การประมาณค่าองค์ประกอบความแปรปรวนสำหรับการทดลองปัจจัยพหุด้วยวิธีบูตสแตรป

นางสาว มนชยา เจียงประดิษฐ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2543

ISBN 974-13-0035-2

BOOTSTRAP ESTIMATION OF VARIANCE COMPONENTS FOR FACTORIAL EXPERIMENT

MISS MONCHAYA CHAINGPRADIT

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 2000
ISBN 974-13-0035-2

หัวข้อวิทยานิพนธ์	การประมาณค่าองค์ประกอบความแปรปรวนสำหรับการทดลองปัจจัยพหุด้วยวิธีบูตสแตรป
โดย	นางสาว มนชยา เจียงประดิษฐ์
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.... คณบดีคณะพาณิชยศาสตร์และการบัญชี
(ผู้ช่วยศาสตราจารย์ ดร. วิรัช อภิเมธีธำรง)

คณะกรรมการสอบวิทยานิพนธ์

.... ประธานกรรมการ
(รองศาสตราจารย์ ศิริพร สาเกทอง)

.... อาจารย์ที่ปรึกษา
(รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา)

.... กรรมการ
(รองศาสตราจารย์ นพรัตน์ รุ่งอุทัยศิริ)

.... กรรมการ
(อาจารย์ ดร. ยงยุทธ ไชยพงศ์)

มนชยา เจียงประดิษฐ์ : การประมาณค่าองค์ประกอบความแปรปรวนสำหรับการทดลองปัจจัยพหุด้วยวิธีบูตสแตรป. (BOOTSTRAP ESTIMATION OF VARIANCE COMPONENTS FOR FACTORIAL EXPERIMENT) อ. ที่ปรึกษา : รศ. ดร.สุพล ดุรงค์วัฒนา, 100 หน้า. ISBN 974-13-0035-2.

วัตถุประสงค์ของการวิจัยครั้งนี้ เพื่อศึกษาและเปรียบเทียบการประมาณค่าองค์ประกอบความแปรปรวนสำหรับการทดลองปัจจัยพหุ กรณีตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม 2 วิธี คือวิธีแบบบูตสแตรปและวิธีแบบคลาสสิก โดยที่ตัวแบบเป็นดังนี้

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \ , i = 1,2,\ldots,a; \ j = 1,2,\ldots b; k = 1,2,\ldots,n$$

Y_{ijk} แทนค่าสังเกตที่ k ระดับที่ i ของปัจจัย A และระดับที่ j ของปัจจัย B, μ แทนค่าเฉลี่ยรวม , α_i แทนอิทธิพลสุ่มระดับที่ i ของปัจจัย A และ α_i มีการแจกแจงที่อิสระจากกันซึ่งมีค่าเฉลี่ย 0 ความแปรปรวน $\sigma_\alpha{}^2$, β_j แทนอิทธิพลสุ่มระดับที่ j ของปัจจัย B และ β_j มีการแจกแจงที่อิสระจากกันซึ่งมีค่าเฉลี่ย 0 ความแปรปรวน $\sigma_\beta{}^2$, $(\alpha\beta)_{ij}$ แทนอิทธิพลสุ่มของผลกระทบร่วมระหว่างระดับที่ i ของปัจจัย A และระดับที่ j ของปัจจัย Bและ $(\alpha\beta)_{ij}$ มีการแจกแจงที่อิสระจากกันซึ่งมีค่าเฉลี่ย 0 ความแปรปรวน $\sigma_{\alpha\beta}{}^2$, ε_{ijk} แทนความคลาดเคลื่อนสุ่มของค่าสังเกตที่ k ระดับที่ i ของปัจจัย A และระดับที่ j ของปัจจัย Bและ ε_{ijk} มีการแจกแจงที่อิสระจากกันซึ่งมีค่าเฉลี่ย 0 ความแปรปรวน $\sigma_\varepsilon{}^2$,a แทนจำนวนระดับของปัจจัย A ,b แทนจำนวนระดับของปัจจัย B,n แทนจำนวนค่าสังเกตในแต่ละวิธีการทดลอง (Treatment Combination),พารามิเตอร์ที่ต้องการประมาณคือค่าองค์ประกอบความแปรปรวน $\sigma_\alpha{}^2, \sigma_\beta{}^2 \sigma_{\alpha\beta}{}^2, \sigma_\varepsilon{}^2$

ข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองด้วยเทคนิคมอนติคาร์โลด้วยโปรแกรม S-PLUS 2000 โดยศึกษาภายใต้การแจกแจงของความคลาดเคลื่อน 2 ลักษณะ คือความคลาดเคลื่อนมีการแจกแจงแบบปกติ ศึกษาภายใต้สถานการณ์ต่าง ๆ ดังนี้ a=b=2,n=3,5,7 a=b=3,n=3,5,7 a=b=4,n=3,5,7 สัมประสิทธิ์การแปรผัน(CV)10%,30% และ50% ความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปนศึกษาภายใต้สถานการณ์ต่าง ๆ ดังนี้ a=b=2,n=3,5,7 a=b=3,n=3,5,7 a=b=4,n=3,5,7 พารามิเตอร์ที่กำหนดความแปรปรวนของความคลาดเคลื่อน σ^2=100 เปอร์เซนต์ของการปลอมปนเป็น 5% 10% และ 25% และสเกลแฟคเตอร์ 2 ระดับ คือ 3 และ 10 เกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการประมาณทั้ง 2 วิธีคือระยะทางยุคลิดเฉลี่ยระหว่างเวคเตอร์ของค่าประมาณองค์ประกอบความแปรปรวนกับค่าจริงของเวกเตอร์ขององค์ประกอบความแปรปรวน

ผลการวิจัยสรุปได้ดังนี้ สำหรับความคลาดเคลื่อนที่มีการแจกแจงแบบปกติ เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้มีค่าต่ำ(a=2,b=2,n=3,5) วิธีบูตสแตรปมีระยะทางยุคลิดเฉลี่ยของเวกเตอร์ขององค์ประกอบความแปรปรวนในการประมาณค่าองค์ประกอบความแปรปรวนต่ำกว่าวิธีคลาสสิก เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้มีขนาดมาก (a=4,b=4,n=5,7) วิธีคลาสสิกมีระยะทางยุคลิดเฉลี่ยของเวกเตอร์ขององค์ประกอบความแปรปรวนต่ำกว่าวิธีบูตสแตรป สำหรับความคลาดเคลื่อนที่มีการแจกแจงแบบปกติปลอมปน โดยส่วนใหญ่ประมาณ98%วิธีบูตสแตรปมีระยะทางยุคลิดเฉลี่ยของเวกเตอร์ขององค์ประกอบความแปรปรวนต่ำกว่าวิธีคลาสสิก

ภาควิชา สถิติ	ลายมือชื่อ
สาขาวิชา สถิติ	ลายมือชื่ออาจารย์ที่ปรึกษา................................
ปีการศึกษา 2543	

4182316026 : MAJOR STATISTICS
KEY WORD: VARIANCE COMPONENTS / BOOTSTRAP ESTIMATION / CLASSIC ESTIMATION / TWO-WAY CROSSED CLASSIFICATION RANDOM-EFFECT

MONCHAYA CHAINGPRADIT : BOOTSTRAP ESTIMATION OF VARIANCE COMPONENTS FOR FACTORIAL EXPERIMENT. THESIS ADVISOR : ASSOCIATE PROFESSOR. SUPOL DURONGWATANA ,Ph.D. 100 pp. ISBN 974-13-0035-2.

The objective of this study is to compare 2 methods of variance components for factorial crossed classification design; the bootstrapping method and the classical method. The model for factorial crossed classification design is as follows:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \ , i = 1,2,...,a; \ j = 1,2,...b; k = 1,2,...,n$$

Y_{ijk} is the kth observation for the ith level of factor A and jth level of factor B; μ is the grand mean; α_i is the ith random effect of factor A and α_i is independently distributed with mean 0 and variance σ_α^2; β_j is the jth random effect of factor B and β_j is independently distributed with mean 0 and variance σ_β^2; $(\alpha\beta)_{ij}$ is the random effect for interaction for the ith level of factor A and the jth level of factor B and $(\alpha\beta)_{ij}$ is also independently distributed with mean 0 and variance σ_β^2; ε_{ijk} is the random error the kth observation for the ith level of factor A and jth level of factor B and ε_{ijk} is also independently distributed with mean 0 and variance σ_ε^2;a is the number of levels for factor A;b is the number of levels for factor B;n is the number of replication for each treatment combination;The parameters; $\sigma_\alpha^{\ 2}, \sigma_\beta^{\ 2} \sigma_{\alpha\beta}^{\ 2}$ and $\sigma_\varepsilon^{\ 2}$, are variance components for the model.

Monte Carlo simulation is done through S-plus 2000 code. It is simulated under situation due to the distribution of random errors. When the distribution of random error is normal distribution ,the simulation is specified at a=b=2,n=3,5 and 7; at a=b=3,n=3,5 and 7;and at a=b=4,n=3,5 and 7 ,the coefficient of variation (CV) is specified at 10%,30% and 50% respectively. When the distribution of random error is contaminated normal distribution, the simulation is specified at a=b=2,n=3,5 and 7; at a=b=3,n=3,5 and 7;and at a=b=4,n=3,5 and 7 ,the variance of random error σ^2 is specified at 100, the percent of contamination is specified at 5%,10% and 25% while the scale factor for contaminated variance of the errors is generated at scale factor 3 and 10. The average of Euclidean distance between the vector of variance component estimates and the vector of true values is a measure for comparison between both methods.

The results of this study show that when the random errors have normal distribution, the number of levels for both factor A and factor B and the number of replication are low(a=2,b=2,and n=3 or 5),the point estimates for variance components using the bootstrapping method, provide shorter averaged distance than the one from the classicical method. When the number of levels for both factor A and factor B, and the number of replication are high(a=4,b=4,and n=5 or 7),the distance from the bootstrapping method is bigger than the one from the classical method. When the distribution of random errors is contaminated normal distribution regardless the percent of contamination and the scale factor , the distance using the bootstrapping method is shorter than the one from classical method in almost of all case (98%).

Department Statistics	Student's signatuere ………………………
Field of study Statistics	Advisor 's signature ………………………..
Academic year 2000	

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของรองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำแนะนำ คำปรึกษา ตลอดจนช่วยเหลือแก้ไขข้อบกพร่องต่างๆจนกระทั่งวิทยานิพนธ์เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณด้วยความรู้สึกซาบซึ้งและสำนึกในพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ศิริพร สาเกทอง รองศาสตราจารย์ นพรัตน์ รุ่งอุทัยศิริ และอาจารย์ ดร.ยงยุทธ ไชยพงศ์ ในฐานะประธานกรรมการและกรรมการสอบวิทยานิพนธ์ ที่กรุณาตรวจแก้วิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น ขอกราบขอบพระคุณคณาจารย์ประจำภาควิชาสถิติ และภาควิชาคณิตศาสตร์ที่ให้โอกาสทางการศึกษา และประสิทธิประสาทความรู้ให้แก่ผู้เขียนจนกระทั่งสำเร็จการศึกษา

ท้ายนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดา มารดา ซึ่งสนับสนุนในด้านการเงิน และให้กำลังใจแก่ผู้วิจัยเสมอมาจนสำเร็จการศึกษา และเนื่องจากทุนการวิจัยครั้งนี้บางส่วนได้รับมาจากทุนอุดหนุนการวิจัยของบัณฑิตวิทยาลัย จึงขอขอบพระคุณบัณฑิตวิทยาลัยมา ณ ที่นี้ด้วย

สารบัญ

สารบัญ (ต่อ)

สารบัญตาราง

สารบัญตาราง (ต่อ)

สารบัญตาราง (ต่อ)

ตารางที่ หน้า

สารบัญตาราง (ต่อ)

ตารางที่ หน้า

สารบัญตาราง (ต่อ)

ตารางที่ หน้า

สารบัญตาราง (ต่อ)

สารบัญภาพ

บทที่1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการศึกษาการวางแผนการทดลอง (Experimental Designs) การทดลองปัจจัยพหุ (Factorial Experiment) เป็นการทดลองที่สามารถนำมาประยุกต์ใช้กับงานในด้านต่าง ๆ เช่นการทดลองด้านการเกษตร การทดลองด้านอุตสาหกรรม และการทดลองในห้องปฏิบัติการทางการแพทย์ เป็นต้น ทั้งนี้เนื่องจากการจัดปัจจัยพหุ (Factorial Arrangement)นั้น ทำให้ศึกษาได้หลาย ๆ ปัจจัยในการทดลองเดียว โดยที่สามารถใช้หน่วยทดลองน้อยกว่าการศึกษาปัจจัยเดียวหลาย ๆ ครั้ง อีกทั้งผู้วิจัยยังสามารถศึกษาปฏิกิริยาสัมพันธ์ระหว่างปัจจัยต่าง ๆ ได้ ซึ่งทำให้ขอบเขตการอนุมานผลสรุปการทดลองเป็นไปได้อย่างกว้างขวาง

สำหรับในการวิจัยครั้งนี้ ผู้วิจัยสนใจศึกษาการประมาณค่าพารามิเตอร์ของการจัดปัจจัยพหุ (Factorial Arrangement) สำหรับตัวแบบข้ามกลุ่มสองปัจจัย (Two-way Crossed Classification) กรณีที่ปัจจัยทั้งสองเป็นปัจจัยสุ่ม โดยที่จัดหน่วยทดลองให้กับแต่ละปัจจัยด้วยวิธีการสุ่ม ซึ่งพารามิเตอร์ที่ต้องการประมาณคือค่าองค์ประกอบความแปรปรวน (Variance Components)

โดยทั่วไปวิธีการประมาณค่าองค์ประกอบความแปรปรวนจะใช้วิธีแบบคลาสสิก (Classical Method) ซึ่งจะเป็นวิธีที่เหมาะสมเมื่อความคลาดเคลื่อนเป็นไปตามข้อตกลงเบื้องต้น กล่าวคือ ความคลาดเคลื่อนมีการแจกแจงแบบปกติ แต่ถ้าแจกแจงของความคลาดเคลื่อนไม่เป็นการแจกแจงแบบปกติ เช่นการแจกแจงที่มีหางยาวหรือมีการกระจายไปทางหางมากกว่าปกติ ซึ่งลักษณะการแจกแจงดังกล่าว จะทำให้ความคลาดเคลื่อนบางค่ามีค่าสูงมาก ๆ หรือต่ำมาก ๆ จากลักษณะดังกล่าวนี้จะทำให้ค่า F ที่ในตารางวิเคราะห์ความแปรปรวน (ANOVA) ผิดไป ซึ่งจะมีผลทำให้ผลสรุปเชื่อถือไม่ได้ ดังนั้นวิธีการนี้อาจไม่เหมาะสมในการประมาณพารามิเตอร์ ซึ่งอาจแก้ปัญหานี้ได้โดยใช้วิธีแบบบูตสแตรป

วิธีแบบบูตสแตรป (Bootstrap Method) เป็นวิธีทางนอนพาราเมตริก (Nonparametric) ซึ่งเป็นวิธีการสุ่มตัวอย่างซ้ำแบบใส่คืน (Resampling with Replacement) โดยใช้ประโยชน์ของการคำนวณด้วยคอมพิวเตอร์ เพื่ออนุมานเชิงสถิติเกี่ยวกับประชากรที่สนใจ ในกรณีที่ไม่ทราบการแจกแจงของความคลาดเคลื่อน ในปี ค.ศ. 1979 แบรดเลย์ เอฟรอน (Bradley Efron) ได้เสนอ

วิธีการนี้เป็นคนแรก ซึ่งวิธีนี้มีแนวความคิดมาจาก วิธีแจกไนฟ (Jackknife) ของเควโนอีล (Queneuille) และตูกี้ (Tukey)

ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยจึงสนใจศึกษาและเปรียบเทียบการประมาณค่าองค์ประกอบความแปรปรวน สำหรับตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม 2 วิธี คือ วิธีแบบคลาสสิกและวิธีแบบบูตสแตรป เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ และแบบปกติปลอมปน

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและเปรียบเทียบการประมาณค่าองค์ประกอบความแปรปรวน สำหรับตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม 2 วิธี คือ

1.2.1 วิธีแบบคลาสสิก

1.2.2 วิธีแบบบูตสแตรป

เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ และแบบปกติปลอมปน

1.3 สมมติฐานของการวิจัย

เมื่อการแจงแจงของความคลาดเคลื่อนเป็นแบบปกติการประมาณค่าองค์ประกอบความแปรปรวนด้วยวิธีแบบคลาสสิกมีค่าใกล้เคียงกับพารามิเตอร์มากกว่าวิธีแบบบูตสแตรป แต่เมื่อการแจกแจงของความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปนวิธีแบบบูตสแตรปมีค่าใกล้เคียงกับพารามิเตอร์มากกว่าวิธีแบบคลาสสิก

1.4 ข้อตกลงเบื้องต้น

1.4.1 พิจารณาตัวแบบ (Model)

<u>ตัวแบบข้ามกลุ่มสองปัจจัยเชิงสุ่ม</u>

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

$i = 1,2,\ldots,a$

โดยที่ $j = 1,2,\ldots b$

$k = 1,2,\ldots,n$

Y_{ijk} แทน ค่าสังเกตที่ k ระดับที่ i ของปัจจัยแรก และระดับที่ j ของปัจจัยสอง

μ แทน ค่าเฉลี่ยรวม

α_i แทน ผลกระทบระดับที่ i ของปัจจัยแรก

β_j แทน ผลกระทบระดับที่ j ของปัจจัยสอง

$(\alpha\beta)_{ij}$ แทน ผลกระทบร่วมระหว่าง ระดับที่ i ของปัจจัยแรก และระดับที่ j ของปัจจัยสอง

ε_{ijk} แทน ความคลาดเคลื่อนของค่าสังเกตที่ k ระดับที่ i ของปัจจัยแรก และระดับที่ j ของปัจจัยสอง

a แทนจำนวนระดับของปัจจัยแรก

b แทนจำนวนระดับของปัจจัยที่ 2

n แทนจำนวนค่าสังเกตในแต่ละวิธีการทดลอง (Treatment Combination)

