

การแก้ไขปัญหาข้อมูลตอบสนองของแผนแบบการทดลองสุ่มตลอดที่ไม่มีการแจกแจงแบบปกติ



นางสาวจงจิต มารุ่งศิริกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศึกษาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2548

ISBN 974-53-2698-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

CORRECTION OF NON-NORMALITY FOR RESPONSE OBSERVATION IN
COMPLETELY RANDOMIZED DESIGN

Miss Jongjit Marungsirikul

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

Department of Statistics

Faculty of Commerce and Accountancy

Chulalongkorn University

Academic Year 2005

ISBN 974-53-2698-4

หัวข้อวิทยานิพนธ์

การแก้ไขปัญหาข้อมูลตอบสนองของแผนแบบการทดลองสุ่มตลอดที่ไม่
มีการแจกแจงแบบปกติ

โดย

นางสาวจงจิต มารุ่งสิริกุล

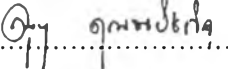
สาขาวิชา

สถิติ

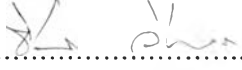
อาจารย์ที่ปรึกษา

รองศาสตราจารย์ ดร.สุพล ดรงค์วัฒนา

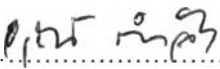
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์
ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

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

คณะกรรมการสอบวิทยานิพนธ์

.......... ประธานกรรมการ
(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)

.......... อาจารย์ที่ปรึกษา
(รองศาสตราจารย์ ดร.สุพล ดรงค์วัฒนา)

.......... กรรมการ
(อาจารย์ ดร.อรุณี กำลัง)

จงจิต มารุ่งสิริกุล : การแก้ไขปัญหาข้อมูลตอบสนองของแผนแบบการทดลองสุ่มตลอดที่ไม่มีมีการแจกแจงแบบปกติ. (CORRECTION OF NON-NORMALITY FOR RESPONSE OBSERVATION IN COMPLETELY RANDOMIZED DESIGN) อ. ที่ปรึกษา : รศ.ดร.สุพล ดุรงค์วัฒนา, 204 หน้า. ISBN 974-53-2698-4.

วัตถุประสงค์ของการวิจัยในครั้งนี้เพื่อศึกษาการแก้ไขปัญหาข้อมูลตอบสนองที่ได้จากการทดลองในแผนแบบการทดลองสุ่มตลอดที่ไม่มีมีการแจกแจงแบบปกติ ด้วยการแปลงข้อมูล เพื่อหารูปแบบการแปลงข้อมูลที่เหมาะสมสำหรับข้อมูลที่มีความคลาดเคลื่อนมีระดับความเบ้ ความโด่งและความแปรปรวนที่แตกต่างกัน ให้มีการแจกแจงแบบปกติและยังเป็นไปตามเงื่อนไขของการวิเคราะห์ความแปรปรวน โดยมีตัวแบบเชิงสถิติดังนี้ $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ เมื่อ $i=1,2,\dots,k$ และ $j=1,2,\dots,n$ โดยที่ Y_{ij} แทนข้อมูลตอบสนองที่ได้จากการทดลองจากวิธีทดลองที่ i หน่วยทดลองที่ j μ แทนค่าเฉลี่ยรวม τ_i แทนผลกระทบจากวิธีทดลองที่ i ε_{ij} แทนความคลาดเคลื่อนสุ่มของการทดลองจากวิธีทดลองที่ i หน่วยทดลองที่ j ที่มีค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ σ^2 ในการวิจัยครั้งนี้การจำลองข้อมูลในแต่ละสถานการณ์ใช้วิธีมอนติคาร์โลด้วยโปรแกรม R โดยกำหนดจำนวนวิธีทดลองเท่ากับ 3 4 และ 5 จำนวนหน่วยทดลองในแต่ละวิธีทดลองเท่ากับ 4 5 และ 6 ความคลาดเคลื่อนของข้อมูลแบ่งออกเป็น 2 กรณีคือเบ้ขวาและเบ้ซ้าย กรณีละ 2 ระดับคือ เบ้น้อยและเบ้มาก ซึ่งในแต่ละระดับความเบ้มี 3 ระดับความโด่งคือ น้อย ปานกลางและมาก และสัมประสิทธิ์ความแปรผันเท่ากับ 20% 40% และ 60% ที่ระดับนัยสำคัญ 0.05 สำหรับเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการแปลงข้อมูลที่เหมาะสมในการแก้ปัญหาดังกล่าวคือ ค่าสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลตอบสนองไม่มีมีการแจกแจงแบบปกติ ค่าสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลตอบสนองไม่มีมีการแจกแจงแบบปกติและมีความแปรปรวนเท่ากัน ค่าสัดส่วนของการปฏิเสธสมมติฐานว่างและอำนาจการทดสอบของการทดสอบเอฟ

