206413

สิวะโชติ ศรีสุทธิยากร : การจำลองตัวแบบความถดถอยเชิงลำดับชั้น เมื่อการแจกแจงความคลาดเคลื่อนสุ่มไม่ได้มีการแจกแจงปกติ (A SIMULATION OF HIERARCHICAL REGRESSION MODEL WITH NONNORMALITY RANDOM ERROR DISTRIBUTED) อ.ที่ปรึกษา: ร.ศ. ดร. สุพล ดุรงค์วัฒนา, 216 หน้า.

วัตถุประสงค์ของการวิจัยเพื่อศึกษาประสิทธิภาพในการประมาณค่าพารามิเตอร์ในแต่ละกลุ่มของตัวแบบการถดถอยเชิงลำดับชั้น เมื่อการแจกแจงของความคลาดเคลื่อนสุ่มในระดับที่ 2 มีการแจกแจงแบบไม่ปกติ และศึกษาวิธีปรับแก้สูตรคำนวณค่าคลาดเคลื่อนมาตรฐาน โดยใช้วิธีประมาณค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่งของฮูเบอร์ไวท์ (Huber/White Robust Standard Error (ROBUST)) ว่าสามารถประมาณค่าคลาดเคลื่อนมาตรฐานได้ผิดพลาดน้อยกว่าค่าคลาดเคลื่อนมาตรฐานที่ได้จากตัวแบบ(MODEL) เมื่อความคลาดเคลื่อนสุ่มในระดับที่ 2 มีการแจกแจงแบบไม่ปกติหรือไม่ วิธีประมาณที่ใช้ในการศึกษานี้ได้แก่วิธี Iterative Generalized Least Square (IGLS) และ วิธี Restricted Iterative Generalized Least Square (RIGLS) จะศึกษาโดยแยกพารามิเตอร์ในตัวแบบออกเป็น 3 กลุ่มคือ กลุ่มพารามิเตอร์อิทธิพลคงที่ (Fix Effects Parameters) กลุ่มพารามิเตอร์ความแปรปรวนในระดับที่ 1 (Level-1 Variance) และกลุ่มพารามิเตอร์ส่วนประกอบความแปรปรวนในระดับที่ 2 (Level-2 Variance Component) เกณฑ์การเปรียบเทียบที่ใช้คือเกณฑ์ค่าความเอนเอียงสัมพัทธ์ (Relative Biases (RB)) และเกณฑ์ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองระหว่างค่าประมาณค่าคลาดเคลื่อนมาตรฐานกับค่าคลาดเคลื่อนมาตรฐานแบบมอนติคาร์โลสัมพัทธ์(Root of Relative Mean Square Error (RMSE) between Estimated Standard Error and Monte Carlo Standard Error) เพื่อเปรียบเทียบความผิดพลาดในการประมาณค่าพารามิเตอร์ และค่าความคลาดเคลื่อนมาตรฐานตามลำดับ

ในการวิจัยนี้กำหนดให้การแจกแจงความคลาดเคลื่อนสุ่มในตัวแปรตามในระดับที่ 1 มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 ความแปรปรวนเป็น 0.5 และในระดับที่ 2 มีการแจกแจงแบบเอ็กซ์โพเนนเชียลพาวเวอร์แบบไม่สมมาตรสองตัวแปร (Bivariate Asymmetric Exponential Power Distribution) ที่มีค่าเฉลี่ย $\underline{0}$ และเมตริกซ์ความแปรปรวนร่วม (Σ) พารามิเตอร์ความโด่ง (α) และพารามิเตอร์ความเบ้ (κ) โดยให้ค่าความแปรปรวนของความคลาดเคลื่อนสุ่มในระดับที่ 2 กำหนดตามค่าสัมประสิทธิ์สหสัมพันธ์ภายในกลุ่ม (Intra Class Correlation) ซึ่งในงานวิจัยนี้กำหนดให้มีค่าเท่ากับ 0.1, 0.2 และ 0.3 ในการศึกษาผลกระทบจากลักษณะการแจกแจงจะพิจารณาใน 2 กรณี กรณีที่ 1 กำหนดให้พารามิเตอร์ความเบ้คงที่โดยมีค่าเท่ากับ 1 และพารามิเตอร์ความโด่งมีค่าเท่ากับ 1, 1.25, 1.6, 3, 8 และ 32 ตามลำดับ กรณีที่ 2 กำหนดพารามิเตอร์ความโด่งคงที่โดยมีค่าเท่ากับ 2 และพารามิเตอร์ความเบ้มีค่าเท่ากับ 0.8518, 0.6437 และ0.1 ตามลำดับ ขนาดตัวอย่างที่ใช้แบ่งออกเป็น 2 ระดับ ขนาดตัวอย่างในระดับที่ 1 คือ 5, 10 และ15 และขนาดตัวอย่างในระดับที่ 2 คือ 15, 30 และ 50 ข้อมูลที่ใช้ในการวิจัยใช้โปรแกรมคอมพิวเตอร์จำลองด้วยเทคนิคมอนติคาร์โลกระทำซ้ำ 500 ครั้ง ในแต่ละสถานการณ์ ซึ่งผลการวิจัยสรุปได้ดังนี้