ดังนั้นจำนวนค่าสังเกตรวม คือ abn

เมื่อ $\alpha_i, \beta_j, (\alpha\beta)_{ij}$ และ ε_{ijk} มีการแจกแจงที่อิสระจากกัน โดยที่

$\alpha_i \overset{iid}{\sim} H$ เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_α^2

$\beta_j \overset{iid}{\sim} H$ เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_β^2

$(\alpha\beta)_{ij} \overset{iid}{\sim} H$ เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ $\sigma_{\alpha\beta}^2$

$\varepsilon_{ijk} \overset{iid}{\sim} H$ เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_ε^2

และจะได้ค่าเฉลี่ยและความแปรปรวนของข้อมูล Y_{ijk} คือ

$$E(Y_{ijk}) = \mu$$

$$V(Y_{ijk}) = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma_\varepsilon^2$$

ซึ่ง พารามิเตอร์ μ เป็นพารามิเตอร์ที่ไม่ทราบค่า และ $\sigma_\alpha^2, \sigma_\beta^2 \sigma_{\alpha\beta}^2, \sigma_\varepsilon^2$ เรียกว่าเป็นพารามิเตอร์องค์ประกอบความแปรปรวน (Variance Components Parameters) ที่ต้องการประมาณ

1.5 ขอบเขตของการวิจัย

1.5.1 ศึกษาวิธีการประมาณค่าองค์ประกอบความแปรปรวนสำหรับตัวแบบข้ามกลุ่มสองปัจจัยเชิงสุ่มด้วยวิธีแบบบูตสแตรปและวิธีแบบคลาสสิก

1.5.2 กำหนดระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ (a,b,n) มีค่าดังตารางข้างล่าง

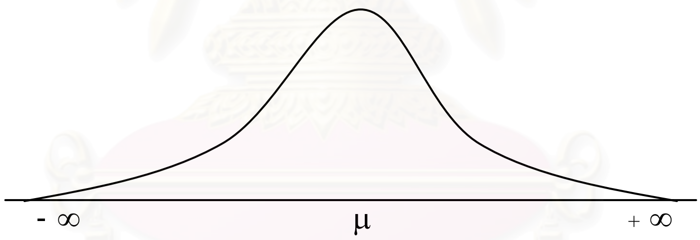
(2,2,3)	(2,2,5)	(2,2,7)
(3,3,3)	(3,3,5)	(3,3,7)
(4,4,3)	(4,4,5)	(4,4,7)

1.5.3 จัดหน่วยทดลองให้กับปัจจัยด้วยวิธีการสุ่มสมบูรณ์ (Completely Randomized Design : CRD)

1.5.4 ลักษณะการแจกแจงของความคลาดเคลื่อนที่ศึกษามีดังนี้

ก. การแจกแจงปกติ (Normal Distribution)

ฟังก์ชันการแจกแจงคือ $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, $-\infty < x < \infty$



รูปที่ 1.1 แสดงเส้นโค้งของการแจกแจงปกติ

ค่าคาดหวัง $E(X) = \mu$

ความแปรปรวน $V(X) = \sigma^2$

ข. การแจกแจงปกติปลอมปน (Scale Contaminated Normal Distribution)

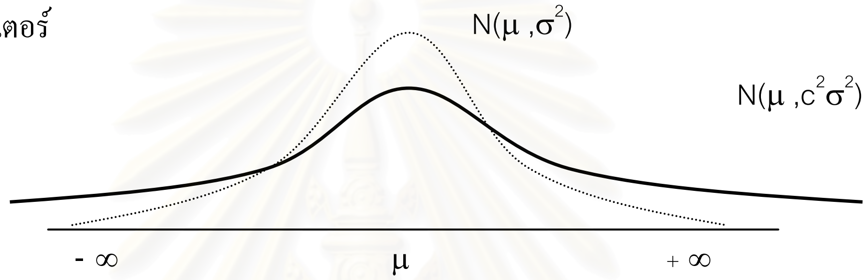
ลักษณะการแจกแจงแบบปกติปลอมปน เป็นการแจกแจงที่แปลงมาจากการแจกแจงแบบปกติ ซึ่งมีฟังก์ชันการแจกแจงคือ

$$f(x) = (1\text{-}p)\, N(0,\sigma^2) + p\, N(0,c^2\sigma^2) \;,\; c > 0 \;,\; 0 \le p \le 1$$

ซึ่งหมายความว่า ค่า X จะมาจากการแจกแจงแบบ $N(0,\sigma^2)$ ด้วยความน่าจะเป็น 1-p และจากการแจกแจงแบบ $N(0,c^2\sigma^2)$ ด้วยความน่าจะเป็น p

เมื่อ ค่าเฉลี่ยเท่ากับ 0 และ σ^2 เป็นพารามิเตอร์ที่กำหนดความแปรปรวนของความคลาดเคลื่อน

p และ c เป็นค่าคงที่ (Fixed Constant) ที่กำหนดเปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์



รูปที่1.2 แสดงเส้นโค้งของการแจกแจงปกติปลอมปนเมื่อเทียบกับการแจกแจงปกติ

ค่าคาดหวัง $E(X) = 0$

ความแปรปรวน $V(X) = (1\text{-}p)\sigma^2 + pc^2\sigma^2$

1.5.5 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ กำหนดให้ข้อมูลมีค่าสัมประสิทธิ์การแปรผัน (Coefficient of Variation : C.V.%) ในระดับต่าง ๆ กันดังนี้ คือ 10% 30% และ50% (พารามิเตอร์ของประชากรมีค่าเฉลี่ย (μ) เท่ากับ 200* และส่วนเบี่ยงเบนมาตรฐาน (σ_y) เท่ากับ 20 60 และ 100)

1.5.6 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน กำหนดให้พารามิเตอร์ที่กำหนดค่าความแปรปรวนของความคลาดเคลื่อน (σ^2) มีค่าเท่ากับ 100 เปอร์เซนต์การปลอมปน เท่ากับ 5% 10% 25% และสเกลแฟคเตอร์เท่ากับ 3 และ 10 ตามลำดับ

---

*สำหรับการวิจัยการกำหนดค่า μ ณ ระดับต่างๆจะไม่ทำให้ผลที่ได้เปลี่ยนแปลงไป ดังนั้นการวิจัยครั้งนี้กำหนดให้ μ เท่ากับ 200

1.5.7 กำหนดค่าความแปรปรวนของความคลาดเคลื่อนของการทดลอง (σ_ε^2) โดยมีหลักการดังนี้

เนื่องจาก $C.V.(y_{ijk}) = \dfrac{SD(Y_{ijk})}{\mu}$

$$= \frac{\sqrt{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma_\varepsilon^2}}{\mu}$$

ในการศึกษาครั้งนี้กำหนดให้ค่าความแปรปรวนของปัจจัยแรก ค่าความแปรปรวนของปัจจัยที่สอง และค่าความแปรปรวนของผลกระทบร่วมมีค่าเป็น k เท่าของความแปรปรวนของความคลาดเคลื่อน ($\sigma_\alpha^2 = \sigma_\beta^2 = \sigma_{\alpha\beta}^2 = k\sigma_\varepsilon^2$)

$$C.V.(y_{ijk}) = \frac{SD(Y_{ijk})}{\mu}$$

$$= \frac{\sqrt{k\sigma_\varepsilon^2 + k\sigma_\varepsilon^2 + k\sigma_\varepsilon^2 + \sigma_\varepsilon^2}}{\mu}$$

$$= \frac{\sigma_\varepsilon\sqrt{3k+1}}{\mu}$$

ดังนั้น $\sigma_\varepsilon^2 = \dfrac{(C.V.(y_{ijk}) \times \mu)^2}{3k+1}$

โดยที่ k เป็นค่าจำนวนเต็มคงที่เท่ากับ 1,3 และ 5

1.5.8 ในการวิจัยครั้งนี้สร้างแบบจำลองข้อมูลโดยใช้เทคนิคการจำลองมอนติคาร์โล (Monte Carlo Simulation) เขียนด้วยโปรแกรม S-PLUS 2000

1.5.9 การจำลองในแต่ละการทดลองกระทำซ้ำ 200 ครั้ง

1.5.10 การสุ่มตัวอย่างแบบใส่คืนในวิธีแบบบูตสแตรปกระทำซ้ำ 100 ครั้ง

1.6 คำจำกัดความที่ใช้ในการวิจัย

1.6.1 หน่วยทดลอง (Experimental Unit) เป็นหน่วยหรือกลุ่มของหน่วยที่ได้รับวิธีการทดลอง (Treatment) อย่างเดียวกัน กล่าวคือหน่วยทดลองจัดเป็นหน่วยที่ใหญ่ที่สุดที่ได้รับวิธีการทดลอง

1.6.2 ปัจจัย(Factor)เป็นตัวแปรอิสระที่กำลังศึกษาว่าสัมพันธ์หรือมีผลกระทบต่อตัวแปรตามหรือไม่

1.6.3 ระดับของปัจจัย(Factor Level) เป็นประเภทต่างๆ ของปัจจัยหรือตัวแปรอิสระที่กำลังศึกษาอยู่

1.6.4 ปัจจัยคงที่ และปัจจัยสุ่ม (Fixed Factor and Random Factor)

ปัจจัยคงที่ คือปัจจัยที่ระดับของปัจจัยนั้น ๆ เป็นจำนวนระดับทั้งหมดที่เป็นไปได้ทั้งหมดที่กำลังทำการศึกษา

ปัจจัยสุ่ม คือปัจจัยที่ระดับของปัจจัยนั้น ๆ เป็นบางส่วนของระดับทั้งหมดที่เป็นไปได้ของปัจจัยนั้น กล่าวคือเป็นตัวอย่างระดับของระดับประชากร

1.6.5 วิธีการทดลอง (Treatment) เป็นการเรียกการจัดระดับต่าง ๆ ของปัจจัย แล้วแต่ว่าปัจจัยที่กำลังศึกษานั้นเป็นปัจจัยเดี่ยว หรือปัจจัยพหุ ในกรณีที่ปัจจัยที่ต้องการศึกษาเป็นปัจจัยเดี่ยว แต่ละระดับของปัจจัยนั้นจัดเป็นแต่ละวิธีการทดลอง ในกรณีที่ปัจจัยที่ต้องการศึกษาเป็นปัจจัยพหุ ส่วนผสมของแต่ละระดับของปัจจัยเป็นวิธีการทดลองหรืออาจเรียกว่า ส่วนผสมของวิธีการทดลอง (Treatment Combinations)

1.6.6 ระยะทางยุคลิด(Euclidean Distance) หมายถึงการเปรียบเทียบขนาดเวกเตอร์ขององค์ประกอบความแปรปรวนแต่ละตัว

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 เพื่อใช้เป็นแนวทางในการประมาณองค์ประกอบความแปรปรวนสำหรับตัวแบบการทดลองปัจจัยพหุ ด้วยวิธีแบบบูตสแตรป

1.7.2 เพื่อใช้เป็นแนวทางในการเลือกใช้วิธีการประมาณองค์ประกอบความแปรปรวนสำหรับตัวแบบการทดลองปัจจัยพหุได้อย่างมีประสิทธิภาพที่สุด ในกรณีที่ความคลาดเคลื่อนมีการแจกแจงเป็นแบบปกติ และการแจกแจงแบบปกติปลอมปน

1.7.3 เพื่อใช้เป็นแนวทางในการศึกษาการประมาณค่าพารามิเตอร์ขององค์ประกอบความแปรปรวนสำหรับแผนการทดลองของตัวแบบอื่น ๆ

1.8 วิธีดำเนินการวิจัย

1.8.1 ศึกษาการประมาณค่าองค์ประกอบความแปรปรวนในตัวแบบสองปัจจัยข้ามกลุ่มเชิงสุ่ม และเขียนโปรแกรมจำลองค่าสังเกตในตัวแบบที่ต้องการศึกษา รวมถึงโปรแกรมสำหรับการคำนวณค่าองค์ประกอบความแปรปรวนของแต่ละวิธี ดังนี้

ก) วิธีแบบคลาสสิก

ข) วิธีแบบบูตสแตรป

1.8.2 ศึกษาเปรียบเทียบวิธีการประมาณค่าองค์ประกอบความแปรปรวนทั้งสองวิธี โดยใช้ระยะทางยุคลิดเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบ

บทที่ 2

แนวคิดและทฤษฎี

ในทางสถิติแนวความคิดเกี่ยวกับการประมาณค่าองค์ประกอบความแปรปรวนสำหรับการทดลองปัจจัยพหุ กรณีตัวแบบข้ามกลุ่มสองปัจจัยเชิงสุ่ม สามารถประมาณได้หลายวิธีโดยวิธีที่เราใช้ทั่วๆไปในทางสถิติคือการประมาณแบบคลาสสิก (Classical Estimation) ซึ่งแนวความคิดแบบคลาสสิกจะมีการกำหนดข้อตกลงเบื้องต้นของข้อมูล (Assumption) แต่ถ้าข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น กล่าวคือเมื่อความคลาดเคลื่อนไม่เป็นการแจกแจงแบบปกติเราจะใช้การประมาณแบบบูตสแตรป (Bootstrap Estimation) ที่เป็นวิธีทางนอนพาราเมตริกโดยการสุ่มตัวอย่างซ้ำแบบใส่คืน ซึ่งเป็นการประมาณค่าองค์ประกอบความแปรปรวนที่นำมาเปรียบเทียบในการศึกษาครั้งนี้ ในขั้นต้นเราจะกล่าวถึงตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม ตารางการวิเคราะห์ความแปรปรวน ลักษณะการแจกแจงของความคลาดเคลื่อนที่ศึกษาและการประมาณค่าองค์ประกอบความแปรปรวนด้วยวิธีแบบคลาสสิกและวิธีแบบบูตสแตรป ได้กล่าวถึงในหัวข้อถัดไป

2.1 ตัวแบบข้ามกลุ่มสองปัจจัยเชิงสุ่ม

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

โดยที่ $i = 1,2,\ldots,a$, $j = 1,2,\ldots b$, $k = 1,2,\ldots,n$

Y_{ijk} แทน ค่าสังเกตที่ k ระดับที่ i ของปัจจัยแรก และระดับที่ j ของปัจจัยสอง

μ แทน ค่าเฉลี่ยรวม

α_i แทน ผลกระทบระดับที่ i ของปัจจัยแรก

β_j แทน ผลกระทบระดับที่ j ของปัจจัยสอง

$(\alpha\beta)_{ij}$ แทน ผลกระทบร่วมระหว่าง ระดับที่ i ของปัจจัยแรก และระดับที่ j ของปัจจัยสอง

ε_{ijk} แทน ความคลาดเคลื่อนของค่าสังเกตที่ k ระดับที่ i ของปัจจัยแรก และระดับที่ j ของปัจจัยสอง

a แทนจำนวนระดับของปัจจัยแรก

b แทนจำนวนระดับของปัจจัยที่สอง

n แทนจำนวนค่าสังเกตในแต่ละวิธีการทดลองผสม(Treatment Combination)

ดังนั้นจำนวนค่าสังเกตรวม คือ abn

เมื่อ $\alpha_i, \beta_j, (\alpha\beta)_{ij}$ และ ε_{ijk} มีการแจกแจงที่อิสระจากกัน โดยที่

$\alpha_i \overset{iid}{\sim}$ H เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_α^2

$\beta_j \overset{iid}{\sim}$ H เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_β^2

$(\alpha\beta)_{ij} \overset{iid}{\sim}$ H เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ $\sigma_{\alpha\beta}^2$

$\varepsilon_{ijk} \overset{iid}{\sim}$ H เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบของการแจกแจง โดยมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_ε^2

ตารางที่ 2.1 ตารางวิเคราะห์ความแปรปรวนสำหรับตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม

สาเหตุ	องศาความเป็นอิสระ	ผลรวมกำลังสอง	ผลรวมกำลังสองเฉลี่ย	ค่าคาดหวังของผลรวมกำลังสองเฉลี่ย
ปัจจัยแรก	$(a-1)$	$SSA = bn(\sum \bar{y}_{i..} - \bar{y}_{...})^2$	$MSA = SSA/(a-1)$	$\sigma_1^2 = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + bn\sigma_\alpha^2$
ปัจจัยที่สอง	$(b-1)$	$SSB = an(\sum \bar{y}_{.j.} - \bar{y}_{...})^2$	$MSB = SSB/(b-1)$	$\sigma_2^2 = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + an\sigma_\beta^2$
ผลกระทบร่วม	$(a-1)(b-1)$	$SSAB = n\sum\sum(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$	$MSAB = SSAB/(a-1)(b-1)$	$\sigma_3^2 = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2$
ความคลาดเคลื่อน	$ab(n-1)$	$SSE = \sum\sum\sum(y_{ijk} - \bar{y}_{ij.})^2$	$MSE = SSE/ab(n-1)$	σ_ε^2
รวม	$abn-1$	$\sum\sum\sum(y_{ijk} - \bar{y}_{...})^2$		

2.2 ลักษณะการแจกแจงของความคลาดเคลื่อนที่ศึกษา

ลักษณะการแจกแจงของความคลาดเคลื่อนที่ศึกษามีดังนี้

ก. การแจกแจงแบบปกติ (Normal Distribution)

ฟังก์ชันการแจกแจงคือ $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, $-\infty < x < \infty$

ค่าคาดหวัง $E(X) = \mu$

ความแปรปรวน $V(X) = \sigma^2$

ข. การแจกแจงแบบปกติปลอมปน (Scale Contaminated Normal Distribution)

ลักษณะการแจกแจงแบบปกติปลอมปน เป็นการแจกแจงที่แปลงมาจากการแจกแจงแบบปกติ ซึ่งมีฟังก์ชันการแจกแจงคือ

$$f(x) = (1-p)\,N(0,\sigma^2) + p\,N(0,c^2\sigma^2)\ ,\ c > 0\ ,\ 0 \le p \le 1$$

ซึ่งหมายความว่า ค่า X จะมาจากการแจกแจงแบบ $N(0,\sigma^2)$ ด้วยความน่าจะเป็น 1-p และจากการแจกแจงแบบ $N(0,c^2\sigma^2)$ ด้วยความน่าจะเป็น p

เมื่อ ค่าเฉลี่ยเท่ากับ 0 และ σ^2 เป็นพารามิเตอร์ที่กำหนดความแปรปรวนของความคลาดเคลื่อน

p และ c เป็นค่าคงที่ (Fixed Constant) ที่กำหนดเปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์

