การศึกษาเปรียบเทียบการประมาณค่าจากตัวแบบการถดถอย
สำหรับข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ที่มีการแจกแจงแบบล็อกนอร์มอล

น.ส.ศิวพร ทิพย์พันธุ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561

การศึกษาเปรียบเทียบการประมาณค่าจากตัวแบบการถดถอย
สำหรับข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ที่มีการแจกแจงแบบล็อกนอร์มอล

น.ส.ศิวพร ทิพย์พันธุ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561

หัวข้อวิทยานิพนธ์	การศึกษาเปรียบเทียบการประมาณค่าจากตัวแบบการถดถอย สำหรับข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ที่มีการแจกแจงแบบล็อกนอร์มอล
โดย	น.ส.ศิวพร ทิพย์พันธุ์
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.อนุภาพ สมบูรณ์สวัสดี

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

.... คณบดีคณะพาณิชยศาสตร์และการบัญชี

(รองศาสตราจารย์ ดร.พสุ เดชะรินทร์)

คณะกรรมการสอบวิทยานิพนธ์

.... ประธานกรรมการ

(อาจารย์ ดร.อัครินทร์ ไพบูลย์พานิช)

.... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์ ดร.อนุภาพ สมบูรณ์สวัสดี)

.... กรรมการ

(อาจารย์ ดร.นัท กุลวานิช)

.... กรรมการภายนอกมหาวิทยาลัย

(ผู้ช่วยศาสตราจารย์ ดร.ธิดาพร ศุภภากร)

ศิวพร ทิพย์พันธุ์ : การศึกษาเปรียบเทียบการประมาณค่าจากตัวแบบการถดถอยสำหรับข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ที่มีการแจกแจงแบบล็อกนอร์มอล. (A COMPARATIVE STUDY ON ESTIMATION FROM REGRESSION MODEL FOR TYPE-I RIGHT-CENSORED DATA FROM LOGNORMAL DISTRIBUTION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร.อนุภาพ สมบูรณ์สวัสดี

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบวิธีการประมาณค่าจากตัวแบบการถดถอย เมื่อตัวแปรตามมีการแจกแจงแบบล็อกนอร์มอลและตัวแปรตามบางค่าเป็นข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ด้วยวิธีกำลังสองต่ำสุด (OLS) วิธีภาวะน่าจะเป็นสูงสุด (MLE) วิธีของแชตเทอร์จีและแมคลีช (CM) และวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม (MLE_EM) ข้อมูลในการศึกษาได้จากการจำลองข้อมูลจำนวน 81 สถานการณ์ สถานการณ์ละ 10,000 รอบ ขนาดตัวอย่าง (n) เท่ากับ 30, 50, 100 และเปอร์เซ็นต์การถูกตัดปลายทางขวาของตัวแปรตาม (r) เท่ากับ 10%, 20%, 30% และอัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 คือ 1:1, 1:2, 1:5 และอัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความคลาดเคลื่อน คือ 2:1, 1:1, 1:2

จากการศึกษาพบว่า 1) วิธี MLE และวิธี MLE_EM มีประสิทธิภาพสูงสุดเมื่อตัวอย่างมีขนาดใหญ่ (n=100) หรือตัวแปรตามถูกตัดปลายทางขวามาก (r=30%) ในทางกลับกัน 2) วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดเมื่อตัวอย่างมีขนาดเล็ก (n=30) หรือตัวแปรตามถูกตัดปลายทางขวาน้อย (r=10%) และ 3) วิธี CM เป็นวิธีที่มีประสิทธิภาพสูงสุดในสถานการณ์ที่เหลือ กล่าวคือ เมื่อตัวอย่างมีขนาดปานกลาง (n=50) หรือตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r=20%) นอกจากนั้นพบว่า 4) ทุกวิธีมีประสิทธิภาพมากขึ้นเมื่อตัวอย่างมีขนาดใหญ่ขึ้นหรือตัวแปรตามถูกตัดปลายทางขวาน้อยลงหรือความคลาดเคลื่อนกระจายตัวน้อยกว่าตัวแปรอิสระ

ภาควิชา	ภาควิชาสถิติ	ลายมือชื่อนิสิต ....
สาขาวิชา	สถิติ	ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก ....
ปีการศึกษา	2561	

5981543026 : MASTER OF SCIENCE

Regression Model, Type-I Right-Censoring, Variance

Siwaporn Tippan :

A COMPARATIVE STUDY ON ESTIMATION FROM REGRESSION MODEL FOR TYPE-I RIGHT-CENSORED DATA FROM LOGNORMAL DISTRIBUTION.

ADVISOR: Asst. Prof. Anupap Somboonsavatdee, Ph.D.

The objective of this research is to compare the estimation methods for log-linear regression model with dependent variable under type-I right-censoring: 1) Ordinary Least Squares Method (OLS); 2) Maximum Likelihood Estimation (MLE); 3) Chatterjee and McLeish Method (CM); and 4) Maximum Likelihood Estimation using the EM algorithm (MLE_EM). The results on this research are from 81 simulated scenarios with simulation size of 10,000. The sample sizes (n) are 30, 50, 100; the censoring proportions of data (r) are 10%, 20%, 30%; the ratio of variances of two independent variables of 2:1, 1:1, 1:2; the ratio of the sum of variances of two independent variables to error variance of 2:1, 1:1, 1:2.

The findings are: 1) MLE_EM and MLE perform the best at large sample size (n=100) or high censoring proportion (r=30%); on the other hand, 2) OLS performs the best at small sample size (n=30) or low censoring proportion (r=10%); and 3) CM performs the best generally for the rest of scenarios such that the sample size is moderate (n=50) or censoring proportion is moderate (r=20%); moreover, 4) The efficiency of all of the methods increase when sample size increase or proportion of right-censored data on dependent variable decrease or ratio of the sum of variances of two independent variables to error variance increase.

Department:	Department of Statistics	Student's Signature ....
Field of Study:	Statistics	Advisor's Signature ....
Academic Year:	2018	

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณผู้ช่วยศาสตราจารย์ ดร.อนุภาพ สมบูรณ์สวัสดี เป็นอย่างสูง ที่กรุณาสละเวลาให้คำแนะนำสั่งสอนและคำปรึกษาที่มีประโยชน์ รวมถึงช่วยแก้ไขข้อบกพร่องต่างๆ เพื่อให้ผู้วิจัยนำมาปรับปรุงจนกระทั่งวิทยานิพนธ์เสร็จสมบูรณ์

ผู้วิจัยขอกราบขอบพระคุณอาจารย์ ดร.อัครินทร์ ไพบูลย์พานิช ประธานกรรมการสอบวิทยานิพนธ์ อาจารย์ ดร.นัท กุลวานิช และผู้ช่วยศาสตราจารย์ ดร.ธิดาพร ศุภภากร กรรมการสอบวิทยานิพนธ์ที่กรุณาสละเวลาตรวจสอบและแก้ไขวิทยานิพนธ์ให้ถูกต้องสมบูรณ์ยิ่งขึ้น