1. เมื่อพิจารณาจากค่า RB พบว่าเมื่อการแจกแจงความคลาดเคลื่อนสุ่มในระดับที่ 2 มีการแจกแจงแบบไม่ปกติ ค่าประมาณของพารามิเตอร์ในตัวแบบจะไม่ได้รับผลกระทบจากปัจจัยดังกล่าว วิธี IGLS และวิธี RIGLS จะประมาณค่าของพารามิเตอร์อิทธิพลคงที่ และความแปรปรวนในระดับที่ 1 ได้ใกล้เคียงกัน แต่วิธี RIGLS จะประมาณค่าส่วนประกอบความแปรปรวนในระดับที่ 2 ได้แม่นยำกว่าวิธี IGLS เสมอ ปัจจัยที่ส่งผลกระทบต่อความแม่นยำของการประมาณพารามิเตอร์อิทธิพลคงที่ และส่วนประกอบความแปรปรวนในระดับที่ 2 ได้แก่ ขนาดตัวอย่าง และค่าสัมประสิทธิ์สหสัมพันธ์ภายในกลุ่ม ส่วนปัจจัยที่ส่งผลกระทบต่อความแม่นยำในการประมาณพารามิเตอร์ความแปรปรวนในระดับที่ 1 ได้แก่ขนาดตัวอย่าง

2. เมื่อพิจารณาจากค่า RMSE พบว่าเมื่อการแจกแจงความคลาดเคลื่อนสุ่มในระดับที่ 2 มีการแจกแจงแบบไม่ปกติ ค่าประมาณของค่าคลาดเคลื่อนมาตรฐานของพารามิเตอร์ในตัวแบบทั้งจากวิธี MODEL และวิธี ROBUST มีความผิดพลาดอยู่ในระดับสูง โดยวิธี ROBUST จะมีความผิดพลาดในการประมาณค่าคลาดเคลื่อนมาตรฐานของพารามิเตอร์ส่วนประกอบความแปรปรวนในระดับที่ 2 น้อยกว่าวิธี MODEL ในกรณีที่การแจกแจงความคลาดเคลื่อนสุ่มในระดับที่ 2 มีความแบนราบมากกว่าปกติเท่านั้น ส่วนในกรณีอื่นๆวิธี MODEL จะประมาณค่าได้ผิดพลาดน้อยกว่าวิธี ROBUST ปัจจัยที่ส่งผลกระทบต่อการประมาณค่าคลาดเคลื่อนมาตรฐานของพารามิเตอร์อิทธิพลคงที่ และส่วนประกอบความแปรปรวนในระดับที่ 2 ได้แก่ ขนาดตัวอย่าง และค่าสัมประสิทธิ์สหสัมพันธ์ภายในกลุ่ม ส่วนปัจจัยที่ส่งผลกระทบต่อการประมาณค่าคลาดเคลื่อนมาตรฐานของพารามิเตอร์ความแปรปรวนในระดับที่ 1 ได้แก่ ขนาดตัวอย่าง