ค่าคาดหวัง $E(X) = 0$

ความแปรปรวน $V(X) = (1-p)\sigma^2 + pc^2\sigma^2$

2.3 การประมาณค่าองค์ประกอบความแปรปรวนแบบคลาสสิก

การประมาณค่าองค์ประกอบความแปรปรวนในตัวแบบของแผนการทดลอง ตามแนวคิดแบบคลาสสิกเป็นค่าที่ไม่ทราบค่า วิธีการประมาณค่าแบบจุดขององค์ประกอบความแปรปรวนนี้ เรียกว่า " วิธีวิเคราะห์ความแปรปรวน" ทั้งนี้เพราะนำค่าต่างๆในตารางวิเคราะห์ความแปรปรวนมาใช้ในการประมาณนั่นเอง วิธีนี้ทำโดยการสร้างสมการที่สมมติว่าค่าเฉลี่ยกำลังสองของสาเหตุของความแปรปรวนเป็นค่าประมาณของค่าคาดหวังของค่าเฉลี่ยกำลังสองที่หาได้ดังตารางที่ 2.1 แล้วใช้หลักการคำนวณทางคณิตศาสตร์เพื่อคำนวณค่าประมาณขององค์ประกอบความแปรปรวน ซึ่งแสดงได้ดังนี้

จาก $$E(MSE) = \sigma_\varepsilon^2$$

ค่าประมาณของ σ_ε^2 คือ MSE

และ
$$E(MSAB) = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2$$
$$E(MSAB) - E(MSE) = (\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2) - \sigma_\varepsilon^2$$
$$E(MSAB) - E(MSE) = n\sigma_{\alpha\beta}^2$$

ค่าประมาณของ $\sigma_{\alpha\beta}^2$ คือ $\dfrac{MSAB - MSE}{n}$

$$E(MSA) = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + bn\sigma_\alpha^2$$
$$E(MSA) - E(MSAB) = (\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + bn\sigma_\alpha^2) - (\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2)$$
$$E(MSA) - E(MSAB) = bn\sigma_\alpha^2$$

ค่าประมาณของ σ_α^2 คือ $\dfrac{MSA - MSAB}{bn}$

$$E(MSB) = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + an\sigma_\beta^2$$
$$E(MSB) - E(MSAB) = (\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + an\sigma_\beta^2) - (\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2)$$
$$E(MSB) - E(MSAB) = an\sigma_\beta^2$$

ค่าประมาณของ σ_β^2 คือ $\dfrac{MSB - MSAB}{an}$

2.4 การประมาณค่าองค์ประกอบความแปรปรวนแบบบูตสแตรป

วิธีการหาตัวประมาณของพารามิเตอร์วิธีนี้เป็นวิธีที่เสนอขึ้นโดย แบรดเลย์ เอฟรอน (Bradley Efron) ในปี ค.ศ. 1979 ซึ่งมีแนวคิดมาจาก วิธีแจกไนฟ (Jackknife) ของเควโนอีล (Queneuille) และตูกี้ (Tukey) ซึ่งวิธีบูตสแตรปนี้เป็นวิธีที่นำมาแก้ปัญหาการไม่สามารถหาค่าประมาณ ในกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น(Asssumptions) เช่นเมื่อความคลาดเคลื่อนมีการแจกแจงที่ไม่เป็นแบบปกติ และการหาค่าประมาณนั้นทำได้ยาก เช่นการประมาณส่วนเบี่ยงเบนมาตรฐานของสัมประสิทธิ์สหสัมพันธ์เป็นต้น โดยใช้ประโยชน์ของการคำนวณด้วยคอมพิวเตอร์

แนวคิดที่สำคัญของวิธีแบบบูตสแตรป กล่าวว่า ตัวอย่างคือสิ่งที่เราทราบทั้งหมดเกี่ยวกับประชากร การสุ่มตัวอย่างจากตัวอย่างที่เรามีอยู่จะเหมือนกับการสุ่มตัวอย่างจากประชากร และตัวอย่างแต่ละตัวอย่างจะสามารถอธิบายลักษณะของประชากรด้วยความน่าจะเป็นที่เท่า ๆ กัน ซึ่งแนวคิดนี้อาจจะทำให้ได้ข้อสรุปที่ดีเกี่ยวกับลักษณะของประชากร โดยที่อาจจะดีมากกว่าการกำหนดข้อตกลงเบื้องต้นเกี่ยวกับประชากร ซึ่งวิธิแบบบูตสแตรปมีหลักเกณฑ์ดังนี้คือ เราจะให้ตัวอย่างที่ถูกเก็บรวบรวมมาจากประชากรเปรียบเสมือนประชากร แล้วทำการสุ่มตัวอย่างจากตัวอย่างที่มีอยู่แบบใส่คืน(Resampling with Replacement) ด้วยจำนวนครั้ง (Bootstrap Replications)ที่มากเพียงพอ เพื่อสร้างการแจกแจงของตัวสถิติตัวอย่าง (Sampling Distribution) และนำไปใช้ประมาณค่าพารามิเตอร์ที่สนใจ

สำหรับตัวอย่างง่าย ๆ ที่อาจจะแสดงเทคนิคของวิธีแบบบูตสแตรปได้อย่างชัดเจนยิ่งขึ้นคือการประมาณค่าเฉลี่ยของประชากร ซึ่งแม้ว่าการประมาณค่าเฉลี่ยของประชากรนี้จะคำนวณได้ง่ายโดยไม่ต้องใช้วิธีแบบบูตสแตรปก็ตาม แต่มันก็จะเป็นตัวอย่างที่ดีที่จะทำให้เราเข้าใจเทคนิคนี้ได้มากขึ้น

ตารางที่ 2.2 ตัวอย่างการสุ่มซ้ำแบบใส่คืนด้วยวิธีบูตสแตรป

ลำดับที่	ตัวอย่างชุดที่มีอยู่เดิม* (Original Sample)	การสุ่มตัวอย่างแบบใส่คืน(Resample)			
		1	2	3	4
1	0.697	-0.270	-1.768	-0.270	-0.152
2	-1.395	0.697	-0.152	-0.152	-1.583
3	1.408	-1.768	-0.270	-1.779	-0.787
4	0.875	0.697	-0.133	2.204	-0.101
5	-2.039	-0.133	-1.395	0.875	-0.914
6	-0.727	0.587	0.587	-0.914	0.697
7	-0.366	-0.016	-1.234	-1.779	-0.727
8	2.204	0.179	-0.152	-2.039	-0.727
9	0.179	0.714	-1.395	2.204	-0.787
10	0.261	0.714	1.099	-0.366	-1.779
11	1.099	-0.097	-1.121	0.875	-0.787
12	-0.787	-2.039	-0.787	-0.457	-1.121
13	-0.097	-1.768	-0.016	-1.121	-1.583
14	-1.779	-0.101	0.739	-0.016	-0.914
15	-0.152	1.099	-1.395	-0.270	-1.234
16	-1.768	-0.727	-1.415	-0.914	-1.395
17	-0.956	-1.121	-0.097	-0.860	2.204
18	0.587	-0.097	-0.101	-0.914	-1.779
19	-0.270	2.204	-1.779	-0.457	-0.366
20	-0.101	0.875	-1.121	-0.697	0.875
21	-1.415	-0.016	-0.101	0.179	2.204
22	-0.860	-0.727	-0.914	-0.366	2.204
23	-1.234	1.408	-2.039	0.875	-0.101
24	-0.457	2.204	-0.366	-1.395	-1.121
25	-0.133	-1.779	2.204	-1.234	2.204
26	-1.583	-1.415	-0.016	-1.121	-0.097
27	-0.914	-0.860	-0.457	1.408	-0.914
28	-1.121	-0.860	2.204	0.261	-0.101
29	0.739	-1.121	-0.133	-1.583	-1.779
30	0.714	-0.101	0.697	-2.039	0.714
ค่าเฉลี่ย	-0.282	-0.121	-0.361	-0.349	-0.325
ส่วนเบี่ยงเบนมาตรฐาน	1.039	1.120	1.062	1.147	1.234

* ประชากรของตัวอย่างชุดที่มีอยู่เดิม(Original Sample) : X ~ N(0,1)

พิจารณาตารางที่ 2.2 เมื่อคอลัมน์ที่ 2 แสดงค่า 30 ค่า ที่สร้างจากประชากรที่มีการแจกแจงปกติมาตรฐานโดยวิธีการสุ่ม กล่าวได้ว่าค่าเหล่านี้เป็นตัวอย่างชุดที่มีอยู่เดิม(Original Sample) คอลัมน์ที่3-6 แสดงตัวอย่างสุ่มแบบใส่คืน(Resample) 4 ชุด จากตัวอย่างชุดที่มีอยู่เดิม แถวรองสุดท้ายและแถวสุดท้ายแสดงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแต่ละคอลัมน์

สังเกตว่าตัวอย่างสุ่มแบบใส่คืนแต่ละชุดจะต่างจากตัวอย่างชุดที่มีอยู่เดิมบ้าง กล่าวคือ ไม่มีค่าของตัวอย่างสุ่มแบบใส่คืนชุดใด ๆ ไม่เคยปรากฏในตัวอย่างชุดที่มีอยู่เดิม แต่ในตัวอย่างสุ่มแบบใส่คืนของตัวอย่างชุดที่มีอยู่เดิมจะมีค่าบางค่าของตัวอย่างชุดที่มีอยู่เดิมที่ปรากฏ มากกว่า 1 ครั้ง และบางค่าจะไม่ปรากฏเลย เช่นในตัวอย่างสุ่มแบบใส่คืนที่1 ค่า x = -1.395 (ลำดับที่ 2 ใน ตัวอย่างชุดที่มีอยู่เดิม)จะไม่ปรากฏเลย ในทางตรงกันข้ามค่าของ x = -0.860 (ลำดับที่ 22 ในตัวอย่างชุดที่มีอยู่เดิม) ปรากฏ 2 ครั้ง ซึ่งแต่ละชุดของตัวอย่างสุ่มแบบใส่คืนจะสร้างโดยสมาชิกในเซตของตัวอย่างชุดที่มีอยู่เดิม จากที่กล่าวมาแล้วทำให้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแต่ละชุดของตัวอย่างสุ่มแบบใส่คืนจะมีค่าต่าง ๆ กันและจะมีค่าต่างจากตัวอย่างชุดที่มีอยู่เดิมถึงแม้ว่าค่าอาจจะใกล้เคียงกัน

ถ้าทำการสุ่มตัวอย่างแบบใส่คืนซึ่งเลือกจากตัวอย่างชุดที่มีอยู่เดิม 1,000 ครั้ง และคำนวณค่าเฉลี่ยของแต่ละตัวอย่างสุ่มแบบใส่คืน การแจกแจงของค่าเฉลี่ยของตัวอย่างสุ่มแบบใส่คืนทั้ง1,000ครั้งนี้เรียกว่าเป็นการประมาณแบบบูตสแตรปของการแจกแจงตัวอย่าง (Sampling distribution) การแจกแจงของค่าเฉลี่ยของตัวแปรนี้ สำหรับขนาดตัวอย่าง n=30 จะได้ว่าการแจกแจงแบบบูตสแตรปนี้มีลักษณะเป็นแบบปกติ

สำหรับการประมาณค่าองค์ประกอบความแปรปรวนของตัวแบบข้ามกลุ่มสองปัจจัยเชิงสุ่มมีตัวแบบดังนี้

ตัวแบบ $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$

$i = 1,2,\ldots,a$

โดยที่ $j = 1,2,\ldots,b$

$k = 1,2,\ldots,n$

เมื่อ $E(\varepsilon_{ijk}) = 0$ และ $Cov(\varepsilon_{ijk}) = \sigma_\varepsilon{}^2$

ความคลาดเคลื่อนเป็นตัวแปรสุ่มแบบต่อเนื่องที่มีการแจกแจงแบบเดียวกันและเป็นอิสระซึ่งกันและกัน (Identically independent distribution)

นั่นคือ $\varepsilon_{ijk} \overset{iid}{\sim} H$ เมื่อ H เป็นการแจกแจงที่ไม่ทราบรูปแบบ

เราจะประมาณค่าพารามิเตอร์ โดยที่พารามิเตอร์คือองค์ประกอบความแปรปรวน (Variance components) นั่นคือ $\sigma_\alpha{}^2, \sigma_\beta{}^2, \sigma_{\alpha\beta}{}^2, \sigma_\varepsilon{}^2$

การประมาณค่าองค์ประกอบความแปรปรวนแบบบูตสแตรปมีขั้นตอนการหาตัวประมาณด้วยวิธีแบบบูตสแตรปดังนี้

1. สุ่มตัวแปรสุ่ม Y^*_{ijk} ต่อไปนี้แบบใส่คืน (with Replacement) ขนาด abn จาก Y_{ijk}

2. จากนั้นหาตัวประมาณองค์ประกอบความแปรปรวน $\hat{\sigma}*_{\alpha}{}^2, \hat{\sigma}*_{\beta}{}^2, \hat{\sigma}*_{\alpha\beta}{}^2, \hat{\sigma}*_{\varepsilon}{}^2$ โดยใช้วิธีแบบบูตสแตรป โดยยึดหลักเดียวกันกับวิธีแบบคลาสสิก ซึ่งจะได้

$$\hat{\sigma}*_{\alpha}{}^2 = \frac{MSA - MSAB}{bn}$$

$$\hat{\sigma}*_{\beta}{}^2 = \frac{MSB - MSAB}{an}$$

$$\hat{\sigma}*_{\alpha\beta}{}^2 = \frac{MSAB - MSE}{n}$$

$$\hat{\sigma}*_{\varepsilon}{}^2 = MSE$$

3.ทำขั้นตอนในข้อ 1-2 ซ้ำกัน B ครั้ง จะได้

$$\hat{\sigma}_{\alpha}^{*1\,2}, \hat{\sigma}_{\beta}^{*1\,2}, \hat{\sigma}_{\alpha\beta}^{*1\,2}, \hat{\sigma}_{\varepsilon}^{*1\,2}, \hat{\sigma}_{\alpha}^{*2\,2}, \hat{\sigma}_{\beta}^{*2\,2}, \hat{\sigma}_{\alpha\beta}^{*2\,2}, \hat{\sigma}_{\varepsilon}^{*2\,2}, \ldots, \hat{\sigma}_{\alpha}^{*B\,2}, \hat{\sigma}_{\beta}^{*B\,2}, \hat{\sigma}_{\alpha\beta}^{*B\,2}, \hat{\sigma}_{\varepsilon}^{*B\,2}$$

4. คำนวณหาค่าตัวประมาณบูตสแตรป

$$\hat{\sigma}^2_{\alpha_{BT}} = \sum_{i=1}^{B} \hat{\sigma}_{\alpha}^{*i\,2} \,/\, B$$

$$\hat{\sigma}^2_{\beta_{BT}} = \sum_{i=1}^{B} \hat{\sigma}_{\beta}^{*i\,2} \,/\, B$$

$$\hat{\sigma}^2_{\alpha\beta_{BT}} = \sum_{i=1}^{B} \hat{\sigma}_{\alpha\beta}^{*i\,2} \,/\, B$$

และ $\hat{\sigma}^2_{\varepsilon_{BT}} = \sum_{i=1}^{B} \hat{\sigma}_{\varepsilon}^{*i\,2} \,/\, B$

2.5 เกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการประมาณทั้ง 2 แบบ

การเปรียบเทียบวิธีการประมาณองค์ประกอบความแปรปรวนสำหรับตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม จะพิจารณาโดยการเปรียบเทียบขนาดเวกเตอร์ ขององค์ประกอบความแปรปรวนระหว่างค่าประมาณขององค์ประกอบความแปรปรวนที่ศึกษากับค่าจริงขององค์ประกอบความแปรปรวนแต่ละตัวซึ่งเรียกเกณฑ์นี้ว่าระยะทางยุคลิดเฉลี่ย “Euclidean Distance” โดยสมมติว่าองค์ประกอบความแปรปรวนแต่ละตัวเป็นอิสระต่อกัน โดยมีหลักการดังนี้

กำหนดให้

$\underset{\sim}{\theta}$ เป็นเวกเตอร์ค่าจริงขององค์ประกอบความแปรปรวน

$\hat{\theta}_{\sim CL}$ เป็นเวกเตอร์ค่าประมาณขององค์ประกอบความแปรปรวนจากวิธีแบบคลาสสิก

$\hat{\theta}_{\sim BT}$ เป็นเวกเตอร์ค่าประมาณขององค์ประกอบความแปรปรวนจากวิธีแบบบูตสแตรป

ซึ่ง $$\underset{\sim}{\theta} = \begin{pmatrix} \sigma^2_\alpha \\ \sigma^2_\beta \\ \sigma^2_{\alpha\beta} \\ \sigma^2_\varepsilon \end{pmatrix}, \hat{\theta}_{\sim CL} = \begin{pmatrix} \hat{\sigma}^2_{\alpha_{CL}} \\ \hat{\sigma}^2_{\beta_{CL}} \\ \hat{\sigma}^2_{\alpha\beta_{CL}} \\ \hat{\sigma}^2_{\varepsilon\, CL} \end{pmatrix}, \hat{\theta}_{\sim BT} = \begin{pmatrix} \hat{\sigma}^2_{\alpha_{BT}} \\ \hat{\sigma}^2_{\beta_{BT}} \\ \hat{\sigma}^2_{\alpha\beta_{BT}} \\ \hat{\sigma}^2_{\varepsilon\, BT} \end{pmatrix}$$

ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก

$$Euc = \frac{\sum_{i=1}^{m} \left\| \hat{\theta}_{\sim CL} - \underset{\sim}{\theta} \right\|}{m} = \frac{\sum_{i=1}^{m} \sqrt{(\hat{\sigma}^2_{\alpha_{CL}} - \sigma^2_\alpha)^2 + (\hat{\sigma}^2_{\beta_{CL}} - \sigma^2_\beta)^2 + (\hat{\sigma}^2_{\alpha\beta_{CL}} - \sigma^2_{\alpha\beta})^2 + (\hat{\sigma}_{\varepsilon_{CL}}^{2} - \sigma^2_\varepsilon)^2}}{m}$$

ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป

$$EuB = \frac{\sum_{i=1}^{m} \left\| \hat{\theta}_{\sim BT} - \underset{\sim}{\theta} \right\|}{m} = \frac{\sum_{i=1}^{m} \sqrt{(\hat{\sigma}^2_{\alpha_{BT}} - \sigma^2_\alpha)^2 + (\hat{\sigma}^2_{\beta_{BT}} - \sigma^2_\beta)^2 + (\hat{\sigma}^2_{\alpha\beta_{BT}} - \sigma^2_{\alpha\beta})^2 + (\hat{\sigma}_{\varepsilon_{BT}}^{2} - \sigma^2_\varepsilon)^2}}{m}$$

โดย m จำนวนรอบที่กระทำซ้ำในแต่ละการทดลอง ในการวิจัยครั้งนี้กำหนดให้ m มีค่าเท่ากับ 200

ดังนั้นถ้าวิธีการใดให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าก็จะเป็นวิธีการประมาณที่เหมาะสม กล่าวคือค่าประมาณขององค์ประกอบความแปรปรวนแบบจุดที่ได้มีค่าใกล้เคียงค่าจริงขององค์ประกอบความแปรปรวนนั้น

บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองซึ่งต้องการศึกษาและเปรียบเทียบการประมาณค่าองค์ประกอบความแปรปรวนสำหรับตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม 2 วิธี คือ วิธีแบบคลาสสิกและวิธีแบบบูตสแตรป โดยสร้างความคลาดเคลื่อนให้มีการแจกแจงแบบปกติ และแบบปกติปลอมปน ทั้งนี้เนื่องจากต้องการทราบผลสรุปในกรณีที่ความคลาดเคลื่อนมีการแจกแจงแบบปกติ และไม่เป็นแบบปกติ เพราะว่าการประมาณพารามิเตอร์ด้วยวิธีแบบบูตสแตรปไม่จำเป็นต้องมีข้อกำหนดเกี่ยวกับการแจกแจงของความคลาดเคลื่อน กล่าวคือ ความคลาดเคลื่อนอาจมีลักษณะการแจกแจงแบบใดก็ได้ที่ไม่ทราบ ซึ่งอาจเป็นไปได้ทั้งแบบปกติและไม่เป็นแบบปกติ การจำลองข้อมูลในสถานการณ์จะใช้เทคนิคมอนติคาร์โล โดยใช้โปรแกรม S–PLUS 2000 กับเครื่อง PC เนื่องจากวิธีมอนติคาร์โลเป็นเทคนิคที่ใช้ในการวิจัยครั้งนี้ ดังนั้นในตอนแรกของบทนี้จะกล่าวถึงวิธีมอนติคาร์โลก่อน แล้วจึงแสดงรายละเอียดของแผนการดำเนินการวิจัย ซึ่งรายละเอียดต่าง ๆ เป็นดังนี้

3.1 การจำลองแบบด้วยเทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โลเป็นการทดลองโดยใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหาที่ยังไม่แน่ใจในผลที่จะเกิดขึ้น เพราะเลขสุ่มมีประโยชน์หลายประการ คือ

1.ทำให้การเลือกตัวอย่างไม่มีความเอนเอียงในการสำรวจหรือทดลองในเรื่องนั้น ๆ ทั้งนี้เพราะเลขสุ่มมาจากแนวคิดเกี่ยวกับการคำนวณความน่าจะเป็น

2.เลขสุ่มจะทำให้ได้มาซึ่งรูปแบบต่างๆ หรือวิธีการที่สลับซับซ้อนโดยการสร้างสถานการณ์จำลอง(Simulation)

3.การใช้เลขสุ่มอาจทำเพื่อศึกษาคุณสมบัติทางทฤษฎีของกระบวนการทางสถิติที่มีความสำคัญสำหรับการประมาณค่า ตลอดจนนำไปสู่คำอธิบายเกี่ยวกับอำนาจการทดสอบทางสถิติ(Power of Statistical Tests)

4.เพื่อหาคำตอบในปัญหาทางคณิตศาสตร์ โดยจะพิจารณาจากการแจกแจงความน่าจะเป็นของปัญหานั้นๆ

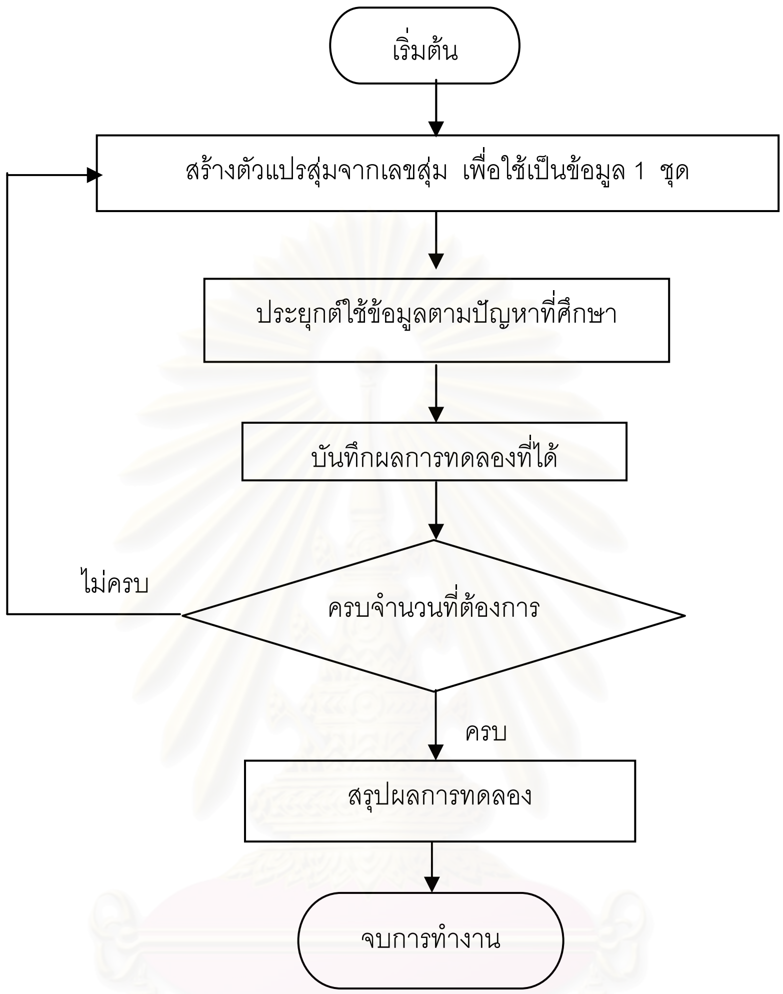
หลักการของเทคนิคมอนติคาร์โล คือการนำเลขสุ่มมาประยุกต์ใช้ในการแก้ไขปัญหาต่างๆ ที่สนใจศึกษาถึงผลสรุปของปัญหานั้นๆ โดยมีขั้นตอนที่สำคัญ 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 **การสร้างเลขสุ่ม** (Generate Random Number)การสร้างเลขสุ่มจะกำหนดให้มีการแจกแจงแบบยูนิฟอร์มในช่วง[0,1] และเป็นอิสระซึ่งกันและกัน จากนั้นนำเลขสุ่มนี้ไปสร้างตัวแปรตามลักษณะการแจกแจงที่ต้องการในปัญหาที่ศึกษา เพื่อเป็นข้อมูลของปัญหานั้นๆ

ขั้นตอนที่ 2 **การประยุกต์ปัญหาที่ต้องการศึกษาโดยใช้เลขสุ่ม** ขั้นตอนนี้ขึ้นอยู่กับปัญหาที่ศึกษา ซึ่งเป็นขั้นตอนที่นำเลขสุ่มมาใช้ในการหาค่าต่าง ๆ ตามปัญหาที่ต้องการตามสูตรการคำนวณในปัญหาที่ศึกษา

ขั้นตอนที่ 3 **การทดลอง** เมื่อประยุกต์ปัญหาที่ต้องการศึกษาโดยใช้เลขสุ่มแล้ว ขั้นต่อไปคือ การทำวิธีการนั้นซ้ำ ๆ กัน (Replication) จำนวนหลายครั้ง โดยถือว่าการทำซ้ำๆ กันนั้นเป็นวิธีการเก็บรวบรวมข้อมูลให้มีจำนวนมากเพื่อลดความไม่แน่นอนของคำตอบ

จากหลักการของเทคนิคมอนติคาร์โล จะเห็นว่าการใช้เลขสุ่มเพื่อเป็นพื้นฐานในการหาคำตอบของปัญหา เป็นวิธีการที่จะนำไปสู่แนวคิดในทางทฤษฎีที่เกี่ยวข้องกับการคำนวณโดยเฉพาะทฤษฎีความน่าจะเป็นที่จะนำไปสู่การอ้างอิงผลสรุปในสถานการณ์ของข้อมูลจริงเพราะไม่มีผลกระทบจากเรื่องอื่น ๆ เข้ามาเกี่ยวข้องในการทดลอง เมื่อทำซ้ำๆ กันเป็นจำนวนมากแล้ว ความคลาดเคลื่อนอย่างสุ่มที่เกิดขึ้นในการวิเคราะห์หาค่าต่าง ๆ ในแต่ละครั้งจะหมดไป (Counter Balance) จากขั้นตอนเทคนิคมอนติคาร์โล สามารถเขียนแผนผังได้ดังนี้



รูปที่ 3.1 **แผนผังเทคนิคมอนติคาร์โล**

3.2 แผนการดำเนินการวิจัย

ในการวิจัยครั้งนี้กำหนดสถานการณ์ต่าง ๆ สำหรับศึกษาและเปรียบเทียบการประมาณค่าองค์ประกอบความแปรปรวนของวิธีแบบคลาสสิก และวิธีแบบบูตสแตรป สำหรับตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม ดังนี้

3.2.1 กำหนดระดับปัจจัยและหน่วยทดลองที่ใช้เป็นดังนี้ a = 2,b = 2, n = 3,5,7 a = 3,b = 3,n = 3,5,7 และ a = 4,b = 4,n = 3,5,7

3.2.2 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ กำหนดให้ข้อมูลมีค่าสัมประสิทธิ์การแปรผัน (Coefficient of Variation : C.V.%) ในระดับต่าง ๆ กันดังนี้ คือ 10% 30% และ50% (พารามิเตอร์ของประชากรมีค่าเฉลี่ย (μ) เท่ากับ 200 และส่วนเบี่ยงเบนมาตรฐาน (σ_y) เท่ากับ 20 60 และ 100)

3.2.3 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน กำหนดให้พารามิเตอร์ที่กำหนดค่าความแปรปรวนของความคลาดเคลื่อน (σ^2) มีค่าเท่ากับ 100 เปอร์เซนต์การปลอมปน เท่ากับ 5% 10% 25% และสเกลแฟคเตอร์เท่ากับ 3 และ 10 ตามลำดับ

3.2.4 ค่าคงที่ k เท่ากับ 1,3 และ 5

3.2.5 กำหนดการกระทำซ้ำในแต่ละสถานการณ์เป็น 200 ครั้ง

3.2.6 สำหรับการประมาณค่าองค์ประกอบความแปรปรวนด้วยวิธีบูตสแตรป กระทำการสุ่มตัวอย่างซ้ำแบบใส่คืน 100 ครั้ง

3.3 ขั้นตอนในการดำเนินการวิจัย

ขั้นตอนในการดำเนินการวิจัย แบ่งออกเป็น 3 ขั้นตอนคือ

3.3.1 การสร้างการแจกแจงของความคลาดเคลื่อนตามที่กำหนด

3.3.2 การสร้างข้อมูลตามตัวแบบดังต่อไปนี้

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

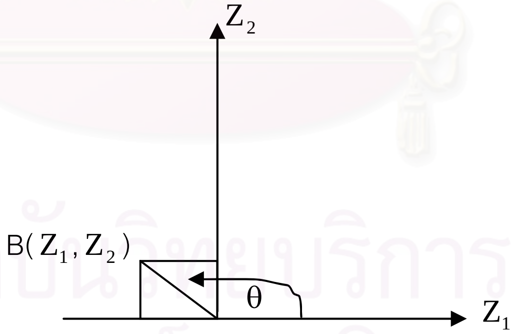
3.3.3 คำนวณหาค่าประมาณองค์ประกอบความแปรปรวนจากทั้งสองวิธี โดยใช้ค่าระยะทางยุคลิดเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบ

รายละเอียดของแต่ละขั้นตอนเป็นดังนี้

3.3.1 การสร้างการแจกแจงของความคลาดเคลื่อนตามที่กำหนด

ก. การสร้างการแจกแจงแบบปกติ

การผลิตตัวเลขสุ่มที่มีการแจกแจงแบบปกติจะใช้วิธีการของ Box Muller(1958) โดยผลิตเลขสุ่มที่มีการแจกแจงปกติมาตรฐาน $N(0,1)$ พร้อมกันทั้งสองค่าและแต่ละค่าเป็นอิสระซึ่งกันและกันโดยใช้ตัวเลข(generator) Z_1 และ Z_2 ซึ่งพิจารณาดังรูปต่อไปนี้



พิจารณาจากรูปจะได้

$$Z_1 = B\cos(\theta) \qquad (3.1)$$

$$Z_2 = B\sin(\theta) \qquad (3.2)$$

เนื่องจาก $B^2 = Z_1^2 + Z_2^2$ มีการแจกแจงไคสแควร์ด้วยระดับขั้นความเสรี 2 และ เทียบเท่ากับการแจกแจงชี้กำลัง (Exponential) ค่าเฉลี่ยเท่ากับ 2 โดยใช้วิธีการแปลงผกผัน (Transformation) ซึ่งสามารถสร้างเลขสุ่มที่มีการแจกแจงชี้กำลัง(Exponential) ได้ดังนี้

$$B = (-2\ln r)^{\frac{1}{2}} \qquad (3.3)$$

เมื่อ r เป็นเลขสุ่มที่มีการแจกแจงสม่ำเสมอ(Uniform) ในช่วง (0,1)

จากการสมมาตรของการแจกแจงปกติ จะได้ว่ามุม θ มีการแจกแจงสม่ำเสมอระหว่าง 0 ถึง 2π เรเดียน และรัศมี B ทำมุมกับ θ เป็นอิสระกัน จาก (3.1) (3.2) และ (3.3) เราสามารถสร้างเลขสุ่มที่มีการแจกแจงปกติมาตรฐาน จากเลขสุ่ม 2 ชุด r_1 และ r_2 กล่าวคือ

$$Z_1 = (-2\ln r_1)^{\frac{1}{2}} \cos(2\pi r_1)$$
$$Z_2 = (-2\ln r_2)^{\frac{1}{2}} \sin(2\pi r_2)$$

ซึ่ง r_1 และ r_2 เป็นตัวเลขสุ่ม เมื่อได้เลขสุ่มที่มีการแจกแจงปกติมาตรฐานแล้ว ทำการแปลงเลขสุ่มดังกล่าวโดยใช้ฟังก์ชัน

$$rn_1 = \mu + \sigma Z_1$$
$$rn_2 = \mu + \sigma Z_2$$

ซึ่งจะได้ว่า rn_1 และ rn_2 มีการแจกแจงปกติ ด้วยค่าเฉลี่ยเท่ากับ μ และมีความแปรปรวนเท่ากับ σ^2

สำหรับโปรแกรม S-plus 2000 จะใช้ฟังก์ชัน rnorm(n, μ, sd) ในการสร้างการแจกแจงแบบปกติ โดย n แทน ขนาดตัวอย่าง μ แทนค่าเฉลี่ย และ sd แทน ค่าเบี่ยงเบนมาตรฐาน ซึ่งผู้วิจัยเสนอไว้ในภาคผนวก

ข. การสร้างการแจกแจงแบบปกติปลอมปน

การผลิตเลขสุ่มที่มีการแจกแจงแบบปกติปลอมปนที่มีค่าเฉลี่ยและความแปรปรวนตามที่กำหนดผู้วิจัยใช้วิธีที่ แรมเซย์ (Ramsay) เสนอไว้โดยการสร้างการแจกแจงที่แปลงมาจากการแจกแจงแบบปกติ ที่มีฟังก์ชันการแจกแจงอยู่ในรูปของ

$$f(x) = (1-p)\, N(0,\sigma^2) + p\, N(0,c^2\sigma^2) \;,\; c > 0 \;,\; 0 \le p \le 1$$

ซึ่งหมายความว่า ค่า X จะมาจากการแจกแจงแบบ $N(0,\sigma^2)$ ด้วยความน่าจะเป็น 1-p และจากการแจกแจงแบบ $N(0,c^2\sigma^2)$ ด้วยความน่าจะเป็น p

เมื่อ ค่าเฉลี่ยเท่ากับ 0 และ σ^2 เป็นพารามิเตอร์ที่กำหนดความแปรปรวนของความคลาดเคลื่อน

p และ c เป็นค่าคงที่ (Fixed Constant) ที่กำหนดเปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์

ค่าคาดหวัง $E(X) = 0$

ความแปรปรวน $V(X) = (1-p)\sigma^2 + p\,c^2\sigma^2$

สำหรับคำสั่งในการสร้างการแจกแจงแบบปกติปลอมปน ผู้วิจัยเสนอไว้ในภาคผนวก

3.3.2 การสร้างข้อมูลตามตัวแบบ

ก. การสร้างตัวแปรสุ่ม α_i

$$\alpha_i \overset{iid}{\sim} N(0, \sigma_\alpha^2)$$

ข. การสร้างตัวแปรสุ่ม β_j

$$\beta_j \overset{iid}{\sim} N(0, \sigma_\beta^2)$$

ค. การสร้างตัวแปรสุ่ม $(\alpha\beta)_{ij}$

$$(\alpha\beta)_{ij} \overset{iid}{\sim} N(0, \sigma_{\alpha\beta}^2)$$

ง. การสร้างตัวแปรสุ่ม ε_{ijk}

$\varepsilon_{ijk} \overset{iid}{\sim} H$ เมื่อ H มีการแจกแจงตามที่กำหนด

3.3.3 การคำนวณค่าองค์ประกอบความแปรปรวน

ก. วิธีแบบคลาสสิก

ค่าประมาณของ σ_α^2 คือ $\hat{\sigma}_\alpha^2 = \dfrac{MSA - MSAB}{bn}$

ค่าประมาณของ σ_β^2 คือ $\hat{\sigma}_\beta^2 = \dfrac{MSB - MSAB}{an}$

ค่าประมาณของ $\sigma_{\alpha\beta}^2$ คือ $\hat{\sigma}_{\alpha\beta}^2 = \dfrac{MSAB - MSE}{n}$