นอกจากนี้ผู้วิจัยขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ให้ความรู้ทางด้านวิชาการ รวมไปถึงเจ้าหน้าที่ของภาควิชาสถิติที่ช่วยจัดทำเอกสาร และอำนวยความสะดวกในด้านต่างๆ

สุดท้ายนี้ผู้วิจัยขอกราบขอบพระคุณบิดา มารดา และทุกคนในครอบครัว รวมถึงขอบคุณน้องสาว เพื่อนและรุ่นพี่ที่คอยสนับสนุนและให้กำลังใจเสมอมา

ศิวพร ทิพย์พันธุ์

สารบัญ

หน้า

สารบัญตาราง

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันเทคนิคทางสถิติที่เป็นที่นิยม คือ การวิเคราะห์การถดถอย ซึ่งเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ สมการถดถอยที่ได้เป็นสมการที่ใช้ในการประมาณค่าเฉลี่ยของตัวแปรตามเมื่อรู้ค่าตัวแปรอิสระ และสามารถนำมาใช้ในการวางแผนในการดำเนินงานได้อีกด้วย เช่น เกศินี พรมธ (2555) ได้นำข้อมูลอุณหภูมิและความชื้นสัมพัทธ์ของห้องสะอาดซึ่งเป็นตัวแปรอิสระ และข้อมูลปริมาณการใช้พลังงานไฟฟ้าซึ่งเป็นตัวแปรตามของอุตสาหกรรมผลิตวอยซ์ คอยล์ มอเตอร์ในปี พ.ศ.2552 ถึง พ.ศ.2553 มาวิเคราะห์การถดถอย และจากสมการถดถอยที่ได้พบว่า หากกำหนดอุณหภูมิควบคุมของห้องสะอาดเพิ่มขึ้น 1 องศาเซลเซียส และให้ความชื้นสัมพัทธ์ของห้องสะอาดเพิ่มขึ้น 1% ในเวลา 24 เดือน จะทำให้ประหยัดค่าไฟฟ้าได้เท่ากับ 5,9695,556 บาท จากงานวิจัยนี้ นอกจากผู้วิจัยได้สมการถดถอยที่ทำให้ทราบระดับความสัมพันธ์ของตัวแปรอิสระกับตัวแปรตามแล้วผู้วิจัยยังสามารถนำผลวิจัยดังกล่าวมาใช้วางแผนในการใช้ห้องเพื่อให้ค่าไฟฟ้าของอุตสาหกรรมผลิตวอยซ์ คอยล์ มอเตอร์ลดลงอีกด้วยอีกด้วย เป็นต้น

สำหรับการวิเคราะห์การถดถอย ตัวแปรตามบางค่าอาจจะเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ซึ่งเป็นข้อมูลที่มีการกำหนดเวลาในการสังเกตไว้ล่วงหน้า โดยข้อมูลที่เกิดขึ้นในช่วงเวลาในการสังเกตนั้นจะเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ส่วนข้อมูลที่เกิดขึ้นหลังจากเวลาในการสังเกตจะเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ซึ่งเป็นข้อมูลที่ไม่สามารถระบุค่าที่แน่นอนได้ ดังนั้นจึงให้ข้อมูลดังกล่าวมีค่าเท่ากับระยะเวลาในการสังเกตที่กำหนดไว้ล่วงหน้า เช่น ผู้วิจัยต้องการศึกษาปัจจัยที่ส่งผลต่อการรอดชีวิต และค่าเฉลี่ยของระยะเวลารอดชีวิตของหนูที่มีเซลล์มะเร็งในร่างกาย จึงได้กำหนดระยะเวลาศึกษา 1 ปี โดยให้ตัวแปรตาม คือ ระยะเวลารอดชีวิตของหนู ส่วนตัวแปรอิสระ เช่น อายุ น้ำหนัก ระยะของโรคมะเร็ง ขนาดของก้อนมะเร็ง จำนวนหรือตำแหน่งของต่อมน้ำเหลืองที่เซลล์มะเร็งแพร่เข้าไป ความถี่ของการใช้ยาเพื่อรักษาอาการป่วยของหนูและชนิดของยาที่ใช้ในการรักษา เป็นต้น เมื่อสิ้นสุดการวิจัยแล้วพบว่าหนูบางตัวตายระหว่างการวิจัยแต่หนูบางตัวกลับมีชีวิตรอดแม้จะสิ้นสุดการวิจัยแล้ว ดังนั้นผู้วิจัยจึงสามารถบันทึกระยะเวลารอดชีวิตของหนูได้แค่บางตัวซึ่งเสียชีวิตในระหว่างดำเนินการวิจัยได้เท่านั้น ระยะเวลารอดชีวิตของหนูที่เสียชีวิตระหว่างดำเนินการวิจัย ถือว่าเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ส่วนระยะเวลารอดชีวิตของหนูที่ตายหลังสิ้นสุดการทำวิจัยแล้วถือว่าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภท

ที่ 1 ซึ่งผู้วิจัยไม่ทราบระยะเวลารอดชีวิตของหนูเหล่านี้ ดังนั้นจึงให้ระยะเวลารอดชีวิตของหนูเท่ากับ 1 ปี เป็นต้น

ในการวิเคราะห์การถดถอยนั้น ต้องมีการประมาณค่าพารามิเตอร์ของสมการถดถอยและวิธีที่ง่ายที่สุดสำหรับการประมาณค่าพารามิเตอร์ของสมการถดถอยเมื่อตัวแปรตามบางค่าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 คือ การตัดข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 นั้นทิ้งไป แต่วิธีนี้จะทำให้ค่าเฉลี่ยของค่าประมาณของตัวแปรตามที่ได้ต่ำกว่าค่าเฉลี่ยของตัวแปรตามที่แท้จริง และความคลาดเคลื่อนของการประมาณจะมีค่าสูง เนื่องจากขนาดตัวอย่างลดลง และข้อมูลที่เหลือหลังจากตัดข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ทิ้งไปไม่ได้เป็นตัวแทนที่ดีของประชากรที่สนใจศึกษา ดังนั้นผู้วิจัยจึงมีความสนใจในการหาตัวประมาณที่ทำให้ตัวแปรตามที่ได้จากการประมาณค่ามีค่าใกล้เคียงกับค่าที่แท้จริงของตัวแปรตามมากที่สุด จากการทบทวนงานวิจัยที่ผ่านมา พบว่า