3. เมื่อพิจารณาจากค่า RB และค่า RMSE จากการวิจัยพบว่าขนาดตัวอย่างในระดับที่ 1 ควรมีขนาดมากกว่า 15 หน่วยขึ้นไป และขนาดตัวอย่างในระดับที่ 2 ควรมีขนาดมากกว่า 50 หน่วยขึ้นไป จึงจะมีความเหมาะสมในการประมาณค่าพารามิเตอร์ และค่าคลาดเคลื่อนมาตรฐานของตัวแบบ เมื่อการแจกแจงความคลาดเคลื่อนสุ่มในระดับที่ 2 มีการแจกแจงแบบไม่ปกติ

จ
206413

##4882283626 : MAJOR STATISTICS

KEY WORD: HIERARCHICAL REGRESSION MODEL/ NONNORMALITY/ ROBUST STANDARD ERROR/ ITERATIVE GENERALIZED LEAST SQUARE/IGLS/RESTRICTED ITERATIVE GENERALIZED LEAST SQUARE/ RIGLS.

SIWACHOAT SRISUTTIYAKORN: A SIMULATION OF HIERARCHICAL REGRESSION MODEL WITH NONNORMAL DISTRIBUTED RANDOM ERROR. THESIS ADVISOR: ASSOC. PROF. SUPOL DURONGWATANA, Ph.D., 216 pp.

The purpose of this research is to study the efficiency of parameter estimator in hierarchical regression models by using iterative generalized least square (IGLS), restricted iterative generalized least square (RIGLS) and the robustness of Huber/White robust standard error estimator (ROBUST), when non-normal distributed random errors in level-two are occurred. The parameters in model were divided into fix effects parameters, level-1 variance and level-2 variance component. The criterion in this study are Relative Biased (RB), Root of Relative Mean Square Error (RMSE) between Estimated Standard Error and Monte Carlo Standard Error were used to compare the accuracy of estimated parameters and standard error in such model.

In this research, the distribution of random errors in level-one are normal distribution with mean equal to 0 and variance equal to 0.5. The distribution of random errors in level-two are bivariate asymmetric exponential power distribution with mean equal to $\underline{0}$, variance-covariance matrix is Σ, kurtosis parameter is α and skewed parameter is κ.The variance of random errors in level-two are indicated by intra-class correlation (ICC) and divided into two separated cases. First, κ equals to 1 and α equals to 1, 1.25, 1.6, 3, 8, and 32. Secondly, α equals to 2 and κ equals to 0.8518, 0.6437, and 0.1. The level-1 sample-sizes which are used in this study are 5, 10, and 15, and the level-2 sample-sizes are 15, 30, and 50. The data for this experiment was generated through the Monte Carlo simulation technique and were conducted repeatedly 500 times for each case. The results of this research are as follow:

1. From RB criterion, non-normal random errors result from level-two yielded no effect to accuracy of estimated parameter. IGLS and RIGLS methods produced the same value of estimated fixed effects parameters and level-one variance parameter, however RIGLS has sustained a higher accurate estimated variance-covariance components parameters in level-two compared to IGLS. The influential factors of estimated fixed effects parameters and level-two variance-covariance component parameters are sample-size and ICC. As such, level-one variance parameter is also effected by the sample-size.

2. From RMSE criterion, non-normal random errors in level-two resulted in a relatively high inaccuracy of estimated standard error by both MODEL and ROBUST methods. For ROBUST, the estimated standard error result showed lower error compared to the MODEL, where level-two random errors have yielded a platykurtic profile. In other case, a MODEL can estimate less error than ROBUST. Estimated standard error of fix effects parameters and level-two variance-covariance parameters are effected by sample-size and ICC. As such, level-one variance parameter is also effected by the sample-size.

3. From RB and RMSE criterion, when the non-normal random errors in level-two were shown, then the level-one sample-size was suggested to obtain 15 units or more, while the level-two sample-size was suggested to have 50 units or more. These sample-size numbers yielded a preferable result of estimated parameter and standard errors in hierarchical regression model under the scope of this research.