ผลการศึกษาจะสรุปได้ดังนี้

1. ที่ระดับความเบ้น้อย จะมีสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลไม่มีมีการแจกแจงแบบปกติและมีความแปรปรวนเท่ากันภายหลังการแปลงข้อมูลมากกว่าที่ระดับความเบ้มาก การแปลงข้อมูลด้วยค่า $\lambda = 0.5$ จะให้ค่าความสำเร็จมากในกรณีเบ้ขวาที่ระดับความเบ้น้อย การแปลงข้อมูลด้วยค่า $\lambda = 0$ จะให้ค่าความสำเร็จมากในกรณีเบ้ขวาที่ระดับความเบ้มาก ยกเว้นกรณีที่มีสัมประสิทธิ์ความแปรผันน้อย การแปลงด้วยค่า $\lambda = 0$ $\lambda = -0.5$ และ $\lambda = -1.0$ จะให้ค่าความสำเร็จมาก ส่วนกรณีเบ้ซ้าย การแปลงด้วยค่า $\lambda = 1.5$ และ $\lambda = 2.0$ จะให้ค่าความสำเร็จสูง
2. ที่ระดับความเบ้เดียวกัน กรณีที่มีค่าความโด่งมาก จะให้ค่าความสำเร็จน้อยกว่ากรณีอื่นและความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 จะน้อย
3. เมื่อสัมประสิทธิ์ความแปรผันมีค่าสูงขึ้น ความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 จะลดลง

ภาควิชา.....สถิติ.....

สาขาวิชา.....สถิติ.....

ปีการศึกษา2548.....

ลายมือชื่อนิสิต.....จงจิต มารุ่งสิริกุล.....

ลายมือชื่ออาจารย์ที่ปรึกษา..........

4682183326 : MAJOR STATISTICS

KEY WORD: BOX-COX TRANSFORMATION / NORMALITY / COMPLETELY RANDOMIZED DESIGN

JONGJIT MARUNGSIRIKUL : CORRECTION OF NON-NORMALITY FOR RESPONSE OBSERVATION IN COMPLETELY RANDOMIZED DESIGN. THESIS ADVISOR : ASSOC. PROF.SUPOL DURONGWATANA,Ph.D., 204 pp. ISBN 974-53-2698-4.

The objective of this study is to correct of non-normality for response observation in completely randomized design. The result of this study present the appropriate function for transformation of data when residuals have different skewness,kurtosis and variance.The statistical model is $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}; i = 1, 2, \dots, k$ and $j = 1, 2, \dots, n$ ε_{ij} is a random error of the j^{th} response data from i^{th} treatment with mean 0 and variance σ^2 . In this study,data was simulated by the Monte Carlo method with R program. Defining numbers of treatment are 3,4 and 5. The numbers of replication in each treatment are 4,5 and 6. Residual was divided in 2 cases,positive skewness and negative skewness.Each of case has 2 levels,less skewness and more skewness. Each of level has 3 levels of kurtosis,low,medium and high. Coefficient of Variation(C.V.) is 20%,40% and 60% at 0.05 significance. The criterion of determination is proportion of success for correction of normality,proportion of success for correction of normality which has homogeneity of variances, proportion of null hypothesis rejection and power of the test. The result of this study can be summarized as follow:

1. When residuals have less skewness,proportion of success for correction of normality which has homogeneity of variances after transform data is more than the residuals have more skewness. For positive skewness case,if residuals are less skewed,transformation of data with $\lambda = 0.5$ will has most successful proportion. If residuals are more skewed,transformation of data with $\lambda = 0$ will has most successful proportion. Except at low C.V.,transformation of data with $\lambda = 0$, $\lambda = -0.5$ and $\lambda = -1.0$ are suitable. For negative skewness,transformation of data with $\lambda = 1.5$ and $\lambda = 2.0$ have more successful.
2. At the same level of skewness,high kurtosis will has successful proportion for correction of normality which has homogeneity of variances after transform data less than another case and ability for control Type I Error is little.
3. When C.V. is increased,the ability for control Type I Error will decrease.

