the number of replication for treatment is small, the weighting method using estimated variance of $\bar{y}_{..}$ is found as the best estimator. In addition to this small k case, when the number of levels for factor and the number of replication for treatment are large, the least absolute value method is found as the best estimator. For large k , the least absolute value method is the best estimator for almost all of the cases but when the number of levels for factor is small, the weighting method using estimated variance of estimator σ_{τ}^2 is the best estimator. For the estimation of $\sigma_{\tau}^2 + \sigma_{\varepsilon}^2$, when k is small and the number of levels for factor and the number of replication for treatment are small, it is found that the least absolute value method is the best estimator. In the case of small k , when the number of levels for factor is large and the number of replication for treatment is small, the weighting method using estimated variance of $\bar{y}_{..}$ is found as the best estimator. And also, in case of small k , when the number of replication for treatment is large, the weighting method using estimated variance of variance components is the best estimator. As final findings, when k is large what are found in the estimation of $\sigma_{\tau}^2 + \sigma_{\varepsilon}^2$ are similar to the results obtained in the estimation of σ_{τ}^2 .

Department Statistics

Field of study Statistics

Academic year 2001

Student's signature *kanogkan leerojana prapa*

Advisor's signature *SP Dgk*

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของรองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำแนะนำตลอดจนช่วยเหลือแก้ไขข้อบกพร่องต่างๆ ด้วยดีตลอดมา จนกระทั่งวิทยานิพนธ์เสร็จสมบูรณ์ จึงใคร่ขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้

สำหรับการศึกษาของผู้วิจัยที่สามารถสำเร็จไปได้ด้วยดีนั้นเพราะเนื่องมาจากได้รับการประสิทธิ์ประสาทความรู้จากคณาจารย์ประจำภาควิชาสถิติทุกท่าน ประกอบกับได้รับการสนับสนุนทุนการศึกษาจากโครงการพัฒนาอาจารย์ สาขาขาดแคลนเพื่อศึกษาในประเทศ จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง นอกจากนี้ทุนการวิจัยบางส่วนยังได้รับมาจากทุนอุดหนุนการวิจัยของบัณฑิตวิทยาลัย จึงขอขอบพระคุณมา ณ ที่นี้

ท้ายนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดา-มารดา ซึ่งสนับสนุนด้านการเงินและให้กำลังใจแก่ผู้วิจัยเสมอมาจนกระทั่งสำเร็จการศึกษา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญ	ช
สารบัญตาราง	ฎ
สารบัญภาพ	ณ
บทที่	
1 บทนำ	
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	3
1.3 สมมติฐานการวิจัย	3
1.4 ข้อตกลงเบื้องต้น	3
1.5 ขอบเขตของการวิจัย	5
1.6 คำจำกัดความที่ใช้ในงานวิจัย	6
1.7 ประโยชน์ที่คาดว่าจะได้รับ	6
2 การประมาณค่า	
2.1 ตัวประมาณด้วยวิธีพื้นฐาน	8
2.1.1 ตัวประมาณความควรจะเป็นสูงสุด (Maximum likelihood : ML) ...	9
2.1.2 ตัวประมาณแบบกำลังสองไม่แปรเปลี่ยน (Invariance quadratic estimator : IQE)	15
2.1.3 ตัวประมาณแบบความควรจะเป็นสูงสุดแบบมีข้อจำกัด (Restricted maximum likelihood : REML)	21
2.2 ตัวประมาณด้วยวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวน	23
2.2.1 วิธีหาค่า a_i และ b_i ที่เหมาะสม	23
2.2.1.1 วิธีการถ่วงน้ำหนักที่เท่ากัน	24
2.2.1.2 วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวนของ $\bar{y}_{..}$	24

สารบัญ (ต่อ)

	หน้า
2.2.1.3 วิธีการถ่วงน้ำหนักโดยอาศัยค่าประมาณความแปรปรวน ของตัวประมาณองค์ประกอบความแปรปรวน (ประกอบ ด้วย σ_1^2 และ σ_2^2)	25
2.2.2 วิธีค่าสัมบูรณ์ต่ำสุด (Least absolute value method)	31
2.3 เกณฑ์การเปรียบเทียบวิธีการประมาณองค์ประกอบความแปรปรวน	35
2.3.1 กรณีการประมาณพารามิเตอร์ σ_1^2	35
2.3.2 กรณีการประมาณองค์ประกอบความแปรปรวน	35
3 วิธีดำเนินการวิจัย	36
3.1 การผลิตเลขสุ่มจากรูปแบบการแจกแจงประชากรแบบปกติ	36
3.2 การคำนวณค่าองค์ประกอบความแปรปรวน	38
3.2.1 วิธีพื้นฐาน	38
3.2.1.1 ตัวประมาณแบบความควรจะเป็นสูงสุด (Maximum likelihood : ML)	38
3.2.1.2 ตัวประมาณแบบกำลังสองไม่แปรเปลี่ยน (Invariance quadratic estimator : IQE)	38
3.2.1.3 ตัวประมาณแบบความควรจะเป็นสูงสุดแบบมีข้อจำกัด (Restricted maximum likelihood : REML)	38
3.2.2 วิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวน	39
3.2.2.1 กลุ่ม 1 (เฉลี่ยวิธี ML, IQE และ REML)	39
3.2.2.2 กลุ่ม 2 (เฉลี่ยวิธี ML และ IQE)	41
3.3 การเปรียบเทียบตัวประมาณองค์ประกอบความแปรปรวน	44
3.3.1 กรณีการประมาณพารามิเตอร์ σ_1^2	44
3.3.2 กรณีการประมาณองค์ประกอบความแปรปรวน	44
3.4 ขั้นตอนการทำงานของโปรแกรม	45
4 ผลการวิจัย	47
4.1 ผลการเปรียบเทียบตัวประมาณพารามิเตอร์ σ_1^2	49
4.1.1 ผลการเปรียบเทียบตัวประมาณพารามิเตอร์ σ_1^2 ด้วยวิธีพื้นฐาน	49