และ ค่าประมาณของ σ_ε^2 คือ $\hat{\sigma}_\varepsilon^2 = MSE$

ข. วิธีแบบบูตสแตรป

ค่าประมาณของ σ_α^2 คือ $\hat{\sigma}_{\alpha_{BT}}^2 = \sum_{i=1}^{B} \hat{\sigma}_\alpha^{*i\,2} / B$

ค่าประมาณของ σ_β^2 คือ $\hat{\sigma}_{\alpha_{BT}}^2 = \sum_{i=1}^{B} \hat{\sigma}_\alpha^{*i\,2} / B$

ค่าประมาณของ $\sigma_{\alpha\beta}^2$ คือ $\hat{\sigma}_{\alpha\beta_{BT}}^2 = \sum_{i=1}^{B} \hat{\sigma}_{\alpha\beta}^{*i\,2} / B$

และ ค่าประมาณของ σ_ε^2 คือ $\hat{\sigma}_{\varepsilon_{BT}}^2 = \sum_{i=1}^{B} \hat{\sigma}_\varepsilon^{*i\,2} / B$

เมื่อ B คือจำนวนครั้งในการสุ่มตัวอย่างแบบใส่คืน สำหรับการวิจัยครั้งนี้ผู้วิจัยกำหนดให้ B=100

ค. การคำนวณหาระยะทางยุคลิดเฉลี่ย

วิธีแบบคลาสสิก

$$Educ = \frac{\sum_{i=1}^{m}\left\| \underset{\sim}{\hat{\theta}}_{CL} - \underset{\sim}{\theta} \right\|}{m} = \frac{\sum_{i=1}^{m}\sqrt{(\hat{\sigma}^2_{\alpha_{CL}} - \sigma^2_{\alpha})^2 + (\hat{\sigma}^2_{\beta_{CL}} - \sigma^2_{\beta})^2 + (\hat{\sigma}^2_{\alpha\beta_{CL}} - \sigma^2_{\alpha\beta})^2 + (\hat{\sigma}_{\varepsilon_{CL}}^{2} - \sigma^2_{\varepsilon})^2}}{m}$$

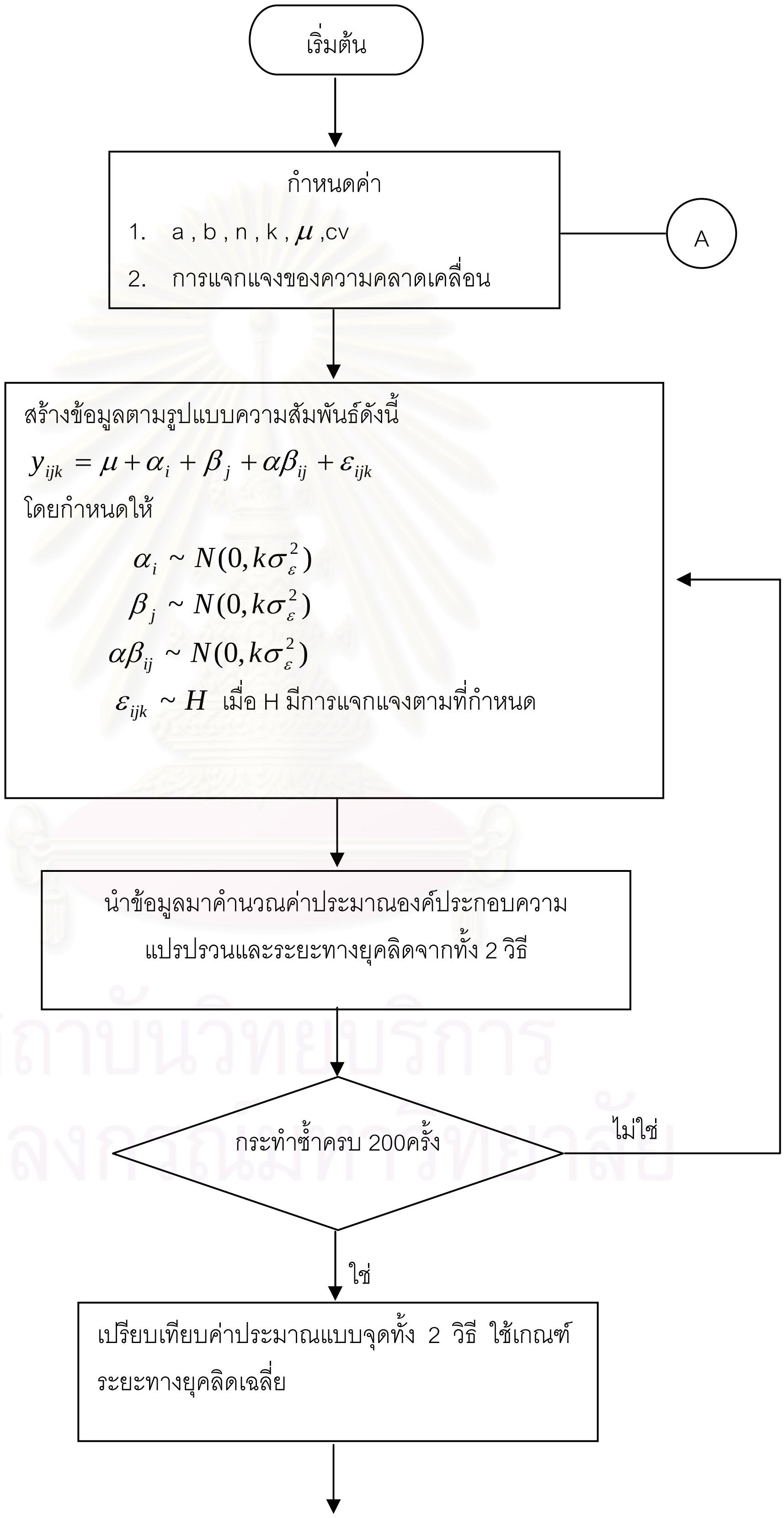
วิธีแบบบูตสแตรป

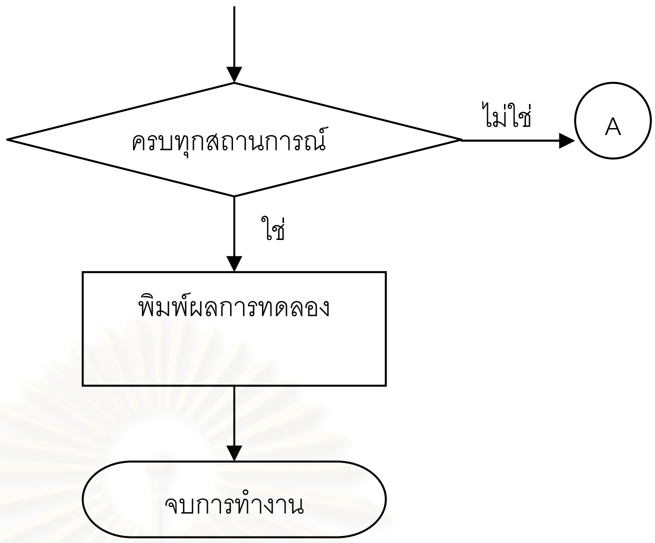
$$Educ = \frac{\sum_{i=1}^{m}\left\| \underset{\sim}{\hat{\theta}}_{BT} - \underset{\sim}{\theta} \right\|}{m} = \frac{\sum_{i=1}^{m}\sqrt{(\hat{\sigma}^2_{\alpha_{BT}} - \sigma^2_{\alpha})^2 + (\hat{\sigma}^2_{\beta_{BT}} - \sigma^2_{\beta})^2 + (\hat{\sigma}^2_{\alpha\beta_{BT}} - \sigma^2_{\alpha\beta})^2 + (\hat{\sigma}_{\varepsilon_{BT}}^{2} - \sigma^2_{\varepsilon})^2}}{m}$$

เมื่อ m คือจำนวนการกระทำซ้ำในแต่ละสถานการณ์ สำหรับการวิจัยครั้งนี้ผู้วิจัยกำหนดให้ m=200

ซึ่งถ้าวิธีการประมาณแบบใดที่ให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าก็จะเป็นวิธีการประมาณที่เหมาะสม

3.4 ขั้นตอนในการทำงานของโปรแกรม




ครบทุกสถานการณ์
ไม่ใช่
A
ใช่
พิมพ์ผลการทดลอง
จบการทำงาน

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบวิธีการประมาณค่าองค์ประกอบความแปรปรวน สำหรับการทดลองปัจจัยพหุ กรณีตัวแบบการทดลองข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม 2 วิธี คือ 1) วิธีแบบคลาสสิก (Classical Estimation) 2) วิธีแบบบูตสแตรป (Bootstrap Estimation) โดยศึกษาภายใต้การแจกแจงของความคลาดเคลื่อนเป็นแบบปกติ และแบบปกติปลอมปนในสถานการณ์ต่าง ๆ กล่าวคือทำการศึกษาในสถานการณ์ที่ปัจจัยทดลอง A มีอยู่ a ระดับ ปัจจัยทดลอง B มีอยู่ b ระดับ และขนาดหน่วยทดลองที่ใช้ n คือ a=2,b=2,n=3,5,7 , a=3,b=3,n=3,5,7 , a=4,b=4,n=3,5,7 และเมื่อ $\sigma_\alpha^2 = \sigma_\beta^2 = \sigma_{\alpha\beta}^2 = k\sigma_\varepsilon^2$ เมื่อ k เป็นค่าจำนวนเต็มคงที่ซึ่งกำหนดให้ k =1,3,5 โดยที่เมื่อความคลาดเคลื่อนมีการแจกแจงเป็นแบบปกติ ผู้วิจัยทำการจำลองข้อมูล ให้มีสัมประสิทธิ์การแปรผัน (CV%)ต่างๆ กัน คือ CV =10%,30%, 50%และเมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน จะทำการจำลองข้อมูล ให้มีพารามิเตอร์ที่กำหนดความแปรปรวนของความคลาดเคลื่อน σ^2 =100 เปอร์เซนต์การปลอมปน p=5%,10%,25% และสเกลแฟคเตอร์ 3,10 โดยวิธีการจำลองข้อมูลนั้นจะอาศัยเทคนิคมอนติคาร์โลซีมูเลชัน (Monte Carlo Simulation) ซึ่งจะกระทำซ้ำทั้งหมด 200 รอบ และสำหรับการคำนวณในวิธีแบบบูตสแตรป จะทำการสุ่มตัวอย่างซ้ำแบบใส่คืนทั้งหมด100 รอบ

ในการศึกษาวิจัยครั้งนี้เกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการประมาณทั้ง 2 แบบ เพื่อหาวิธีการประมาณที่เหมาะสมที่สุด คือระยะทางยุคลิดเฉลี่ย กล่าวคือ ถ้าวิธีการประมาณแบบใดให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าก็จะเป็นวิธีการประมาณที่เหมาะสม ซึ่งแสดงได้ว่าค่าประมาณขององค์ประกอบความแปรปรวนที่ได้มีค่าใกล้เคียงค่าจริงขององค์ประกอบความแปรปรวนนั้น

สำหรับการนำเสนอผลการวิจัยจะนำเสนออยู่ในรูปของตารางซึ่งแบ่งเป็น 2 ตอน คือ

4.1 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ

4.2 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน

และเพื่อสะดวกในการนำเสนอผลการวิจัย ผู้วิจัยขอใช้สัญลักษณ์ต่อไปนี้แทนต่อไปนี้แทนความหมายต่าง ๆ ดังนี้

a แทน ระดับของปัจจัยทดลอง A

b แทน ระดับของปัจจัยทดลอง B

n แทน ขนาดตัวอย่างที่ใช้ในแต่ละปัจจัยทดลอง

CV แทนสัมประสิทธ์การแปรผัน(%)

p แทนเปอร์เซนต์การปลอมปน

c แทนสเกลแฟคเตอร์

4.1 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ

การนำเสนอผลการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิก และวิธีแบบบูตสแตรป ผู้วิจัยจะนำเสนอผลที่ได้ในกรณีที่ระดับปัจจัยและขนาดตัวอย่างที่ใช้ ค่าสัมประสิทธิ์การแปรผัน ค่าคงที่ k มีค่าต่างๆ กันดังตาราง ที่ 4.1 – 4.18

จากตาราง ที่ 4.1-4.9 พบว่า

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ CV=30%,k=5 และ CV=50%,k=5

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ CV=30%,k=5 และ CV=50%,k=5

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=7 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ CV=50%,k=3

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ CV=10%,k=3 และ CV=30%,k=3

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ CV=10%,k=3 CV=30%,k=1,3 CV=50%,k=1

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=7 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ CV=10%,k=3 CV=50%,k=5

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ CV=10%,k=5 CV=50%,k=3

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ CV=30%,k=3

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=7 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ CV=10%,k=3,5

จากตารางที่4.10-4.18 พบว่า

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=10% ค่าคงที่ k=1 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ a=2,b=2,n=3,5,7 และ a=3,b=3,n=7

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=30% ค่าคงที่ k=1 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=3,b=3,n=3 และ a=4,b=4,n=3,5,7

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=50% ค่าคงที่ k=1 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=3,b=3,n=3 และ a=4,b=4,n=3,5,7

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=10% ค่าคงที่ k=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=3,b=3,n=7 และ a=4,b=4,n=3,5

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=30% ค่าคงที่ k=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=4,b=4,n=3,7

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=50% ค่าคงที่ k=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ a=2,b=2,n=3,5 a=3,b=3,n=7 และ a=4,b=4,n=3

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=10% ค่าคงที่ k=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=3,b=3,n=3,5 และ a=4,b=4,n=5

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=30% ค่าคงที่ k=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ a=2,b=2,n=7 และ a=3,b=3,n=7

เมื่อค่าสัมประสิทธิ์การแปรผัน CV=50% ค่าคงที่ k=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปทุกกรณี ยกเว้นในกรณีที่ a=2,b=2,n=7

จากตารางทั้งหมด สรุปได้ว่าเมื่อระดับปัจจัยทดลองและขนาดหน่วยทดลองที่ใช้มีจำนวนน้อย วิธีแบบบูตสแตรปจะให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าวิธีแบบคลาสสิก แต่เมื่อระดับปัจจัยทดลองและขนาดหน่วยทดลองที่ใช้มีจำนวนเพิ่มขึ้น วิธีแบบคลาสสิกจะให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าวิธีแบบบูตสแตรป

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ ค่าสัมประสิทธิ์การแปรผันมีค่าหนึ่งๆ และค่าคงที่ k เพิ่มขึ้น ระยะทางยุคลิคเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มเพิ่มขึ้น

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ ค่าคงที่ k มีค่าหนึ่งๆ และค่าสัมประสิทธิ์การแปรผันเพิ่มขึ้น ระยะทางยุคลิคเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มเพิ่มขึ้น

เมื่อค่าสัมประสิทธิ์การแปรผัน ค่าคงที่ k และระดับปัจจัยทดลองมีค่าหนึ่ง ๆ แต่ขนาดหน่วยทดลองที่ใช้มีค่าเพิ่มขึ้น ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มลดลง

เมื่อค่าสัมประสิทธิ์การแปรผัน ค่าคงที่ k และขนาดหน่วยทดลองที่ใช้มีค่าหนึ่ง ๆ แต่ระดับปัจจัยทดลองมีค่าเพิ่มขึ้น ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มลดลง

ตารางที่ 4.1 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=2,b=2,n=3 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	10	1	299.7293	297.3827**
		3	306.6197	303.9689**
		5	348.1708	345.5251**
	30	1	2748.1201	2716.9155**
		3	2896.6418	2868.8082**
		5	3283.2525**	3285.7134
	50	1	7410.6939	7299.6408**
		3	8034.9188	7998.7784**
		5	8178.1122**	8184.4014

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.2 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=2,b=2,n=5 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=5	10	1	283.4166	277.6445**
		3	288.0783	287.0783**
		5	333.8287	332.3155**
	30	1	2348.5341	2319.0981**
		3	2727.1130	2718.5522**
		5	2820.3786**	2821.9057
	50	1	6556.1771	6491.7783**
		3	7395.1909	7346.0183**
		5	7677.5384**	7685.4930

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.3 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=2,b=2,n=7 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=7	10	1	246.4767	240.2480**
		3	276.9037	273.9888**
		5	312.0121	310.1863**
	30	1	2322.5327	2284.6272**
		3	2660.5318	2644.3254**
		5	2692.0189	2678.5496**
	50	1	6439.4654	6392.2133**
		3	7207.0092**	7210.8280
		5	7584.3065	7531.3909**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.4 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=3,b=3,n=3 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=3,b=3,n=3	10	1	192.7570**	193.8016
		3	208.0720	207.0945**
		5	209.7725**	215.7266
	30	1	1689.6481**	1699.4401
		3	1826.3211	1822.4283**
		5	1872.4986**	1880.3120
	50	1	4574.4396**	4588.7139
		3	5451.9921**	5471.9008
		5	5610.9693**	5612.7052

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.5 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=3,b=3,n=5 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=3,b=3,n=5	10	1	183.7465**	183.9473
		3	190.5695	189.5656**
		5	206.6062**	206.7172
	30	1	1654.6287	1649.3679**
		3	1725.0189	1718.7067**
		5	1843.4642**	1851.9395
	50	1	4255.5942	4238.5317**
		3	4880.5744**	4884.8010
		5	5354.6324**	5357.0379

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.6 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=3,b=3,n=7 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=3,b=3,n=7	10	1	173.2871	172.0707**
		3	177.6092**	177.9124
		5	200.6640	200.1712**
	30	1	1432.9327	1426.8091**
		3	1669.8505	1657.4261**
		5	1817.8967	1809.3335**
	50	1	4170.6612	4130.1624**
		3	4873.2637	4846.9086**
		5	5121.2209**	5133.9977

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.7 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=4,b=4,n=3 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=4,b=4,n=3	10	1	146.9801**	152.0108
		3	165.9333**	166.1642
		5	174.6679	174.4146**
	30	1	1327.2909**	1381.8712
		3	1453.7576**	1488.7573
		5	1643.4022**	1644.0005
	50	1	3581.9736**	3695.9091
		3	3968.7921	3959.3701**
		5	4104.7171**	4115.6938