Department.....Statistics.....Student's signature.....*Jongjit Marungsirikul*
 Field of study.....Statistics.....Advisor's signature.....*Supol Durongwatana*
 Academic year.....2005.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วย ความกรุณาและความช่วยเหลือเป็นอย่างดีจากท่าน อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร.สุพล ดุงศ์วัฒนา ที่กรุณาให้คำปรึกษา ตลอดจน ข้อคิดเห็นต่างๆ ในการวิจัยรวมถึงได้ช่วยตรวจผลการวิจัยโดยรวม ผู้เขียนขอขอบพระคุณด้วยความซาบซึ้งและสำนึกในพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณรองศาสตราจารย์ ดร.ธีระพร วีระถาวร และอาจารย์ ดร.อรุณี กำลิ่ง ที่กรุณาเป็นคณะกรรมการสอบวิทยานิพนธ์ฉบับนี้ และในช่วงการศึกษาเพื่อเสนอโครงร่าง วิทยานิพนธ์ได้กรุณาให้คำแนะนำเป็นอย่างดี

ขอขอบพระคุณอาจารย์ ดร.เสกสรร เกียรติสุโขทัย ที่กรุณาให้คำแนะนำในการจำลอง ข้อมูลและความรู้ต่างๆ เพื่อใช้ในการวิจัยครั้งนี้

ขอขอบพระคุณอาจารย์ทุกท่านที่ให้โอกาสทางการศึกษาและประสิทธิ์ประสาทความรู้ ให้แก่ผู้เขียน

ขอขอบคุณเพื่อนๆ หลักสูตรสถิตศาสตร์มหาบัณฑิต รหัส 46 ทุกๆ คนและรุ่นพี่ที่ให้ความช่วยเหลือและคำแนะนำในเรื่องต่างๆ ทั้งทางด้านวิชาการและด้านธุรการ

เนื่องจากวิทยานิพนธ์ฉบับนี้ ผู้เขียนได้รับทุนอุดหนุนการวิจัยจากบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย จึงขอขอบพระคุณไว้ ณ ที่นี้ด้วย นอกจากนี้ผู้เขียนขอขอบคุณเครือข่าย ห้องสมุดของจุฬาลงกรณ์มหาวิทยาลัย ที่ช่วยอำนวยความสะดวกแก่การค้นคว้าอย่างเต็มที่

และวิทยานิพนธ์ฉบับนี้คงไม่สามารถสำเร็จลุล่วงไปด้วยดี ถ้าปราศจากกำลังใจ คำแนะนำ และความช่วยเหลือในทุกๆ ด้าน จาก คุณพ่อ และพี่ชายทุกๆ คน

สุดท้ายนี้ขอรำลึกถึงพระคุณ คุณแม่ ผู้ที่ได้ลุล่วงไปแล้วและจะอยู่ในความทรงจำตลอดไป

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
บทที่	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตของเบื้องต้น.....	2
1.4 ขอบเขตของการวิจัย.....	3
1.5 วิธีดำเนินการวิจัย.....	4
1.6 เกณฑ์ที่ใช้ในการตัดสินใจ.....	5
1.7 คำจำกัดความต่างๆที่ใช้ในการวิจัย.....	5
1.8 ประโยชน์ที่คาดว่าจะได้รับ.....	5
2 แนวคิดและทฤษฎี.....	6
2.1 แผนแบบการทดลองสุ่มตลอด.....	6
2.2 การแจกแจงแลมดาของตุ๊กกี้.....	6
2.3 การทดสอบการแจกแจง.....	10
2.4 วิธีตรวจสอบของเลขวิน.....	10
2.5 การแปลงข้อมูล.....	11
2.6 การวิเคราะห์ความแปรปรวนหรือการทดสอบเอฟสำหรับ แผนแบบการทดลองสุ่มตลอด.....	11
3 วิธีดำเนินการวิจัย.....	13
3.1 การจำลองข้อมูล.....	13
3.2 แผนการดำเนินการวิจัย.....	14
3.3 ขั้นตอนในการดำเนินการวิจัย.....	14