บังอร กุมพล (2539) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยเมื่อค่าสังเกตของตัวแปรตามและความคลาดเคลื่อนมีการแจกแจงปกติ และค่าสังเกตของตัวแปรตามบางค่าเป็นข้อมูลที่ถูกตัดทิ้งทางขวาประเภทที่ 1 วิธีประมาณค่าที่ใช้ในการศึกษา คือ วิธีการของสมิธ วิธีการโมดิไฟด์แอคชัวเรียล และวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม ขนาดตัวอย่างเท่ากับ 10, 20, 30, 50, 100 ข้อมูลถูกตัดปลายทางขวาร้อยละ 10, 20, 30, 40, 50 จำนวนตัวแปรอิสระเท่ากับ 3 ตัว 5 ตัว และ 7 ตัว จากการศึกษาพบว่า เมื่อข้อมูลถูกตัดปลายทางขวาร้อยละ 10, 20 และ 30 ทุกขนาดตัวอย่างและจำนวนตัวแปรที่ศึกษา ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ที่ได้จากการประมาณค่าด้วยวิธีการของสมิธมีค่าต่ำที่สุด ยกเว้นเมื่อข้อมูลถูกตัดปลายทางขวาเท่ากับร้อยละ 10 ขนาดตัวอย่างเท่ากับ 10 และจำนวนตัวแปรอิสระเท่ากับ 5 ตัว ซึ่งค่า RMSE ที่ได้จากการประมาณค่าด้วยวิธีการโมดิไฟด์แอคชัวเรียลมีค่าต่ำกว่าวิธีอื่น และยกเว้นเมื่อข้อมูลถูกตัดปลายทางขวาเป็นร้อยละ 20 และ 30 จำนวนตัวแปรอิสระเท่ากับ 7 ตัว ซึ่งค่า RMSE ที่ได้จากการประมาณค่าด้วยวิธีด้วยภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็มมีค่าต่ำกว่าวิธีการอื่น และเมื่อข้อมูลถูกตัดปลายทางขวาเป็นร้อยละ 40 และ 50 ค่า RMSE ที่ได้จากการประมาณค่าด้วยวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็มมีค่าต่ำที่สุด นอกจากนั้น พบว่า ค่า RMSE แปรผกผันกับขนาดตัวอย่าง และแปรผันตรงกับร้อยละของข้อมูลที่ถูกตัดปลายทางขวา

จำเนียน จำนงค์รักษ์ (2539) ศึกษาเปรียบเทียบการประมาณค่าตัวแปรตามในสมการถดถอยโดยที่ตัวแปรตามบางค่าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 และความคลาดเคลื่อนแจกแจงแบบดับเบิลเอกซ์โพเนนเชียล แบบปกติ และแบบล็อกนอร์มอล เมื่อจำนวนตัวแปรอิสระเท่ากับ 2 ตัว โดยที่ตัวแปรอิสระตัวที่ 1 มีการแจกแจงแบบไวบูลล์และตัวแปรอิสระตัวที่ 2 มีการแจกแจงแบบปกติ

ขนาดตัวอย่างเท่ากับ 10, 20, 30, 50, 100 ข้อมูลถูกตัดปลายทางขวาร้อยละ 10, 20, 30, 40, 50 วิธีการประมาณที่ใช้ในการศึกษาวิจัยครั้งนี้คือ วิธีกำลังสองต่ำสุด วิธีการของแชตเทอร์จีและแมคลีช วิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม วิธีการของบัคเลย์และเจมส์ เมื่อสิ้นสุดการศึกษา พบว่า ค่า RMSE ที่ได้จากการประมาณค่าด้วยวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็มมีค่าต่ำกว่าวิธีอื่น ในทุกสถานการณ์ที่ทำการศึกษา และเมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติและแบบดับเบิลเอกซ์โพเนนเชียล กรณีที่จุดที่ข้อมูลถูกตัดปลายมีค่ามากกว่าค่าเฉลี่ยของข้อมูลเป็น σ_T และ $1.5\sigma_T$ ค่า RMSE ที่ได้จากการประมาณค่าด้วยวิธีของบัคเลย์และเจมส์มีค่าน้อยกว่าวิธีกำลังสองต่ำสุด และค่า RMSE ของทุกวิธีมีค่าลดลงเมื่อข้อมูลถูกตัดปลายทางขวามากขึ้น และเมื่อจุดที่ข้อมูลถูกตัดทิ้งมีค่ามากกว่าค่าเฉลี่ยของข้อมูลเป็น $2\sigma_T$ ค่า RMSE ที่ได้จากการประมาณค่าด้วยวิธีการของบัคเลย์และเจมส์มีค่ามากกว่าวิธีกำลังสองต่ำสุด และค่า RMSE ของทุกวิธีเพิ่มขึ้นเมื่อข้อมูลถูกตัดปลายทางขวามากขึ้น และเมื่อความคลาดเคลื่อนมีการแจกแจงแบบล็อกนอร์มอล ค่า RMSE ของทุกวิธีเพิ่มขึ้นเมื่อข้อมูลถูกตัดปลายทางขวามากขึ้น

สำหรับงานวิจัยในครั้งนี้ ผู้วิจัยสนใจศึกษาและเปรียบเทียบวิธีการประมาณค่าจากตัวแบบการถดถอย เมื่อตัวแปรตามบางค่าเป็นข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ที่มีความสัมพันธ์กับตัวแปรอิสระและความคลาดเคลื่อนที่มีการแจกแจงแบบล็อกนอร์มอลด้วยวิธีการประมาณค่าแบบกำลังสองต่ำสุด วิธีภาวะน่าจะเป็นสูงสุด วิธีของแชตเทอร์จีและแมคลีช และวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม ส่วนการออกแบบการทดลองในงานวิจัยครั้งนี้แตกต่างจากงานวิจัยผ่านมา กล่าวคือ จากงานวิจัยที่ผ่านมาได้พิจารณาผลของขนาดตัวอย่าง สัดส่วนของข้อมูลที่ถูกตัดปลายทางขวา จำนวนตัวแปรอิสระ และการแจกแจงของความคลาดเคลื่อน แต่ในวิจัยครั้งนี้ ผู้วิจัยได้พิจารณาผลของขนาดตัวอย่าง สัดส่วนของข้อมูลที่ถูกตัดปลายทางขวา อัตราส่วนของความแปรปรวนของตัวแปรอิสระแต่ละตัว และอัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อน รวมถึงตัวแบบค่าสังเกตของตัวแปรตามและเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่าจากตัวแบบการถดถอยมีความแตกต่างจากงานวิจัยที่ผ่านมา

1.2 วัตถุประสงค์การวิจัย

เพื่อศึกษาและเปรียบเทียบการประมาณค่าจากตัวแบบการถดถอยเมื่อตัวแปรตามบางค่าเป็นข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ที่มีความสัมพันธ์กับตัวแปรอิสระและความคลาดเคลื่อนที่มีการแจกแจงแบบล็อกนอร์มอล ด้วยวิธีกำลังสองต่ำสุด (Ordinary Least Squares Method; OLS)

วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation; MLE) วิธีของแชตเทอร์จีและแมคลีช (Chatterjee and McLeish Method; CM) และวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม (Maximum Likelihood Estimation using the EM algorithm; MLE using the EM algorithm)

1.3 นิยามและสัญลักษณ์

ธีระพร วีระถาวร (2536) ได้กล่าวว่า

1.3.1 ความแปรปรวน (Variance) ของค่าประมาณพารามิเตอร์ θ คือ ค่าคาดหวังของกำลังสองของผลต่างระหว่างค่าประมาณของพารามิเตอร์ กับค่าคาดหวังของค่าประมาณของพารามิเตอร์ กล่าวคือ $Var(\theta) = E\left[(\theta - E(\theta))^2\right]$