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.8 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=4,b=4,n=5 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=4,b=4,n=5	10	1	135.9941**	136.5764
		3	153.5784**	154.0743
		5	159.7572**	160.4143
	30	1	1257.2435**	1277.1924
		3	1360.1097	1358.0151**
		5	1469.6787**	1470.0782
	50	1	3497.2256**	3509.2695
		3	3910.7076**	3921.0612
		5	4075.2399**	4081.9412

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.9 แสดงการเปรียบเทียบระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับของปัจจัยและขนาดหน่วยทดลองที่ใช้ คือ a=4,b=4,n=7 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	สัมประสิทธิ์การแปรผัน(CV%)	ค่าk	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=4,b=4,n=7	10	1	134.1448**	134.9325
		3	147.9107	147.8185**
		5	157.7051	157.3778**
	30	1	1256.3296**	1259.7520
		3	1358.1989**	1360.0151
		5	1435.7960**	1438.7064
	50	1	3443.6066**	3452.9387
		3	3654.8596**	3657.1504
		5	3921.9599**	3924.1166

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.10 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 10% ค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	299.7293	297.3827**
a=2,b=2,n=5	283.4166	277.6445**
a=2,b=2,n=7	246.4767	240.2480**
a=3,b=3,n=3	192.7570**	193.8016
a=3,b=3,n=5	183.7465**	183.9473
a=3,b=3,n=7	173.2871	172.0707**
a=4,b=4,n=3	146.9801**	152.0108
a=4,b=4,n=5	135.9941**	136.5764
a=4,b=4,n=7	134.1448**	134.9325

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.11 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 30% ค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	2748.1201	2716.9155**
a=2,b=2,n=5	2348.5341	2319.0981**
a=2,b=2,n=7	2322.5327	2284.6272**
a=3,b=3,n=3	1689.6481**	1699.4401
a=3,b=3,n=5	1654.6287	1649.3679**
a=3,b=3,n=7	1432.9327	1426.8091**
a=4,b=4,n=3	1327.2909**	1381.8712
a=4,b=4,n=5	1257.2435**	1277.1924
a=4,b=4,n=7	1256.3296**	1259.7520

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.12 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 50% ค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	7410.6939	7299.6408**
a=2,b=2,n=5	6556.1771	6491.7783**
a=2,b=2,n=7	6439.4654	6392.2133**
a=3,b=3,n=3	4574.4396**	4588.7139
a=3,b=3,n=5	4255.5942	4238.5317**
a=3,b=3,n=7	4170.6612	4130.1624**
a=4,b=4,n=3	3581.9736**	3695.9091
a=4,b=4,n=5	3497.2256**	3509.2695
a=4,b=4,n=7	3443.6066**	3452.9387

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.13 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 10% ค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	306.6197	303.9689**
a=2,b=2,n=5	288.0783	287.0783**
a=2,b=2,n=7	276.9037	273.9888**
a=3,b=3,n=3	208.0720	207.0945**
a=3,b=3,n=5	190.5695	189.5656**
a=3,b=3,n=7	177.6092**	177.9124
a=4,b=4,n=3	165.9333**	166.1642
a=4,b=4,n=5	153.5784**	154.0743
a=4,b=4,n=7	147.9107	147.8185**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.14 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 30% ค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	2896.6418	2868.8082**
a=2,b=2,n=5	2727.1130	2718.5522**
a=2,b=2,n=7	2660.5318	2644.3254**
a=3,b=3,n=3	1826.3211	1822.4283**
a=3,b=3,n=5	1725.0189	1718.7067**
a=3,b=3,n=7	1669.8505	1657.4261**
a=4,b=4,n=3	1453.7576**	1488.7573
a=4,b=4,n=5	1360.1097	1358.0151**
a=4,b=4,n=7	1358.1989**	1360.0151

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.15 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 50% ค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	8034.9188	7998.7784**
a=2,b=2,n=5	7395.1909	7346.0183**
a=2,b=2,n=7	7207.0092**	7210.8280
a=3,b=3,n=3	5451.9921**	5471.9008
a=3,b=3,n=5	4880.5744**	4884.8010
a=3,b=3,n=7	4873.2637	4846.9086**
a=4,b=4,n=3	3968.7921	3959.3701**
a=4,b=4,n=5	3910.7076**	3921.0612
a=4,b=4,n=7	3654.8596**	3657.1504

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.16 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 10% ค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	348.1708	345.5251**
a=2,b=2,n=5	333.8287	332.3155**
a=2,b=2,n=7	312.0121	310.1863**
a=3,b=3,n=3	209.7725**	215.7266
a=3,b=3,n=5	206.6062**	206.7172
a=3,b=3,n=7	200.6640	200.1712**
a=4,b=4,n=3	174.6679	174.4146**
a=4,b=4,n=5	159.7572**	160.4143
a=4,b=4,n=7	157.7051	157.3778**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.17 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 30% ค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	3283.2525**	3285.7134
a=2,b=2,n=5	2820.3786**	2821.9057
a=2,b=2,n=7	2692.0189	2678.5496**
a=3,b=3,n=3	1872.4986**	1880.3120
a=3,b=3,n=5	1843.4642**	1851.9395
a=3,b=3,n=7	1817.8967	1809.3335**
a=4,b=4,n=3	1643.4022**	1644.0005
a=4,b=4,n=5	1469.6787**	1470.0782
a=4,b=4,n=7	1435.7960**	1438.7064

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.18 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าสัมประสิทธิ์การแปรผัน(CV%)เท่ากับ 50% ค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติ

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	8178.1122**	8184.4014
a=2,b=2,n=5	7677.5384**	7685.4930
a=2,b=2,n=7	7584.3065	7531.3909**
a=3,b=3,n=3	5610.9693**	5612.7052
a=3,b=3,n=5	5354.6324**	5357.0379
a=3,b=3,n=7	5121.2209**	5133.9977
a=4,b=4,n=3	4104.7171**	4115.6938
a=4,b=4,n=5	4075.2399**	4081.9412
a=4,b=4,n=7	3921.9599**	3924.1166

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

4.2 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน

การนำเสนอผลการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิก และวิธีแบบบูตสแตรป ผู้วิจัยจะนำเสนอผลที่ได้ในกรณีที่ระดับปัจจัยและขนาดตัวอย่างที่ใช้ ค่าคงที่ k เปอร์เซนต์การปลอมปน และสเกลแฟคเตอร์ มีค่าต่างๆ กันดังตาราง ที่ 4.19 – 4.45

จากตารางที่ 4.19-4.45 พบว่า

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=3,5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=7 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ k=1,p=5%,c=3

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ k=5,p=5%,c=3

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=7 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ k=1,p=5%,c=3

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=5 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=7 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ k=3,p=5%,c=3

จากตารางที่ 4.28-4.54 พบว่า

เมื่อค่าคงที่ k=1 เปอร์เซนต์การปลอมปน p=5% และสเกลแฟคเตอร์ c=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=2,b=2,n=7 และ a=4,b=4,n=3

เมื่อค่าคงที่ k=3 เปอร์เซนต์การปลอมปน p=5% และสเกลแฟคเตอร์ c=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=4,b=4,n=7

เมื่อค่าคงที่ k=5 เปอร์เซนต์การปลอมปน p=5% และสเกลแฟคเตอร์ c=3 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี ยกเว้นในกรณีที่ a=3,b=3,n=5

เมื่อค่าคงที่ k=1,3,5 เปอร์เซนต์การปลอมปน p=10%,25% และสเกลแฟคเตอร์ c=3,10 ค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบบูตสแตรปมีค่าต่ำกว่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีแบบคลาสสิกทุกกรณี

จากตารางทั้งหมด อาจสรุปได้ว่า

สำหรับการแจกแจงปกติปลอมปนเมื่อเปอร์เซนต์การปลอมปนและค่าสเกลแฟคเตอร์อยู่ในระดับที่สูง วิธีแบบบูตสแตรปจะให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าวิธีแบบคลาสสิก แต่เมื่อเปอร์เซนต์การปลอมปนและค่าสเกลแฟคเตอร์อยู่ในระดับที่ต่ำวิธีแบบคลาสสิกอาจจะให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าวิธีแบบบูตสแตรป

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์มีค่าหนึ่งๆ แต่ค่าคงที่ k มีค่าเพิ่มขึ้น ระยะทางยุคลิคเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มเพิ่มขึ้น

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ ค่าคงที่ k และเปอร์เซนต์การปลอมปนมีค่าหนึ่งๆ แต่สเกลแฟคเตอร์มีค่าเพิ่มขึ้น ระยะทางยุคลิคเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มเพิ่มขึ้น

เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ ค่าคงที่ k และสเกลแฟคเตอร์มีค่าหนึ่งๆ แต่เปอร์เซนต์การปลอมปนมีค่าเพิ่มขึ้น ระยะทางยุคลิคเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มเพิ่มขึ้น

เมื่อค่าคงที่ k ระดับปัจจัยทดลอง เปอร์เซนต์การปลอมปน และสเกลแฟคเตอร์มีค่าหนึ่งๆ แต่จำนวนหน่วยทดลองที่ใช้มีค่าเพิ่มขึ้น ระยะทางยุคลิคเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มลดลง

เมื่อค่าคงที่ k จำนวนหน่วยทดลองที่ใช้ เปอร์เซนต์การปลอมปน และสเกลแฟคเตอร์มีค่าหนึ่งๆ แต่ระดับปัจจัยทดลองมีค่าเพิ่มขึ้น ระยะทางยุคลิคเฉลี่ยที่คำนวณได้จากทั้งสองวิธีมีแนวโน้มลดลง

ตารางที่ 4.19 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	1	5%,3	417.56944	403.4591**
		5%,10	1968.2110	1880.7580**
		10%,3	514.9052	501.2716**
		10%,10	3023.9118	2943.0154**
		25%,3	848.3456	835.6557**
		25%,10	7430.4470	7359.0000**
	3	5%,3	1097.9464	1078.3450**
		5%,10	5000.4955	4889.1501**
		10%,3	1486.9344	1470.7988**
		10%,10	8368.7406	8295.6572**
		25%,3	2525.5189	2501.4501**
		25%,10	20899.6394	20650.5616**
	5	5%,3	1781.8041	1761.7435**
		5%,10	7576.2691	7493.3037**
		10%,3	2428.6067	2410.1663**
		10%,10	15239.9496	15180.9914**
		25%,3	3873.0479	3848.6693**
		25%,10	33202.7225	32997.5079**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.20 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=5	1	5%,3	368.5050	365.9295**
		5%,10	1629.3689	1614.0008**
		10%,3	500.3653	494.4215**
		10%,10	2918.0704	2826.1658**
		25%,3	845.9495	824.1540**
		25%,10	7290.2611	7165.6059**
	3	5%,3	1065.6963	1061.4582**
		5%,10	4661.0991	4643.1367**
		10%,3	1375.3080	1366.4624**
		10%,10	8235.6700	8173.7615**
		25%,3	2416.5211	2393.5070**
		25%,10	19509.3181	19326.2052**
	5	5%,3	1765.551	1760.6041**
		5%,10	7327.5122	7307.7623**
		10%,3	2416.7002	2410.0516**
		10%,10	14063.2192	13975.8675**
		25%,3	3613.0868	3581.3361**
		25%,10	31137.1736	30870.9977**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.21 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=2,b=2,n=7 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=7	1	5%,3	363.5582**	365.1190
		5%,10	1608.3689	1588.6727**
		10%,3	467.7929	458.2086**
		10%,10	2245.6604	2185.8888**
		25%,3	811.3360	794.3242**
		25%,10	6985.3542	6859.1884**
	3	5%,3	1033.3320	1023.8180**
		5%,10	4394.9133	4343.0965**
		10%,3	1263.6612	1250.7834**
		10%,10	7956.7613	7875.6572**
		25%,3	2181.1833	2164.0895**
		25%,10	19113.8137	19002.2514**
	5	5%,3	1707.1557	1697.4085**
		5%,10	6890.4855	6863.6984**
		10%,3	2267.5553	2262.7150**
		10%,10	13768.2733	13678.5050**
		25%,3	3447.5782	3433.9205**
		25%,10	29554.6919	29421.9495**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.22 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=3,b=3,n=3	1	5%,3	257.3032	256.7786**
		5%,10	1198.4981	1193.5954**
		10%,3	360.2203	353.7301**
		10%,10	2338.9617	2249.1288**
		25%,3	604.5623	593.7584**
		25%,10	5220.1954	5103.3642**
	3	5%,3	737.7934	736.1676**
		5%,10	3270.1837	3246.4352**
		10%,3	977.1154	973.7881**
		10%,10	6269.9656	6210.3864**
		25%,3	1602.9117	1599.6871**
		25%,10	14149.8920	14055.5955**
	5	5%,3	1219.8779	1218.2307**
		5%,10	5316.3360	5294.0917**
		10%,3	1557.9195	1554.1701**
		10%,10	9615.2630	9592.3744**
		25%,3	2655.9915	2652.6202**
		25%,10	21788.1151	21726.9465**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.23 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=3,b=3,n=5	1	5%,3	254.7186	253.1935**
		5%,10	1157.0597	1153.5954**
		10%,3	338.1144	337.5100**
		10%,10	2098.8297	2086.3903**
		25%,3	576.8553	573.6681**
		25%,10	5036.1566	4998.0000**
	3	5%,3	690.4847	688.4316**
		5%,10	3227.3666	3212.6333**
		10%,3	953.2127	950.2110**
		10%,10	5744.1508	5712.5844**
		25%,3	1588.3785	1577.6871**
		25%,10	13978.2019	13900.1531**
	5	5%,3	1143.0002**	1147.1456
		5%,10	5076.4524	5061.8524**
		10%,3	1524.7850	1518.9086**
		10%,10	9418.1501	9398.3201**
		25%,3	2542.8420	2538.3078**
		25%,10	21786.1151	21725.4774**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.24 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=3,b=3,n=7 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=3,b=3,n=7	1	5%,3	241.4314	240.4876**
		5%,10	1120.1645	1104.0299**
		10%,3	316.0447	312.3369**
		10%,10	1804.7330	1768.6950**
		25%,3	529.1199	522.4385**
		25%,10	4567.9603	4498.9811**
	3	5%,3	687.5159	681.7292**
		5%,10	3034.6222	3021.6107**
		10%,3	913.7051	905.2438**
		10%,10	5592.8969	5579.3509**
		25%,3	1449.9087	1439.5780**
		25%,10	12457.9077	12358.8643**
	5	5%,3	1123.8271	1117.5793**
		5%,10	4962.5727	4942.8781**
		10%,3	1446.4789	1441.0194**
		10%,10	8761.8781	8686.5010**
		25%,3	2529.8955	2524.7963**
		25%,10	20535.0322	20446.6275**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.25 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=4,b=4,n=3	1	5%,3	194.6896**	194.9471
		5%,10	965.0472	964.8018**
		10%,3	275.9573	266.512**
		10%,10	1692.1184	1651.6129**
		25%,3	463.5209	446.8966**
		25%,10	3975.5519	3856.5725**
	3	5%,3	553.0875	551.8737**
		5%,10	2569.9687	2556.7837**
		10%,3	778.0713	777.3035**
		10%,10	4924.8458	4900.1285**
		25%,3	1228.1513	1224.4647**
		25%,10	10526.1375	10522.2777**
	5	5%,3	926.9646	925.297**
		5%,10	4217.1514	4216.6135**
		10%,3	1210.6017	1208.4933**
		10%,10	6980.4189	6975.2707**
		25%,3	2068.9870	2067.1612**
		25%,10	19649.0414	19636.9038**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.26 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=4,b=4,n=5	1	5%,3	190.0598	189.9054**
		5%,10	887.1403	881.6759**
		10%,3	272.5871	256.5354**
		10%,10	1565.7202	1555.8791**
		25%,3	432.9399	431.7509**
		25%,10	3589.6560	3579.7086**
	3	5%,3	536.5813	536.2803**
		5%,10	2455.3229	2454.3762**
		10%,3	742.6070	742.0010**
		10%,10	4312.2032	4295.5016**
		25%,3	1180.2815	1179.0553**
		25%,10	10241.9648	10234.0635**
	5	5%,3	921.8632	918.9241**
		5%,10	3943.4174	3934.5543**
		10%,3	1152.6986	1151.4856**
		10%,10	6761.4857	6750.8412**
		25%,3	2026.0535	2021.1593**
		25%,10	17279.2413	17238.4861**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.27 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้คือ a=4,b=4,n=7 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน

ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้	ค่าคงที่ k	เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=4,b=4,n=7	1	5%,3	186.1717	185.9516**
		5%,10	860.2082	846.9741**
		10%,3	243.9381	242.7135**
		10%,10	1544.3653	1535.4210**
		25%,3	406.9865	402.6286**
		25%,10	3515.7423	3491.7189**
	3	5%,3	511.8489**	513.6150
		5%,10	2303.8710	2297.1956**
		10%,3	696.9085	696.0989**
		10%,10	4226.6302	4223.1355**
		25%,3	1093.0587	1091.7520**
		25%,10	9932.39887	9925.30149**
	5	5%,3	889.0714	888.9256**
		5%,10	3911.1326	3893.6369**
		10%,3	1062.6740	1058.9290**
		10%,10	6640.1198	6584.0650**
		25%,3	1853.1531	1850.2017**
		25%,10	16587.6380	16574.7630**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.28 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 5% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	417.56944	403.4591**
a=2,b=2,n=5	368.5050	365.9295**
a=2,b=2,n=7	363.5582**	365.1190
a=3,b=3,n=3	257.3032	256.7786**
a=3,b=3,n=5	254.7186	253.1935**
a=3,b=3,n=7	241.4314	240.4876**
a=4,b=4,n=3	194.6896**	194.9471
a=4,b=4,n=5	190.0598	189.9054**
a=4,b=4,n=7	186.1717	185.9516**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.29 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 5% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	1097.9464	1078.3450**
a=2,b=2,n=5	1065.6963	1061.4582**
a=2,b=2,n=7	1033.3320	1023.8180**
a=3,b=3,n=3	737.7934	736.1676**
a=3,b=3,n=5	690.4847	688.4316**
a=3,b=3,n=7	687.5159	681.7292**
a=4,b=4,n=3	553.0875	551.8737**
a=4,b=4,n=5	536.5813	536.2803**
a=4,b=4,n=7	511.8489**	513.6150