บทที่	หน้า
3.3.1 การสร้างความคลาดเคลื่อนของข้อมูลให้มีความเบ้ ความโด่ง ตามที่กำหนด.....	15
3.3.2 การสร้างข้อมูลตามแผนการทดลองแบบสุ่มตลอด.....	15
3.3.3 การสร้างอิทธิพลของวิธีทดลองให้มีความแตกต่างกัน.....	16
3.3.4 ทดสอบการแจกแจงของความคลาดเคลื่อนของข้อมูล.....	17
3.3.5 ทดสอบความเท่ากันของความแปรปรวนของข้อมูล.....	17
3.3.6 การแก้ไขปัญหาความคลาดเคลื่อนของข้อมูลที่ไม่มีการแจกแจงแบบปกติด้วยวิธีการแปลงข้อมูล.....	17
3.4 แผนผังแสดงขั้นตอนการทำงานของโปรแกรม.....	18
4 ผลการวิจัย.....	23
4.1 การเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลตอบสนองไม่มีการแจกแจงแบบปกติ.....	36
4.2 การเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลไม่มีการแจกแจงแบบปกติและมีความเท่ากันของความแปรปรวนภายหลังการแปลงข้อมูล.....	70
4.3 การเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของการปฏิเสธสมมติฐานว่างของการทดสอบเอฟ.....	96
4.4 การเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากค่าอำนาจการทดสอบของการทดสอบเอฟ.....	126
5 สรุปผลการวิจัยและข้อเสนอแนะ.....	162
5.1 สรุปผลการศึกษา.....	163
5.2 ข้อเสนอแนะ.....	165
รายการอ้างอิง.....	176
ภาคผนวก	177
ภาคผนวก ก.....	178
ภาคผนวก ข.....	181
ภาคผนวก ค.....	190
ภาคผนวก ง.....	195
ประวัติผู้เขียนวิทยานิพนธ์.....	204

สารบัญตาราง

ตาราง	หน้า
2.1 ตารางวิเคราะห์ความแปรปรวนสำหรับแผนแบบทดลองสุ่มตลอดปัจจัยคงที่เมื่อไม่มีหน่วยตัวอย่างย่อยและจำนวนซ้ำของแต่ละวิธีทดลองเท่ากัน.....	11
4.1 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=3$ และ $C.V.=20\%$	27
4.2 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=3$ และ $C.V.=40\%$	28
4.3 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=3$ และ $C.V.=60\%$	29
4.4 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=4$ และ $C.V.=20\%$	30
4.5 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=4$ และ $C.V.=40\%$	31
4.6 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=4$ และ $C.V.=60\%$	32
4.7 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=5$ และ $C.V.=20\%$	33
4.8 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=5$ และ $C.V.=40\%$	34
4.9 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยพิจารณาจากสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลที่ไม่มีการแจกแจงแบบปกติภายหลังการแปลงข้อมูล เมื่อ $k=5$ และ $C.V.=60\%$	35

ตาราง	หน้า
4.71 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยการพิจารณาจากค่าอำนาจการทดสอบของการทดสอบเอฟ เมื่อ $k=5$ และ $C.V.=40%$ (เบ้ซ้าย).....	151
4.72 แสดงการเปรียบเทียบรูปแบบการแปลงข้อมูลโดยการพิจารณาจากค่าอำนาจการทดสอบของการทดสอบเอฟ เมื่อ $k=5$ และ $C.V.=60%$ (เบ้ซ้าย).....	152
5.1 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=3$ และระดับความเบ้ต่ำ (เบ้ขวา).....	167
5.2 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=3$ และระดับความเบ้สูง (เบ้ขวา).....	168
5.3 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=4$ และระดับความเบ้ต่ำ (เบ้ขวา).....	169
5.4 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=4$ และระดับความเบ้สูง (เบ้ขวา).....	170
5.5 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=5$ และระดับความเบ้ต่ำ (เบ้ขวา).....	171
5.6 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=5$ และระดับความเบ้สูง (เบ้ขวา).....	172
5.7 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=3$ (เบ้ซ้าย).....	173
5.8 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=4$ (เบ้ซ้าย).....	174
5.9 แสดงรูปแบบการแปลงข้อมูลเพื่อแก้ปัญหาความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติเมื่อ $k=5$ (เบ้ซ้าย).....	175