1.3.2 ความเอนเอียง (Bias) ของค่าประมาณของพารามิเตอร์ θ คือ ค่าความแตกต่างระหว่างค่าคาดหวังของค่าประมาณพารามิเตอร์กับค่าพารามิเตอร์ กล่าวคือ $Bias(\theta) = E(\theta) - \theta$

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 เป็นแนวทางในการเลือกตัวประมาณที่เหมาะสมกับข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 เพื่อให้ได้ค่าประมาณของตัวแปรตามใกล้เคียงกับค่าที่แท้จริงของตัวแปรตาม

1.4.2 เป็นแนวทางในการศึกษาเมื่อตัวแปรตามเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ที่มีการแจกแจงแบบอื่นๆ เช่น แกมมา เอ็กซ์โปเนนเชียล เป็นต้น

1.4.3 ตัวแปรตามที่ได้จากการประมาณค่าสามารถนำไปวิเคราะห์ทางสถิติต่อไปได้ เช่น การหาค่าเฉลี่ย เป็นต้น

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ความหมายของข้อมูลที่ถูกตัดปลาย (Censored Data)

Klein และ Moeschberger (2003) ได้อธิบายความหมายของข้อมูลที่ถูกตัด 3 ประเภท ได้แก่ ข้อมูลที่ถูกตัดปลายทางขวา ข้อมูลที่ถูกตัดปลายทางซ้าย และข้อมูลที่ถูกตัดทิ้งแบบช่วง ดังต่อไปนี้

2.1.1 ข้อมูลที่ถูกตัดปลายทางขวา (Right Censored Data) แบ่งเป็น 3 ประเภท ได้แก่

2.1.1.1 ข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 คือ ข้อมูลที่มีการกำหนดเวลาในการสังเกตไว้ล่วงหน้า โดยข้อมูลที่เกิดขึ้นระหว่างดำเนินการสังเกตนั้นจะเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ส่วนข้อมูลที่เกิดขึ้นหลังสิ้นสุดการสังเกตจะเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ซึ่งเป็นข้อมูลที่ไม่สามารถระบุค่าที่แน่นอนได้ ดังนั้นจึงให้ข้อมูลดังกล่าวมีค่าเท่ากับระยะเวลาที่ใช้ในการสังเกตที่กำหนดไว้ล่วงหน้า เช่น ผู้วิจัยต้องการศึกษาระยะเวลารอดชีวิตของหนูที่ได้รับประทานยาชนิดหนึ่ง จึงได้กำหนดระยะเวลาสังเกต 6 เดือน และผู้วิจัยพบว่าหนูบางตัวตายระหว่างการสังเกต แต่หนูบางตัวกลับมีชีวิตรอดแม้จะสิ้นสุดการสังเกตแล้ว ดังนั้น ระยะเวลารอดชีวิตของหนูที่ตายระหว่างทำการสังเกตอยู่นั้น จะถือว่าเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 โดยผู้วิจัยจะทำการบันทึกระยะเวลารอดชีวิตของหนูดังกล่าวให้เท่ากับระยะเวลารอดชีวิตที่แท้จริง ส่วนระยะเวลารอดชีวิตของหนูที่ตายหลังสิ้นสุดการสังเกตซึ่งไม่สามารถระบุระยะเวลารอดชีวิตที่แน่นอนได้ จะถือว่าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 โดยผู้วิจัยจะบันทึกระยะเวลาในการรอดชีวิตของหนูดังกล่าวให้เท่ากับ 6 เดือน เป็นต้น

ให้ T เป็นเวลาในการสังเกตที่กำหนดไว้ล่วงหน้า

$\tau_1, \tau_2, ..., \tau_n$ เป็นเวลา<u>จริง</u>จนเกิดเหตุการณ์ที่สนใจศึกษา

$t_1, t_2, ..., t_n$ เป็นเวลาจนเกิดเหตุการณ์ที่สนใจศึกษาที่<u>ถูกบันทึกโดยผู้วิจัย</u>

นิยาม $t_i = min(\tau_i, T)$

เมื่อเวลาจริงจนเกิดเหตุการณ์ที่สนใจศึกษามีค่าน้อยกว่าหรือเท่ากับเวลาในการสังเกตที่กำหนดไว้ล่วงหน้า แล้วเวลาจริงจนเกิดเหตุการณ์ที่สนใจศึกษานั้นจะเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 และจะถูกบันทึกให้เท่ากับเวลาจริงจนเกิดเหตุการณ์ที่สนใจศึกษา แต่หากเวลาจริงจนเกิดเหตุการณ์ที่สนใจศึกษามีค่ามากกว่าเวลาในการสังเกตที่กำหนดไว้ล่วงหน้า เวลาจริงจนเกิดเหตุการณ์ที่สนใจศึกษานั้นจะเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 และจะถูกบันทึกให้เท่ากับเวลาในการสังเกตที่กำหนดไว้ล่วงหน้า

กล่าวคือ $t_i = \begin{cases} \tau_i & , \tau_i \le T \\ T & , \tau_i > T \end{cases}$ (ถ้าเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1) (ถ้าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1)

และตัวระบุว่าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 คือ

$$\delta_i = \begin{cases} 0 \ , \tau_i \le T & \text{(ถ้าเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1)} \\ 1 \ , \tau_i > T & \text{(ถ้าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1)} \end{cases}$$

เพื่อความสะดวกจะเขียนในรูป (t_i, δ_i)

Jöreskog (2002) กล่าวถึงฟังก์ชันภาวะน่าจะเป็นของข้อมูล ที่บางส่วนของข้อมูลนั้นถูกตัดปลายทางขวาประเภทที่ 1 ไว้ดังนี้

$$L(t_i) = \begin{cases} f(\tau_i) & , \tau_i \le T \ \text{(ถ้าเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 มีจำนวน } m \text{ ค่า คือ } i = 1,2,\ldots,m\text{)} \\ P(\tau_i > T) = S(T), & \tau_i > T \ \text{(ถ้าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 มีจำนวน } n - m \text{ ค่า คือ } i = m+1,\ldots,n\text{)} \end{cases}$$

โดยที่ $f(\tau_i)$ คือ ฟังก์ชันความหนาแน่นน่าจะเป็นของ τ_i

$P(\tau_i > T) = S(T)$ คือ เป็นความน่าจะเป็นที่ τ_i มีค่ามากกว่า T หรือฟังก์ชันการอยู่รอดของ τ_i