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.30 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 5% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	1781.8041	1761.7435**
a=2,b=2,n=5	1765.551	1760.6041**
a=2,b=2,n=7	1707.1557	1697.4085**
a=3,b=3,n=3	1219.8779	1218.2307**
a=3,b=3,n=5	1143.0002**	1147.1456
a=3,b=3,n=7	1123.8271	1117.5793**
a=4,b=4,n=3	926.9646	925.297**
a=4,b=4,n=5	921.8632	918.9241**
a=4,b=4,n=7	889.0714	888.9256**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.31 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 5% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	1968.2110	1880.7580**
a=2,b=2,n=5	1629.3689	1614.0008**
a=2,b=2,n=7	1608.3689	1588.6727**
a=3,b=3,n=3	1198.4981	1193.5954**
a=3,b=3,n=5	1157.0597	1153.5954**
a=3,b=3,n=7	1120.1645	1104.0299**
a=4,b=4,n=3	965.0472	964.8018**
a=4,b=4,n=5	887.1403	881.6759**
a=4,b=4,n=7	860.2082	846.9741**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.32 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 5% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	5000.4955	4889.1501**
a=2,b=2,n=5	4661.0991	4643.1367**
a=2,b=2,n=7	4394.9133	4343.0965**
a=3,b=3,n=3	3270.1837	3246.4352**
a=3,b=3,n=5	3227.3666	3212.6333**
a=3,b=3,n=7	3034.6222	3021.6107**
a=4,b=4,n=3	2569.9687	2556.7837**
a=4,b=4,n=5	2455.3229	2454.3762**
a=4,b=4,n=7	2303.8710	2297.1956**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.33 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 5% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	7576.2691	7493.3037**
a=2,b=2,n=5	7327.5122	7307.7623**
a=2,b=2,n=7	6890.4855	6863.6984**
a=3,b=3,n=3	5316.3360	5294.0917**
a=3,b=3,n=5	5076.4524	5061.8524**
a=3,b=3,n=7	4962.5727	4942.8781**
a=4,b=4,n=3	4217.1514	4216.6135**
a=4,b=4,n=5	3943.4174	3934.5543**
a=4,b=4,n=7	3911.1326	3893.6369**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.34 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 10% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	514.9052	501.2716**
a=2,b=2,n=5	500.3653	494.4215**
a=2,b=2,n=7	467.7929	458.2086**
a=3,b=3,n=3	360.2203	353.7301**
a=3,b=3,n=5	338.1144	337.5100**
a=3,b=3,n=7	316.0447	312.3369**
a=4,b=4,n=3	275.9573	266.512**
a=4,b=4,n=5	272.5871	256.5354**
a=4,b=4,n=7	243.9381	242.7135**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.35 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 10% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	1486.9344	1470.7988**
a=2,b=2,n=5	1375.3080	1366.4624**
a=2,b=2,n=7	1263.6612	1250.7834**
a=3,b=3,n=3	977.1154	973.7881**
a=3,b=3,n=5	953.2127	950.2110**
a=3,b=3,n=7	913.7051	905.2438**
a=4,b=4,n=3	778.0713	777.3035**
a=4,b=4,n=5	742.6070	742.0010**
a=4,b=4,n=7	696.9085	696.0989**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.36 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 10% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	2428.6067	2410.1663**
a=2,b=2,n=5	2416.7002	2410.0516**
a=2,b=2,n=7	2267.5553	2262.7150**
a=3,b=3,n=3	1557.9195	1554.1701**
a=3,b=3,n=5	1524.7850	1518.9086**
a=3,b=3,n=7	1446.4789	1441.0194**
a=4,b=4,n=3	1210.6017	1208.4933**
a=4,b=4,n=5	1152.6986	1151.4856**
a=4,b=4,n=7	1062.6740	1058.9290**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.37 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 10% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	3023.9118	2943.0154**
a=2,b=2,n=5	2918.0704	2826.1658**
a=2,b=2,n=7	2245.6604	2185.8888**
a=3,b=3,n=3	2338.9617	2249.1288**
a=3,b=3,n=5	2098.8297	2086.3903**
a=3,b=3,n=7	1804.7330	1768.6950**
a=4,b=4,n=3	1692.1184	1651.6129**
a=4,b=4,n=5	1565.7202	1555.8791**
a=4,b=4,n=7	1544.3653	1535.4210**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.38 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 10% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	8368.7406	8295.6572**
a=2,b=2,n=5	8235.6700	8173.7615**
a=2,b=2,n=7	7956.7613	7875.6572**
a=3,b=3,n=3	6269.9656	6210.3864**
a=3,b=3,n=5	5744.1508	5712.5844**
a=3,b=3,n=7	5592.8969	5579.3509**
a=4,b=4,n=3	4924.8458	4900.1285**
a=4,b=4,n=5	4312.2032	4295.5016**
a=4,b=4,n=7	4226.6302	4223.1355**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.39 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 10% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	15239.9496	15180.9914**
a=2,b=2,n=5	14063.2192	13975.8675**
a=2,b=2,n=7	13768.2733	13678.5050**
a=3,b=3,n=3	9615.2630	9592.3744**
a=3,b=3,n=5	9418.1501	9398.3201**
a=3,b=3,n=7	8761.8781	8686.5010**
a=4,b=4,n=3	6980.4189	6975.2707**
a=4,b=4,n=5	6761.4857	6750.8412**
a=4,b=4,n=7	6640.1198	6584.0650**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.40 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 25% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	848.3456	835.6557**
a=2,b=2,n=5	845.9495	824.1540**
a=2,b=2,n=7	811.3360	794.3242**
a=3,b=3,n=3	604.5623	593.7584**
a=3,b=3,n=5	576.8553	573.6681**
a=3,b=3,n=7	529.1199	522.4385**
a=4,b=4,n=3	463.5209	446.8966**
a=4,b=4,n=5	432.9399	431.7509**
a=4,b=4,n=7	406.9865	402.6286**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.41 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 25% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	2525.5189	2501.4501**
a=2,b=2,n=5	2416.5211	2393.5070**
a=2,b=2,n=7	2181.1833	2164.0895**
a=3,b=3,n=3	1602.9117	1599.6871**
a=3,b=3,n=5	1588.3785	1577.6871**
a=3,b=3,n=7	1449.9087	1439.5780**
a=4,b=4,n=3	1228.1513	1224.4647**
a=4,b=4,n=5	1180.2815	1179.0553**
a=4,b=4,n=7	1093.0587	1091.7520**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.42 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 25% และสเกลแฟคเตอร์เท่ากับ 3

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	3873.0479	3848.6693**
a=2,b=2,n=5	3613.0868	3581.3361**
a=2,b=2,n=7	3447.5782	3433.9205**
a=3,b=3,n=3	2655.9915	2652.6202**
a=3,b=3,n=5	2542.8420	2538.3078**
a=3,b=3,n=7	2529.8955	2524.7963**
a=4,b=4,n=3	2068.9870	2067.1612**
a=4,b=4,n=5	2026.0535	2021.1593**
a=4,b=4,n=7	1853.1531	1850.2017**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.43 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=1 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 25% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	7430.4470	7359.0000**
a=2,b=2,n=5	7290.2611	7165.6059**
a=2,b=2,n=7	6985.3542	6859.1884**
a=3,b=3,n=3	5220.1954	5103.3642**
a=3,b=3,n=5	5036.1566	4998.0000**
a=3,b=3,n=7	4567.9603	4498.9811**
a=4,b=4,n=3	3975.5519	3856.5725**
a=4,b=4,n=5	3589.6560	3579.7086**
a=4,b=4,n=7	3515.7423	3491.7189**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.44 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=3 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 25% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ยแบบคลาสสิก	ระยะทางยุคลิดเฉลี่ยแบบบูตสแตรป
a=2,b=2,n=3	20899.6394	20650.5616**
a=2,b=2,n=5	19509.3181	19326.2052**
a=2,b=2,n=7	19113.8137	19002.2514**
a=3,b=3,n=3	14149.8920	14055.5955**
a=3,b=3,n=5	13978.2019	13900.1531**
a=3,b=3,n=7	12457.9077	12358.8643**
a=4,b=4,n=3	10526.1375	10522.2777**
a=4,b=4,n=5	10241.9648	10234.0635**
a=4,b=4,n=7	9932.39887	9925.30149**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

ตารางที่ 4.45 แสดงการเปรียบเทียบค่าระยะทางยุคลิดเฉลี่ยที่คำนวณได้จากวิธีคลาสสิก และวิธีบูตสแตรป เมื่อค่าคงที่ k=5 และความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่เปอร์เซนต์การปลอมปนเท่ากับ 25% และสเกลแฟคเตอร์เท่ากับ 10

ระดับปัจจัย และขนาดหน่วยทดลองที่ใช้	ระยะทางยุคลิดเฉลี่ย แบบคลาสสิก	ระยะทางยุคลิดเฉลี่ย แบบบูตสแตรป
a=2,b=2,n=3	33202.7225	32997.5079**
a=2,b=2,n=5	31137.1736	30870.9977**
a=2,b=2,n=7	29554.6919	29421.9495**
a=3,b=3,n=3	21788.1151	21726.9465**
a=3,b=3,n=5	21786.1151	21725.4774**
a=3,b=3,n=7	20535.0322	20446.6275**
a=4,b=4,n=3	19649.0414	19636.9038**
a=4,b=4,n=5	17279.2413	17238.4861**
a=4,b=4,n=7	16587.6380	16574.7630**

หมายเหตุ ** แทนระยะทางยุคลิดเฉลี่ยที่มีค่าต่ำกว่า

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบวิธีการประมาณค่าองค์ประกอบความแปรปรวน สำหรับตัวแบบแผนแบบการทดลองข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม 2 วิธี คือ 1) วิธีแบบคลาสสิก (Classical Estimation) 2) วิธีแบบบูตสแตรป (Bootstrap Estimation) โดยศึกษาภายใต้การแจกแจงของความคลาดเคลื่อน 2 ลักษณะ คือ ความคลาดเคลื่อนมีการแจกแจงแบบปกติ ซึ่งศึกษาภายใต้สถานการณ์ต่าง ๆ ดังนี้ ปัจจัยทดลอง A มีอยู่ a ระดับ ปัจจัยทดลอง B มีอยู่ b ระดับ และขนาดหน่วยทดลองที่ใช้ n กล่าวคือ a=2,b=2,n=3,5,7 a=3,b=3,n=3,5,7 a=4,b=4,n=3,5,7 สัมประสิทธิ์การแปรผันที่ 10%,30%,50% และค่าคงที่ k เป็น 1,3,5 และความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปนซึ่งศึกษาภายใต้สถานการณ์ต่าง ๆ ดังนี้ ปัจจัยทดลอง A มีอยู่ a ระดับ ปัจจัยทดลอง B มีอยู่ b ระดับ และขนาดหน่วยทดลองที่ใช้ n กล่าวคือ a=2,b=2,n=3,5,7 a=3,b=3,n=3,5,7 a=4,b=4,n=3,5,7 ค่าคงที่ k เป็น 1,3,5 เปอร์เซนต์ของการปลอมปนเป็น 5% 10% และ 25% และสเกลแฟคเตอร์ 2 ระดับ คือ 3 และ 10 ตามลำดับ ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการประมาณทั้ง 2 แบบ เพื่อหาวิธีการประมาณที่เหมาะสมที่สุดคือระยะทางยุคลิดเฉลี่ย กล่าวคือ ถ้าวิธีการประมาณแบบใดให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่า ก็จะเป็นวิธีการประมาณที่เหมาะสม ซึ่งแสดงได้ว่าค่าประมาณขององค์ประกอบความแปรปรวนแบบจุดที่ได้มีค่าใกล้เคียงค่าจริงขององค์ประกอบความแปรปรวนนั้น ซึ่งข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองด้วยเทคนิคมอนติคาร์โลด้วยโปรแกรม S-PLUS 2000

5.1 สรุปผลการวิจัย

5.1.1 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ

5.1.1.1 วิธีการประมาณค่าองค์ประกอบความแปรปรวนด้วยวิธีแบบบูตสแตรปจะให้ค่าระยะทางยุคลิดต่ำกว่าวิธีแบบคลาสสิกที่ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้มีขนาดต่ำ แต่เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้เพิ่มขึ้นวิธีคลาสสิกจะให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่า

5.1.1.2 ที่ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ และค่าสัมประสิทธ์การแปรผันที่ใช้หนึ่งๆ เมื่อค่าคงที่ k เพิ่มขึ้น จะให้ระยะทางยุคลิดเฉลี่ยที่มีแนวโน้มเพิ่มขึ้นสำหรับทั้งสองวิธี

5.1.1.3 ที่ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ และค่าคงที่ k มีค่าหนึ่ง ๆ เมื่อสัมประสิทธิ์การแปรผันมีค่าเพิ่มขึ้น จะให้ระยะทางยุคลิดเฉลี่ยเพิ่มขึ้นด้วยสำหรับสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผันตามสัมประสิทธิ์การแปรผัน

5.1.1.4 ที่ค่าสัมประสิทธิ์การแปรผันและค่าคงที่ k หนึ่ง ๆ เมื่อขนาดหน่วยทดลองที่ใช้คงที่ แต่ระดับปัจจัยมีขนาดเพิ่มขึ้นจะให้ระยะทางยุคลิดเฉลี่ยลดลง สำหรับทั้งสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผกผันกับระดับปัจจัย

5.1.1.5 ที่ระดับสัมประสิทธิ์การแปรผันและค่าคงที่ k หนึ่ง ๆ เมื่อระดับปัจจัยคงที่ แต่ขนาดหน่วยทดลองที่ใช้เพิ่มขึ้นจะให้ระยะทางยุคลิดเฉลี่ยลดลง สำหรับทั้งสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผกผันกับขนาดหน่วยทดลองที่ใช้

5.1.2 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน

5.1.2.1 วิธีการประมาณค่าองค์ประกอบความแปรปรวนด้วยวิธีแบบบูตสแตรปให้ค่าระยะทางยุคลิดเฉลี่ยต่ำกว่าวิธีแบบคลาสสิก ที่ทุกระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ ทุกค่าคงที่ k ทุกค่าเปอร์เซนต์การปลอมปน และทุกค่าสเกลแฟคเตอร์ ยกเว้นในกรณีที่เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์มีค่าต่ำในบางกรณีเท่านั้น

5.1.2.2 ที่ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์มีค่าหนึ่งๆ เมื่อค่าคงที่ k เพิ่มขึ้น จะให้ระยะทางยุคลิดเฉลี่ยที่มีแนวโน้มเพิ่มขึ้นสำหรับทั้งสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผันตามค่าคงที่ k

5.1.2.3 ที่ระดับปัจจัยและขนาดหน่วยทดลองที่ใช้ ค่าคงที่ k มีค่าหนึ่ง ๆ เมื่อเปอร์เซนต์การปลอมปนมีค่าเพิ่มขึ้นแต่ค่าสเกลแฟคเตอร์คงที่ จะให้ระยะทางยุคลิดเฉลี่ยเพิ่มขึ้นด้วยสำหรับสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผันตามเปอร์เซนต์การปลอมปน

5.1.2.4 ที่ระดับปัจจัยและหน่วยทดลองที่ใช้ ค่าคงที่ k มีค่าหนึ่งๆ เมื่อเปอร์เซนต์การปลอมปนมีค่าคงที่แต่ค่าสเกลแฟคเตอร์เพิ่มขึ้น จะให้ระยะทางยุคลิดเฉลี่ยเพิ่มขึ้นด้วยสำหรับสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผันตามสเกลแฟคเตอร์

5.1.2.5 ที่ค่าคงที่ k เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์มีค่าหนึ่ง ๆ เมื่อระดับปัจจัยมีขนาดเพิ่มขึ้นแต่ขนาดหน่วยทดลองที่ใช้คงที่ จะให้ระยะทางยุคลิดเฉลี่ยลดลงสำหรับทั้งสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผกผันกับระดับปัจจัย

5.1.2.6 ที่ค่าคงที่ k เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์มีค่าหนึ่ง ๆ เมื่อระดับปัจจัยคงที่แต่ขนาดหน่วยทดลองที่ใช้เพิ่มขึ้นจะให้ระยะทางยุคลิดเฉลี่ยลดลง สำหรับทั้งสองวิธี สอดคล้องตามทฤษฎี เนื่องจากระยะทางยุคลิดเฉลี่ยแปรผกผันกับขนาดหน่วยทดลองที่ใช้

นั่นคือเราควรประมาณค่าองค์ประกอบความแปรปรวนด้วยวิธีบูตสแตรปในกรณีความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน ที่มีเปอร์เซนต์การปลอมปน และสเกลแฟคเตอร์ในระดับสูง และกรณีความคลาดเคลื่อนมีการแจกแจงปกติที่ระดับปัจจัย และขนาดหน่วยตัวอย่างที่ใช้ในการทดลองต่ำ

5.2 ข้อเสนอแนะ

จากผลการวิจัยจะเสนอแนะเป็น 2 ด้านคือ

5.2.1 ด้านการนำไปใช้

5.2.1.1 สำหรับความคลาดเคลื่อนที่มีการแจกแจงแบบปกติ เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้มีค่าต่ำ ควรใช้วิธีบูตสแตรปในการประมาณค่าองค์ประกอบความแปรปรวน แต่เมื่อระดับปัจจัยและขนาดหน่วยทดลองที่ใช้มีขนาดมากพอ ควรใช้วิธีแบบคลาสสิกในการประมาณค่าองค์ประกอบความแปรปรวน

5.2.1.2 สำหรับความคลาดเคลื่อนที่มีการแจกแจงแบบปกติปลอมปน ควรใช้วิธีบูตสแตรปในการประมาณค่าองค์ประกอบความแปรปรวน ยกเว้นในกรณีที่เปอร์เซนต์การปลอมปนและสเกลแฟคเตอร์อยู่ในระดับต่ำ

5.2.2 ด้านการศึกษาวิจัย

5.2.2.1 จำนวนการสุ่มตัวอย่างซ้ำแบบใส่คืนด้วยวิธีบูตสแตรป (Bootstrap Replications) นอกจากจะหาได้จากวิธีกระทำซ้ำเพื่อให้ค่าที่ต้องการประมาณลู่เข้าสู่ค่าคงที่แล้ว (Convergence) อีกวิธีหนึ่งคืออาจสามารถ หาค่า B ได้จากสูตรในการประมาณค่า B ตามแนวคิดของเอฟรอน

5.2.2.2 ในการวิจัยครั้งนี้ ได้ศึกษาและเปรียบเทียบการประมาณค่าองค์ประกอบความแปรปรวนสำหรับการทดลองปัจจัยพหุ กรณีตัวแบบข้ามกลุ่ม 2 ปัจจัยเชิงสุ่ม ระดับปัจจัยทดลอง A เท่ากับ a ระดับปัจจัยทดลอง Bเท่ากับ b เมื่อขนาดตัวอย่างที่ใช้ เป็น 3,5,7 สำหรับการวิจัยครั้งต่อไปอาจทำการศึกษาในกรณีที่ ระดับปัจจัยทดลอง A ไม่เท่ากับ ระดับปัจจัยทดลอง B

5.2.2.3 ในการวิจัยครั้งนี้กำหนดให้ $\sigma_\alpha^2 = \sigma_\beta^2 = \sigma_{\alpha\beta}^2 = k\sigma_\varepsilon^2$ โดยที่ k เป็นจำนวนเต็มที่คงที่ เท่ากับ 1,3,5 ในการวิจัยครั้งต่อไปอาจทำการศึกษาโดย $\sigma_\alpha^2 = k_1\sigma_\varepsilon^2$, $\sigma_\beta^2 = k_2\sigma_\varepsilon^2$, $\sigma_{\alpha\beta}^2 = k_3\sigma_\varepsilon^2$ โดยที่ $k_1 \neq k_2 \neq k_3$ และ $k_1 > 0$, $k_2 > 0$, $k_3 > 0$

5.2.2.4 ในการวิจัยครั้งนี้ได้ทำการศึกษาและเปรียบเทียบเมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ และความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน ในการวิจัยครั้งต่อไปอาจทำการศึกษาความคลาดเคลื่อนที่มีการแจกแจงอื่น เช่นการแจกแจงแบบดับเบิ้ลเอกซ์โปเนนเชียล

รายการอ้างอิง

ภาษาไทย

มาลี ตระการศิรินนท์. การเปรียบเทียบการประมาณค่าพารามิเตอร์ของรูปแบบสมการถดถอยเชิงเส้นด้วยวิธีกำลังสองต่ำสุดและวิธีบูตสแตรป. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2531.