สารบัญ (ต่อ)

	หน้า
4.1.2 ผลการเปรียบเทียบตัวประมาณพารามิเตอร์ σ^2 ด้วยวิธีการเฉลี่ย ค่าองค์ประกอบความแปรปรวนกลุ่ม1	49
4.1.3 ผลการเปรียบเทียบตัวประมาณพารามิเตอร์ σ^2 ด้วยวิธีการเฉลี่ย ค่าองค์ประกอบความแปรปรวนกลุ่ม2	50
4.1.4 ผลการเปรียบเทียบตัวประมาณพารามิเตอร์ σ^2 ด้วยวิธีพื้นฐาน กับวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนที่เหมาะสม ในมุมมอง ของการนำไปใช้	52
4.2 ผลการเปรียบเทียบตัวประมาณองค์ประกอบความแปรปรวน	93
4.2.1 ผลการเปรียบเทียบตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีพื้นฐาน	93
4.2.2 ผลการเปรียบเทียบตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนกลุ่ม1	93
4.2.3 ผลการเปรียบเทียบตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนกลุ่ม2	95
4.2.4 ผลการเปรียบเทียบตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีพื้นฐานกับวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวน ที่เหมาะสม ในมุมมองของการนำไปใช้	98
5 สรุปผลการวิจัยและข้อเสนอแนะ	139
5.1 สรุปผลการวิจัย	140
5.1.1 ตัวประมาณด้วยวิธีพื้นฐาน	140
5.1.2 ตัวประมาณด้วยวิธีเฉลี่ยค่าองค์ประกอบความแปรปรวนกลุ่ม1	141
5.1.3 ตัวประมาณด้วยวิธีเฉลี่ยค่าองค์ประกอบความแปรปรวนกลุ่ม2	141
5.1.4 การเปรียบเทียบระหว่างตัวประมาณด้วยวิธีพื้นฐานและตัวประมาณ ด้วยวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนทั้งกลุ่ม1 และกลุ่ม2	142
5.2 ข้อเสนอแนะ	146
5.2.1 ด้านการนำไปใช้	146
5.2.2 ด้านการศึกษาวิจัย	146
รายการอ้างอิง	148

สารบัญ (ต่อ)

	หน้า
ภาคผนวก	150
ภาคผนวก ก	150
ภาคผนวก ข	157
ภาคผนวก ค	159
ประวัติผู้เขียนวิทยานิพนธ์	170

สารบัญตาราง

ตาราง	หน้า
ตารางที่ 1	4
ตารางที่ 3.1	41
ตารางที่ 3.2	43
ตารางที่ 4.1	53
ตารางที่ 4.2	55
ตารางที่ 4.3	57
ตารางที่ 4.4	59
ตารางที่ 4.5	61
ตารางที่ 4.6	63
ตารางที่ 4.7	65
ตารางที่ 4.8	65

สารบัญญภาพ (ต่อ)

หน้า

รูปที่ 4.36	แสดงการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE) ของตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีพื้นฐานกับวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนที่เหมาะสม ในมุมมองของการนำไปใช้ ตามจำนวนระดับของปัจจัยที่ใช้ในการทดลอง(a) และจำนวนค่าสังเกตในแต่ละระดับของปัจจัย(n) เมื่อ $k = 0.1$	130
รูปที่ 4.37	แสดงการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE) ของตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีพื้นฐานกับวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนที่เหมาะสม ในมุมมองของการนำไปใช้ ตามจำนวนระดับของปัจจัยที่ใช้ในการทดลอง(a) และจำนวนค่าสังเกตในแต่ละระดับของปัจจัย(n) เมื่อ $k = 0.5$	132
รูปที่ 4.38	แสดงการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE) ของตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีพื้นฐานกับวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนที่เหมาะสม ในมุมมองของการนำไปใช้ ตามจำนวนระดับของปัจจัยที่ใช้ในการทดลอง(a) และจำนวนค่าสังเกตในแต่ละระดับของปัจจัย(n) เมื่อ $k = 1$	134
รูปที่ 4.39	แสดงการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE) ของตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีพื้นฐานกับวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนที่เหมาะสม ในมุมมองของการนำไปใช้ ตามจำนวนระดับของปัจจัยที่ใช้ในการทดลอง(a) และจำนวนค่าสังเกตในแต่ละระดับของปัจจัย(n) เมื่อ $k = 4$	136
รูปที่ 4.40	แสดงการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE) ของตัวประมาณองค์ประกอบความแปรปรวน ด้วยวิธีพื้นฐานกับวิธีการเฉลี่ยค่าองค์ประกอบความแปรปรวนที่เหมาะสม ในมุมมองของการนำไปใช้ ตามจำนวนระดับของปัจจัยที่ใช้ในการทดลอง(a) และจำนวนค่าสังเกตในแต่ละระดับของปัจจัย(n) เมื่อ $k = 9$	138