ส่วนฟังก์ชันภาวะน่าจะเป็นรวม คือ

$$L = \prod_{i=1}^{m} f(\tau_i) \cdot \prod_{i=m+1}^{n} S(\mathrm{T})$$

2.1.1.2 ข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 คือ ข้อมูลที่มีการกำหนดจำนวนเหตุการณ์ที่สนใจศึกษาไว้ล่วงหน้า โดยข้อมูลที่เกิดขึ้นหลังจากได้จำนวนเหตุการณ์ที่สนใจศึกษาครบตามที่กำหนดแล้วจะเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ส่วนมากจะใช้ในการทดสอบระยะเวลาใช้งานของเครื่องมือหรืออุปกรณ์ โดยต้องเริ่มต้นทดสอบ ณ เวลาเดียวกัน และจะทำการหยุดทดลองเมื่อเกิดเหตุการณ์ที่สนใจครบตามที่กำหนดไว้ การออกแบบการทดลองแบบนี้จะช่วยประหยัดเวลาและค่าใช้จ่ายเพราะหากต้องการศึกษาตัวอย่างทั้งหมดจะต้องใช้ระยะเวลานานในการทดสอบจนเกิดเหตุการณ์ที่สนใจศึกษา เช่น ผู้วิจัยต้องการศึกษาอายุการใช้งานของหลอดไฟโดยมีการกำหนดจำนวนหลอดไฟที่ศึกษาไว้จำนวน 100 หลอด จากนั้นนำหลอดไฟมา 200 หลอด แล้วเริ่มต้นทำการทดลองโดยให้กระแสไฟฟ้าเข้าไปในหลอดไฟพร้อมกัน แล้วบันทึกเวลาเริ่มตั้งแต่เปิดสวิตซ์จนกระทั่งหลอดไฟ

เสียหลอดที่ 1 จนถึงหลอดที่ 100 และหยุดทำการทดลอง โดยอายุการใช้งานหลอดไฟของหลอดที่ 1 จนถึงหลอดที่ 100 จะเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 2 ส่วนอายุการใช้งานหลอดไฟของหลอดที่ 101 จนถึงหลอดที่ 200 จะถือเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ซึ่งผู้วิจัยไม่รู้อายุการใช้งานที่แท้จริงของหลอดไฟดังกล่าว ดังนั้นจึงบันทึกอายุการใช้งานของหลอดไฟดังกล่าวให้เท่ากับอายุการใช้งานของหลอดไฟหลอดที่ 100 เป็นต้น

2.1.1.3 ข้อมูลที่ถูกตัดปลายทางขวาแบบสุ่ม คือ ข้อมูลที่มีการกำหนดระยะเวลาในการสังเกตไว้ล่วงหน้า โดยข้อมูลที่เราสนใจศึกษาที่เกิดขึ้นระหว่างดำเนินการสังเกตนั้นจะเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาแบบสุ่ม ส่วนข้อมูลที่เกิดจากสาเหตุอื่นที่ไม่ใช่สิ่งที่เราสนใจศึกษา เช่น ผู้ป่วยขอถอนตัวออกจากการทดลองหรือผู้ป่วยตายด้วยสาเหตุอื่น เป็นต้น หรือข้อมูลที่เกิดขึ้นหลังสิ้นสุดการสังเกตจะเป็นข้อมูลที่ถูกตัดปลายทางขวาแบบสุ่ม เช่น ผู้วิจัยต้องการศึกษาระยะเวลารอดชีวิตของคนไข้ที่ป่วยเป็นโรคหัวใจที่ได้รับการผ่าตัดแบบใหม่ชนิดหนึ่ง จึงได้กำหนดระยะเวลาศึกษา 5 ปี และผู้วิจัยพบว่ามีคนไข้บางคนที่เสียชีวิตด้วยโรคหัวใจระหว่างดำเนินการวิจัย ระยะเวลารอดชีวิตของผู้ป่วยเหล่านี้ถือว่าเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาแบบสุ่ม ส่วนผู้ป่วยบางคนที่มีชีวิตอยู่แม้จะสิ้นสุดการทำวิจัยหรือผู้ป่วยที่ไม่ได้เสียชีวิตด้วยโรคหัวใจหรือผู้ป่วยที่ขอถอนตัวจากงานวิจัย ระยะเวลารอดชีวิตของผู้ป่วยเหล่านี้ถือว่าเป็นข้อมูลที่ถูกตัดปลายทางขวาแบบสุ่ม เป็นต้น

2.1.2 ข้อมูลที่ถูกตัดปลายทางซ้าย (Left censored data) คือ ข้อมูลที่เกิดขึ้นก่อนที่ผู้วิจัยจะเริ่มทำการสังเกต ซึ่งเป็นข้อมูลที่ไม่สามารถระบุค่าที่แท้จริงได้ เช่น ผู้วิจัยต้องการศึกษาอายุของผู้เข้าร่วมวิจัยที่เริ่มดื่มแอลกอฮอล์เป็นครั้งแรก เมื่อสอบถามแล้วพบว่า ผู้เข้าร่วมวิจัยบางคนไม่สามารถจำได้ว่าตนเองดื่มแอลกอฮอล์ครั้งแรกเมื่ออายุเท่าไร จึงไม่สามารถระบุอายุแท้จริงที่เริ่มดื่มแอลกอฮอล์ได้ ดังนั้นข้อมูลดังกล่าวจึงเป็นข้อมูลที่ถูกตัดปลายทางซ้าย เป็นต้น

2.1.3 ข้อมูลที่ถูกตัดทิ้งแบบช่วง (Interval censored data) คือ ข้อมูลที่ไม่สามารถระบุเวลาที่แน่นอนเมื่อเกิดเหตุการณ์ที่สนใจศึกษา แต่สามารถระบุช่วงเวลาที่เกิดเหตุการณ์ที่สนใจศึกษาได้ เช่น การศึกษาระยะเวลาการหดตัวของมะเร็งของผู้ป่วยโรคมะเร็งเต้านมที่ได้รับการรักษาด้วยวิธีการฉายรังสี ในความเป็นจริงผู้วิจัยไม่สามารถเฝ้าดูผู้ป่วยได้ตลอดเวลา แต่สามารถนัดผู้ป่วยเพื่อตรวจดูอาการทุก 1 เดือน ดังนั้นผู้วิจัยจึงไม่สามารถบันทึกวันและเวลาที่แน่นอนในขณะที่มะเร็งหดตัวได้ เพียงแต่สามารถบันทึกเดือนที่มะเร็งเกิดหดตัวได้ เป็นต้น

2.2 การแจกแจงที่ใช้ในงานวิจัย

ธีระพร วีระถาวร (2537) กล่าวว่า

2.2.1 การแจกแจงแบบปกติ (Normal Distribution)

ให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ เขียนแทนด้วย $X \sim N(\mu,\sigma)$ เมื่อ μ เป็นพารามิเตอร์แสดงสเกล และ σ เป็นพารามิเตอร์แสดงรูปร่าง แล้วฟังก์ชันความหนาแน่นน่าจะเป็น (Probability Density Function; PDF) จะอยู่ในรูป

$$f\left(x;\mu,\sigma\right)=\frac{1}{\sigma\sqrt{2\pi}}e^{\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]}$$