มนูญ ศรีวิรัตน์. การเปรียบเทียบวิธีสำหรับการประมาณพารามิเตอร์ระหว่างวิธีกำลังสองน้อยที่สุดกับวิธีบูตสแตรปในการวิเคราะห์ความแปรปรวนร่วม. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2534.

สุพล ดุรงค์วัฒนา. การวิเคราะห์เชิงสถิติ : การวิเคราะห์ความแปรปรวน. ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพมหานคร, 2537.

สุพล ดุรงค์วัฒนา. การวางแผนการทดลองขั้นสูง. เอกสารประกอบการสอนวิชาการวางแผนการทดลองขั้นสูง ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2541.

สุรพล อุปดิสกุล. สถิติการวางแผนการทดลองเบื้องต้น. กรุงเทพฯ: สหมิตรออฟเซท, 2526.

ภาษาอังกฤษ

Canty, A. J. Lecture Notes MAST679T: Introduction to Bootstrap. Department of Mathematics and Statistics, Concordia University, 2000.

Cochran, W. G., and Cox, G. M. Experimental Design. New York: John Hiley and Sons,1976.

Cook, D., and Duckworth, W. M. Resampling methods for Inference. University of Iowa, 2001.

Dean, A. M. and Voss, D. T. Design and Analysis of Experiments. New York: Spinger-Verlag, 1999.

Efron, B. Better Bootstrap Confidence Intervals. The American Statistical Association Journal of the American Statistical Association Vol. 82, No. 39 (1987): 171-200.

Efron, B. Bootstrap Methods Another Look at the Jackknife. The Annals of Statistics 7(1979): 1-26.

Efron, B., and Tibshirani, R. J. An Introduction to the Bootstrap. New York: Chapman & Hall, 1993.

Getting Started with S-PLUS 2000. Data Analysis Products Division.MathSoft. US: Seattle, 1988-1999.

Hall, P. On efficient bootstrap Simulation. Biometrika 76(1989): 613-617.

Hall, P. Theoretical Comparison of Bootstrap Confidence Intervals. The Annals of Statistics 3(1988): 927-953.

Hall, P., and Martion, M. A. On Bootstrap resampling and Iteration. Biometrika 75 (1998): 661-671.

Hinkelmann, K., and Hempthorne, O. Design and Analysis of Experiments: volume 1. Canada: John Wiley and Sons, 1994.

Hjorth, J. S. U. Computer Intensive Statistical Methods. London: Chapman & Hall, 1994.

Horowitz, J. L. The Bootstrap. Department of Economics, University of Iowa, 1999.

Lagnado, R., Delianedis, G., Tikhonov, S. Monte Carlo Simulation of Non-Normal Processes. London: Madas-Kapiti, 2000.

Montgomery, C. D. Design and analysis of experiments. 4th ed. Canada: John Wiley and Sons, 1997.

Mooney, C. Z. ,and Duval, R. D. Bootstrapping: A Nonparametric Approach to Statistical Inference. California: Sage, 1993.

Pardoe, I. An Introduction to Bootstrap Methods using Arc. Department of Applied Statistics, University of Minnessota, Technical Report 631 (1999).

S-PLUS 2000 Programmer's Guide. Data Analysis Products Division. MathSoft US: Seattle,1992-1999.

S-PLUS User's Guide. Data Analysis Products Division.MathSoft. US: Seattle, 1987-1999.

Stine, R. A. Bootstrap Resampling Lectures ICPSR Summer Program. Department of Statistics of Wharton School, University of Pennsylvania, 1999.

ภาคผนวก

โปรแกรมการคำนวณหาค่าประมาณองค์ประกอบความแปรปรวน
ทั้งวิธีคลาสสิกและวิธีบูตสแตรป

(* ขั้นตอนในการกำหนดสถานการณ์ต่างๆ *)

```
sim.dat _ function(n,a,b,k1,k2,k3,u,cv,loopc,d,p,o)
```

(* สร้างความคลาดเคลื่อนมีการแจกแจงแบบปกติ *)

```
    if (d == "no" )
    {
        vare _ (cv*u)^2/(k1+k2+k3+1)
        q1 _ k1*vare
        q2 _ k2*vare
        q3 _ k3*vare
        q4 _ vare
        f1 _ array(rnorm(a,0,sqrt(q1)), dim = c(a))
        f2 _ array(rnorm(b,0,sqrt(q2)), dim = c(b))
        f3 _ array(rnorm(a*b,0,sqrt(q3)), dim = c(a,b))
        f4 _ array(rnorm(a*b*n,0,sqrt(q4)), dim =c(a,b,n))
    }
```

(* สร้างความคลาดเคลื่อนมีการแจกแจงแบบปกติปลอมปน *)

```
    if (d == "conta")
    {
        q5 _ 100
        vare _ ((1-p)*q5)+ (p*(o^2)*(q5))
        q1 _ k1*vare
        q2 _ k2*vare
        q3 _ k3*vare
        q4 _ vare
        f1 _ array(rnorm(a,0,sqrt(q1)), dim = c(a))
        f2 _ array(rnorm(b,0,sqrt(q2)), dim = c(b))
```

```
f3 _ array(rnorm(a*b,0,sqrt(q3)), dim = c(a,b))
m _ round((1-p)*a*b*n)
s _ (a*b*n)-m
t _ 0
w _ 0
tt _ 0

repeat
{
    tt _ tt+1
    x_ runif(1,0,1)

    if(x >=0 &&x < (1-p))
    {
        if(tt==1)
        {
            write(rnorm(1,0,sqrt(q5)),"conta.dat",,F)
            t_ t+1
        }
        if(tt !=1)
        {
            if (t!=m)
            {
                write(rnorm(1,0,sqrt(q5)),"conta.dat",,T)
                t_ t+1
            }
            if(t == m)
            {
                if(w != s)
                {
                    write(rnorm(1,0,sqrt((o^2)*q5)),"conta.dat",,T)
                    w_ w+1
                }
                if(w == s)
                {
                    break
                }
            }
```

```
                }
            }

            if(x>=(1-p)&& x<1)
            {
                if(tt==1)
                {
                    write(rnorm(1,0,sqrt((o^2)*q5)),"conta.dat",,F)
                    w_ w+1
                }
                if(tt !=1)
                {
                    if( w!=s)
                    {
                        write(rnorm(1,0,sqrt((o^2)*q5)),"conta.dat",,T)
                        w_ w+1
                    }
                    if(w == s)
                    {
                        if(t != m)
                        {
                            write(rnorm(1,0,sqrt(q5)),"conta.dat",,T)
                            t_ t+1
                        }
                        if(t == m)
                        {
                            break
                        }
                    }
                }
            }
        }

    f4 _ scan("conta.dat")
    f4 _ array(f4,dim=c(a,b,n))
}
```

(* คำนวณหาค่าประมาณองค์ประกอบความแปรปรวนสำหรับวิธีแบบคลาสสิก *)

```
y _ array(, dim = c(a,b,n))
for (i in 1 : a)
{
    for (j in 1 : b)
    {
        for (k in 1 : n)
        {
            y[i,j,k] _ u + f1[i] + f2[j] + f3[i,j] + f4[i,j,k]
        }
    }
}

ss _ 0
for (i in 1 : a)
{
    for (j in 1 : b)
    {
        for (k in 1 : n)
        {
            ss _ ss + y[i,j,k]
        }
    }
}
ss _ (ss^2)/(a*b*n)

se _ 0
for (i in 1 : a)
{
    for (j in 1 : b)
    {
        for (k in 1 : n)
        {
            se _ se + (y[i,j,k])^2
        }
    }
}
```

```
sa _ 0
sat _ 0
for (i in 1 : a)
{
      for (j in 1 : b)
      {
            for (k in 1 : n)
            {
                  sat _ sat + y[i,j,k]
            }
      }
      sa _ sa + (sat^2)
      sat _ 0
}
sa _ (sa)/(b*n)

sb _ 0
sbt _ 0
for (j in 1 : b)
{
      for (i in 1 : a)
      {
            for (k in 1 : n)
            {
                  sbt _ sbt + y[i,j,k]
            }
      }
      sb _ sb + (sbt^2)
      sbt _ 0
}
sb _ (sb)/(a*n)

si _ 0
sit _ 0
for (i in 1 : a)
{
      for (j in 1 : b)
```

```
        {
            for (k in 1 : n)
            {
                sit _ sit + y[i,j,k]
            }
            si _ si +(sit^2)
            sit _ 0
        }
    }
    si _ (si)/(n)

    ssa _ sa-ss
    ssb _ sb-ss
    ssi _ si-ss-ssa-ssb
    sse _ se-ss-ssa-ssb-ssi

    va _ a-1
    vb _ b-1
    vi _ (a-1)*(b-1)
    ve _ a*b*(n-1)

    msa _ round(ssa/va,6)
    msb _ round(ssb/vb,6)
    msi _ round(ssi/vi,6)
    mse _ round(sse/ve,6)

#classic
    cla _ round((msa-msi)/(b*n),6)
    clb _ round((msb-msi)/(a*n),6)
    cli _ round((msi-mse)/n,6)
    cle _ mse

    if(cla < 0)
    {
        cla _ 0
    }

    if(clb < 0)
```

```
    {
          clb _ 0
    }
if(cli < 0)
    {
          cli _ 0
    }
    if(cle < 0)
    {
          cle _ 0
    }
```

(* คำนวณค่าองค์ประกอบความแปรปรวนด้วยวิธีแบบบูตสแตรป *)

```
            #bootstrap
            ba _ array( , dim=c(loopc))
            bb _ array( , dim=c(loopc))
            bi _ array( , dim=c(loopc))
            be _ array( , dim=c(loopc))
            x _ array( , dim=c(a*b))
            z _ array( , dim=c(a,b,n))

            for(l in 1:loopc)
            {
                  for(i in 1 : a)
                  {
                        for(j in 1 : b)
                        {
                              x _ sample(y[i , j , ] , n , replace = T)
                              for(k in 1 : n)
                              {
                                    z[i,j,k] _ x[k]
                              }
                        }
                  }

                  ss _ 0
                  for (i in 1 : a)
                  {
```

```
        for (j in 1 : b)
        {
                for (k in 1 : n)
                {
                        ss _ ss + z[i,j,k]
                }
        }
}
ss _ (ss^2)/(a*b*n)

se _ 0
for (i in 1 : a)
{
        for (j in 1 : b)
        {
                for (k in 1 : n)
                {
                        se _ se + z[i,j,k]^2
                }
        }
}

sa _ 0
sat _ 0
for (i in 1 : a)
{
        for (j in 1 : b)
        {
                for (k in 1 : n)
                {
                        sat _ sat + z[i,j,k]
                }
        }
sa _ sa + (sat^2)
sat _ 0
}
sa _ (sa)/(b*n)
```

```
sb _ 0
sbt _ 0
for (j in 1 : b)
{
	for (i in 1 : a)
	{
		for (k in 1 : n)
		{
			sbt _ sbt + z[i,j,k]
		}
	}
		sb _ sb + (sbt^2)
		sbt _ 0
}
sb _ (sb)/(a*n)

si _ 0
sit _ 0
for (i in 1 : a)
{
	for (j in 1 : b)
	{
		for (k in 1 : n)
		{
			sit _ sit + z[i,j,k]
		}
		si _ si + (sit^2)
		sit _ 0
	}
}
si _ (si)/(n)

ssa _ sa-ss
ssb _ sb-ss
ssi _ si-ss-ssa-ssb
sse _ se-ss-ssa-ssb-ssi

va _ a-1
```

```
        vb _ b-1
        vi _ (a-1)*(b-1)
        ve _ a*b*(n-1)

        msa _ ssa/va
        msb _ ssb/vb
        msi _ ssi/vi
        mse _ sse/ve

        be[l] _ mse
        bi[l] _ (msi-mse)/n
        ba[l] _ (msa-msi)/(b*n)
        bb[l] _ (msb-msi)/(a*n)

        if(ba[l] < 0)
        {
                ba[l] _ 0
        }
        if(bb[l] < 0)
        {
                bb[l] _ 0
        }
        if(bi[l] < 0)
        {
                bi[l] _ 0
        }
        if(be[l] < 0)
        {
                be[l] _ 0
        }
}
bes _ 0
bis _ 0
bbs _ 0
bas _ 0
i _ 0

for    ( l in 1: loopc)
```

```
            {
                        bas _ bas + ba[I]
                        bbs _ bbs + bb[I]
                        bis _ bis + bi[I]
                        bes _ bes + be[I]
                        i _ i + 1

            }
            bas _ round(bas/i,6)
            bbs _ round(bbs/i,6)
            bis _ round(bis/i,6)
            bes _ round(bes/i,6)

            return(cla,clb,cli,cle,bas,bbs,bis,bes,q1,q2,q3,q4)
}



euclidean.cal_function(n,a,b,k1,k2,k3,u,cv,loopc,imax,d,p,o)
{
      imin _ 0
      eucl _ 0
      eub _ 0
      oldeucl _ 0
      oldeub _ 0
      scla _ 0
      sclb _ 0
      scli _ 0
      scle _ 0
      sba _ 0
      sbb _ 0
      sbi _ 0
      sbe _ 0
      h _ 0
      write(imin,"loop.dat",,F)
      repeat
      {
            h _ h+1
            tmp_sim.dat(n,a,b,k1,k2,k3,u,cv,loopc,d,p,o)
```

```
if
((tmp$cla>=0)&&(tmp$clb>=0)&&(tmp$cli>=0)&&(tmp$cle>=0)&&(tmp$bas>=0)&&(tmp$bbs>=0)&&(tmp$bis>=0)
&&(tmp$bes>=0))
{
    imin _ imin + 1
    cla _ tmp$cla
    clb _ tmp$clb
    cli _ tmp$cli
    cle _ tmp$cle
    scla _ cla + scla
    sclb _ clb + sclb
    scli _ cli + scli
    scle _ cle + scle
    bas _ tmp$bas
    bbs _ tmp$bbs
    bis _ tmp$bis
    bes _ tmp$bes
    sba _ bas + sba
    sbb _ bbs + sbb
    sbi _ bis + sbi
    sbe _ bes + sbe
    ecl _ round(sqrt((tmp$cla-tmp$q1)^2+(tmp$clb-tmp$q2)^2+(tmp$cli-tmp$q3)^2+(tmp$cle-
tmp$q4)^2),6)
    eb _ round(sqrt((tmp$bas-tmp$q1)^2+(tmp$bbs-tmp$q2)^2+(tmp$bis-tmp$q3)^2+(tmp$bes-
tmp$q4)^2),6)
    eucl _ eucl + ecl
    eub _ eub + eb
    meucl_eucl/imin
    meub_eub/imin
    print(imin)
    if (h >= 200)
        break
}
}
```

(* พิมพ์ผลลัพธ์ที่ต้องการ *)

```
    cat("Estimation",fill=T)
    cat("a:",a,"b:",b,"n:",n,"k:",k1,fill=T)
    cat("distribution:",d,"p",p,"o",o,fill=T)
    cat("cv",cv,"mean",u,fill=T)
    cat("The average of euclidean of classic is ",round(eucl/imin,6),fill=T)
    cat("The average of euclidean of bootstrap is ",round(eub/imin,6),fill=T)
}
```

(* จบการทำงานของโปรแกรม*)

ประวัติผู้เขียนวิทยานิพนธ์

นางสาว มนชยา เจียงประดิษฐ์ เกิดวันที่ 19 มีนาคม พ.ศ.2519 ที่เขตพญาไท จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาปริญญาตรีวิทยาศาสตรบัณฑิต วิชาเอกคณิตศาสตร์ มหาวิทยาลัยมหิดล ในปีการศึกษา 2540 และเข้าศึกษาต่อในหลักสูตรสถิติศาสตรมหาบัณฑิต ที่จุฬาลงกรณ์มหาวิทยาลัย เมื่อ พ.ศ. 2541