โดย $-\infty<x<\infty, -\infty<\mu<\infty$ และ $\sigma>0$

ค่าเฉลี่ย คือ μ และความแปรปรวน คือ σ^2

และมีฟังก์ชันการแจกแจงสะสม (Cumulative Distribution Function; CDF) คือ

$$P\left(X\le x\right)=F\left(x;\mu,\sigma\right)=\Phi\left(\frac{x-\mu}{\sigma}\right)$$

เมื่อ $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{\left(-\frac{t^2}{2}\right)}dt$ เป็น standard normal CDF

2.2.2 การแจกแจงแบบล็อกนอร์มอล (Lognormal Distribution)

ให้ Y เป็นตัวแปรสุ่มที่มีการแจกแจงแบบล็อกนอร์มอล แทนด้วย $Y \sim LOGNOR(\mu,\sigma)$ เมื่อ μ เป็นพารามิเตอร์แสดงสเกล และ σ เป็นพารามิเตอร์แสดงรูปร่าง ถ้า $Y=e^X \sim LOGNOR(\mu,\sigma)$ เมื่อ $X \sim N(\mu,\sigma)$ จะได้ว่า $\log\left(Y\right)\sim N(\mu,\sigma)$ แล้วฟังก์ชันความหนาแน่นน่าจะเป็น (Probability Density Function; PDF) ของ Y จะอยู่ในรูป

$$f\left(y;\mu,\sigma\right)=\frac{1}{y\sigma\sqrt{2\pi}}e^{\left[-\frac{1}{2}\left(\frac{\log(y)-\mu}{\sigma}\right)^2\right]}$$

โดย $y>0, -\infty<\mu<\infty$ และ $\sigma>0$

ค่าเฉลี่ย คือ $e^{\left(\mu+\frac{\sigma^2}{2}\right)}$ และความแปรปรวน คือ $[e^{\left(\sigma^2\right)}-1]e^{\left(2\mu+\sigma^2\right)}$

และมีฟังก์ชันการแจกแจงสะสม (Cumulative Distribution Function; CDF) คือ

$$P(Y \le y) = F(y;\mu,\sigma) = \Phi\left(\frac{\log(y)-\mu}{\sigma}\right)$$

เมื่อ $\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{\left(-\frac{t^2}{2}\right)} dt$ เป็น standard normal CDF

2.3 การวิเคราะห์การถดถอย

สุพล ดุรงค์วัฒนา (2558) ได้กล่าวถึง

2.3.1 ความหมายของการวิเคราะห์การถดถอย

การวิเคราะห์การถดถอยเป็นศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระว่ามีความสัมพันธ์ทางบวกหรือทางลบและมีความสัมพันธ์กันมากน้อยเพียงใด และสร้างสมการถดถอยเพื่อใช้ในการประมาณค่าเฉลี่ยของตัวแปรตามเมื่อรู้ค่าตัวแปรอิสระ

2.3.2 ข้อตกลงเบื้องต้นในการวิเคราะห์การถดถอย

ความคลาดเคลื่อนมีการแจกแจงแบบปกติโดยมีค่าเฉลี่ยเป็น 0 และมีความแปรปรวนคงที่เท่ากับ σ^2 และความคลาดเคลื่อนแต่ละตัวเป็นอิสระกัน

2.3.3 ตัวแบบการถดถอย

ให้ $Y_i^{'}$ คือ ค่าจริงของตัวแปรตาม ลำดับที่ i

Y_i คือ ค่าสังเกตของตัวแปรตาม ลำดับที่ i

$\hat{Y}_i$ คือ ค่าประมาณของตัวแปรตาม ลำดับที่ i

X_{ji} คือ ตัวแปรอิสระตัวที่ j ลำดับที่ i

β_j คือ ค่าที่แท้จริงของค่าสัมประสิทธิ์การถดถอยตัวที่ j

$\hat{\beta}_j$ คือ ค่าประมาณของค่าสัมประสิทธิ์การถดถอยตัวที่ j

β_0 คือ ระยะตัดแกน y ที่แท้จริง

$\hat{\beta}_0$ คือ ค่าประมาณของระยะตัดแกน y

ε_i คือ ความคลาดเคลื่อนที่แท้จริงลำดับที่ i โดย $\varepsilon_i = Y_i - Y_i^{'}$

e_i คือ เศษเหลือหรือค่าประมาณของความคลาดเคลื่อนลำดับที่ i โดย $e_i = Y_i - \hat{Y}_i$

p คือ จำนวนตัวแปรอิสระ

n คือ ขนาดตัวอย่าง เมื่อ i = 1, 2, ..., n และ j = 1, 2, ..., p

$$\mathbf{Y}' = \begin{bmatrix} Y_1' \\ Y_2' \\ \vdots \\ Y_n' \end{bmatrix}_{n\times 1}, \mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1}, \hat{\mathbf{Y}} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{bmatrix}_{n\times 1}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1)\times 1}, \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_p \end{bmatrix}_{(p+1)\times 1},$$

$$\mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix}_{n\times(p+1)}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n\times 1}, \mathbf{e} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}_{n\times 1}$$

2.3.3.1 ตัวแบบการถดถอย คือ

$Y_i' = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_p X_{pi}$ หรือเขียนในรูปเมทริกซ์ได้เป็น $\mathbf{Y}' = \mathbf{X}\boldsymbol{\beta}$

2.3.3.2 ค่าสังเกตตัวแบบการถดถอย

สามารถเขียนได้ 2 รูปแบบ คือ

2.3.3.2.1 $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_p X_{pi} + \varepsilon_i$

หรือเขียนในรูปเมทริกซ์ได้เป็น $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

2.3.3.2.2 $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \ldots + \hat{\beta}_p X_{pi} + e_i$

หรือเขียนในรูปเมทริกซ์ได้เป็น $\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$

2.3.3.3 ค่าพยากรณ์ตัวแบบการถดถอย คือ

$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \ldots + \hat{\beta}_p X_{pi}$ หรือเขียนในรูปเมทริกซ์ได้เป็น $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$

2.4 วิธีประมาณที่ใช้ในการศึกษา

ก่อนจะเริ่มกล่าวถึงวิธีการประมาณค่าที่ใช้ในการศึกษาในครั้งนี้ มีข้อตกลงเพิ่มเติมจากหัวข้อตัวแบบการถดถอย ดังนี้

ให้ Y_c คือ เวลาในการสังเกตที่กำหนดไว้ล่วงหน้า

$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_p X_{pi} + \varepsilon_i$ คือ เวลาจริงจนเกิดเหตุการณ์ที่สนใจศึกษา หรือค่าสังเกตของตัวแปรตามลำดับที่ i

เมื่อ i = 1, 2, ..., m-1, m, m+1, ..., n และ $m \le n$

σ คือ ส่วนเบี่ยงเบนมาตรฐานที่แท้จริง

$\hat{\sigma}$ คือ ค่าประมาณของส่วนเบี่ยงเบนมาตรฐาน

นิยาม $Y_i^* = \min\left(Y_i, Y_c\right)$ เป็นเวลาที่ผู้วิจัยบันทึกได้

กล่าวคือ $Y_i^* = \begin{cases} Y_i & , Y_i \le Y_c \\ Y_c & , Y_i > Y_c \end{cases}$
(ถ้าเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 มีจำนวน m ค่า คือ i =1,2,..., m)
(ถ้าเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 มีจำนวน n - m ค่า คือ i = m +1,..., n)

ต่อมาให้

$$\mathbf{X}^{\mathrm{NC}} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1m} & X_{2m} & \cdots & X_{pm} \end{bmatrix}_{m\times(p+1)}$$

$$\mathbf{X}^{\mathrm{C}} = \begin{bmatrix} 1 & X_{1(m+1)} & X_{2(m+1)} & \cdots & X_{p(m+1)} \\ 1 & X_{1(m+2)} & X_{2(m+2)} & \cdots & X_{p(m+2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix}_{(n-m)\times(p+1)}$$

ดังนั้น

$$\mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1m} & X_{2m} & \cdots & X_{pm} \\ \hdashline 1 & X_{1(m+1)} & X_{2(m+1)} & & X_{p(m+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix}_{n\times(p+1)} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} \mathbf{X}^{\mathrm{NC}} \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} \mathbf{X}^{\mathrm{C}} \end{matrix}$$

และให้

$$\mathbf{Y}^{\mathrm{NC}*} = \begin{bmatrix} Y_1^* \\ Y_2^* \\ \vdots \\ Y_m^* \end{bmatrix}_{m\times1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix}_{m\times1} \qquad \mathbf{Y}^{\mathrm{C}*} = \begin{bmatrix} Y_{m+1}^* \\ Y_{m+2}^* \\ \vdots \\ Y_n^* \end{bmatrix}_{(n-m)\times1} = \begin{bmatrix} Y_c \\ Y_c \\ \vdots \\ Y_c \end{bmatrix}_{(n-m)\times1}$$

ดังนั้น

$$\mathbf{Y}^{*}=\begin{bmatrix} Y_1^{*} \\ \vdots \\ Y_m^{*} \\ \hdashline Y_{m+1}^{*} \\ \vdots \\ Y_n^{*} \end{bmatrix}_{n\times 1}=\begin{bmatrix} Y_1 \\ \vdots \\ Y_m \\ \hdashline Y_c \\ \vdots \\ Y_c \end{bmatrix}_{n\times 1}\begin{matrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \mathbf{Y}^{\mathrm{NC}*} \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \mathbf{Y}^{\mathrm{C}*} \end{matrix}$$

และเวกเตอร์ของข้อมูลที่มีความสมบูรณ์ คือ $\tilde{\mathbf{Y}}=\begin{bmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \\ \vdots \\ \tilde{Y}_n \end{bmatrix}$

2.4.1 วิธีกำลังสองต่ำสุด (Ordinary Least Squares Method)

ถูกคิดค้นโดย Gauss (1809) โดยมีหลักการ คือ ทำให้ผลรวมของกำลังสองของเศษเหลือมีค่าต่ำที่สุด และถ้าข้อมูลที่นำมาวิเคราะห์เป็นไปตามเงื่อนไขของการวิเคราะห์การถดถอยแล้ว ตัวประมาณที่ได้จากวิธีนี้จะเป็นตัวประมาณที่ไม่เอนเอียงและมีความแปรปรวนต่ำที่สุด

<u>**การประมาณด้วยวิธีกำลังสองต่ำสุด**</u>

เริ่มต้นด้วยการหาผลรวมของกำลังสองของเศษเหลือในรูปของเมทริกซ์ จาก $\mathbf{e}=\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}$ จะได้ว่า

$$\sum_{i=1}^{n}\mathbf{e}^2=\left(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)^{\mathrm{T}}\left(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)$$

$$=\mathbf{Y}^{\mathrm{T}}\mathbf{Y}-\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y}-\mathbf{Y}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}+\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}$$

$$=\mathbf{Y}^{\mathrm{T}}\mathbf{Y}-2\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y}+\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}$$

และเนื่องจากต้องการหาค่า $\hat{\boldsymbol{\beta}}$ ที่ให้ผลรวมของกำลังสองของเศษเหลือมีค่าต่ำที่สุด จะได้ว่า

$$\frac{\partial}{\partial \hat{\boldsymbol{\beta}}}\left(\mathbf{Y}^{\mathrm{T}}\mathbf{Y}-2\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y}+\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}\right)=0$$

$$-2\mathbf{X}^{\mathrm{T}}\mathbf{Y}+2\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}=0$$

$$\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}=\mathbf{X}^{\mathrm{T}}\mathbf{Y}$$

$$\hat{\boldsymbol{\beta}}=\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$$

เนื่องจากข้อมูลที่นำมาวิเคราะห์ คือ ตัวแปรตามที่ประกอบไปด้วยข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 และข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 แต่เมื่อนำข้อมูลมาวิเคราะห์จะถือว่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 จะเสมือนเป็นตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ดังนั้นตัวประมาณที่ได้จะเป็นตัวประมาณที่เอนเอียง (ค่าคาดหวังของตัวประมาณกำลังสองต่ำสุดจะมีค่าไม่เท่ากับค่าพารามิเตอร์จริง) โดยค่าประมาณที่ได้จะต่ำกว่าค่าพารามิเตอร์ที่แท้จริง

2.4.2 วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation)

ธีระพร วีระถาวร (2536) กล่าวว่า ตัวประมาณภาวะน่าจะเป็นสูงสุด คือ ตัวประมาณที่ทำให้ความน่าจะเป็นในการเกิด $\boldsymbol{\theta}$ เกิดขึ้นสูงสุดเมื่อสุ่มตัวอย่าง $\mathbf{X}$ ขึ้นมา กล่าวคือ ตัวประมาณภาวะน่าจะเป็นสูงสุด คือ ตัวประมาณที่ทำให้ $\mathbf{L}(\boldsymbol{\theta};\mathbf{X})$ มีค่าสูงสุดหรือ $\log(\mathbf{L}(\boldsymbol{\theta};\mathbf{X}))$ มีค่ามากที่สุด

โดย $\mathbf{L}(\boldsymbol{\theta};\mathbf{X})$ คือ ฟังก์ชันภาวะน่าจะเป็นรวมซึ่งเป็นตัวที่บ่งบอกถึงความน่าจะเป็นที่ $\boldsymbol{\theta}$ จะเกิดขึ้นเมื่อสุ่มตัวอย่าง $\mathbf{X}$ ขึ้นมา

$\boldsymbol{\theta}=\begin{bmatrix}\theta_1\\ \theta_2\\ \vdots\\ \theta_p\end{bmatrix}_{p\times 1}$ คือ เวกเตอร์ของพารามิเตอร์

$\mathbf{X}=\begin{bmatrix}X_1\\ X_2\\ \vdots\\ X_n\end{bmatrix}_{n\times 1}$ คือ เวกเตอร์ของตัวอย่าง

และ $\mathbf{p}(\mathbf{X};\boldsymbol{\theta})$ คือ ฟังก์ชันความหนาแน่นรวมของ $\boldsymbol{\theta}$ เมื่อ $\mathbf{L}(\boldsymbol{\theta};\mathbf{X})=\mathbf{p}(\mathbf{X};\boldsymbol{\theta})$

เมื่อตัวอย่างของข้อมูลที่สมบูรณ์มีขนาดใหญ่ จะทำให้ตัวประมาณภาวะน่าจะเป็นสูงสุดไม่เอนเอียง (ค่าคาดหวังของตัวประมาณภาวะน่าจะเป็นสูงสุดจะมีค่าเท่ากับค่าพารามิเตอร์) และรูปร่างการกระจายของข้อมูลจะคล้ายกับการแจกแจงแบบปกติ

การประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุด

Jöreskog (2002) กล่าวถึงฟังก์ชันภาวะน่าจะเป็นของตัวแปรตามที่ประกอบไปด้วยข้อมูลที่ไม่ถูกตัดปลายทางขวาแบบที่ 1 และข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1 ดังนี้

$$L(Y_i^*) = \begin{cases} f(Y_i) \quad ,Y_i \le Y_c & \text{(ถ้าเป็นตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1)} \\ P(Y_i > Y_c) = S(Y_c), Y_i > Y_c & \text{(ถ้าเป็นตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1)} \end{cases}$$

ดังนั้นฟังก์ชันภาวะน่าจะเป็นรวม $\mathbf{L(\theta;X)}$

$$= \prod_{i=1}^{m} f(Y_i\ ) \cdot \prod_{i=m+1}^{n} S(Y_c)$$

$$= \prod_{i=1}^{m} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{Y_i - \left(\beta_0 + \beta_1 X_{1i} + \ldots + \beta_p X_{pi}\right)}{\sigma}\right]^2} \right] \bullet \prod_{i=m+1}^{n} \left[1 - \Phi\left(\frac{Y_c - \left(\beta_0 + \beta_1 X_{1i} + \ldots + \beta_p X_{pi}\right)}{\sigma} \right) \right]$$

โดย $S(Y_c)$ คือ ความน่าจะเป็นที่ตัวแปรสุ่ม Y_i มีค่ามากกว่า Y_c

m คือ จำนวนตัวแปรตามที่เป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาแบบที่ 1

n-m คือ จำนวนตัวแปรตามที่เป็นข้อมูลที่ถูกตัดปลายทางขวาแบบที่ 1

Tobin (1958) ได้นำเสนอวิธีที่ทำให้ $\mathbf{L(\theta;X)}$ มีค่าสูงสุดโดยไม่ได้ใช้การหาอนุพันธ์อันดับหนึ่งของ $\mathbf{L(\theta;X)}$ เพื่อประมาณค่าของ $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, ..., \beta_p, \sigma)$ โดยตรง แต่จะทำการประมาณค่าของ $\boldsymbol{\theta}^* = \left(\beta_0^*, \beta_1^*, ..., \beta_p^*, \sigma^*\right)$ แทน โดยที่ $\beta_j^* = \frac{\beta_j}{\sigma}\ , \sigma^* = \frac{1}{\sigma}$ และ j = 0, 1, 2, ..., p ดังนั้นฟังก์ชันภาวะน่าจะเป็นรวม $\mathbf{L(\theta^*;X)}$

$$= \prod_{i=1}^{m} f(Y_i) \cdot \prod_{i=m+1}^{n} S(Y_c)$$

$$= \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}} \sigma^* e^{-\frac{1}{2}\left[Y_i\ \sigma^* - \left(\beta_0^* + \beta_1^* X_{1i} + \ldots + \beta_p^* X_{pi}\right)\right]^2} \bullet \prod_{i=m+1}^{n} \left[1 - \Phi\left(Y_c \sigma^* - \left(\beta_0^* + \beta_1^* X_{1i} + \ldots + \beta_p^* X_{pi}\right)\right)\right]$$

จากนั้นใช้วิธีกำลังสองต่ำสุดในการหาค่าประมาณพารามิเตอร์เริ่มต้น และใช้วิธีการของ Newton-Raphson ในการทำให้ $\mathbf{L}(\boldsymbol{\theta}^*;\mathbf{X})$ มีค่าสูงสุดหรือทำให้ $-\log(\mathbf{L}(\boldsymbol{\theta}^*;\mathbf{X}))$ มีค่าต่ำสุด จนได้ค่าประมาณของ $\boldsymbol{\theta}^*$ ต่อมาหาค่าประมาณของ $\boldsymbol{\theta}$ ซึ่ง $\hat{\beta}_j = \dfrac{\hat{\beta}_j^*}{\hat{\sigma}^*}$ $,\hat{\sigma} = \dfrac{1}{\hat{\sigma}^*}$

2.4.3 วิธีของแชตเทอร์จีและแมคลีช (Chatterjee and McLeish Method)

Chatterjee และ McLeish (1986) นำหลักการ การแทนค่าที่สูญหายด้วยค่าที่ทำให้ผลรวมกำลังสองของความคลาดเคลื่อนมีค่าต่ำที่สุด มาประยุกต์กับการประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองต่ำสุดโดยใช้วิธีการวนซ้ำ มีขั้นตอนดังนี้

ขั้นตอน 0 ใช้เฉพาะข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 เพื่อคำนวณค่า $\hat{\boldsymbol{\beta}}^{(0)}$ ด้วยวิธีกำลังสองต่ำที่สุด ดังนั้น $\hat{\boldsymbol{\beta}}^{(0)} = \left(\left(\mathbf{X}^{NC}\right)^{\mathbf{T}}\mathbf{X}^{NC}\right)^{\mathbf{-1}}\left(\mathbf{X}^{NC}\right)^{\mathbf{T}}\mathbf{Y}^{NC*}$

ขั้นตอน a แทนค่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 โดย

- ถ้า $\hat{\beta}_0^{(r)} + \hat{\beta}_1^{(r)} X_{1i} + \ldots + \hat{\beta}_p^{(r)} X_{pi} > Y_c$ จะแทนค่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ด้วยค่า $\hat{\beta}_0^{(r)} + \hat{\beta}_1^{(r)} X_{1i} + \ldots + \hat{\beta}_p^{(r)} X_{pi}$
- ถ้า $\hat{\beta}_0^{(r)} + \hat{\beta}_1^{(r)} X_{1i} + \ldots + \hat{\beta}_p^{(r)} X_{pi} \le Y_c$ จะแทนค่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ด้วยค่า Y_c เหมือนเดิม

เมื่อ r คือ รอบที่ของการประมาณค่าพารามิเตอร์ , r = 0, 1, 2, 3, ...

r =0 คือ ขั้นตอน 0

และ $i = m+1, m+2, \ldots, n$

ขั้นตอน b นำตัวแปรตามที่ได้จากการแทนค่าใน**ขั้นตอน a** มารวมกับตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 เพื่อคำนวณค่า $\hat{\boldsymbol{\beta}}^{(r+1)}$ โดยใช้วิธีกำลังสองต่ำสุด

ทำ**ขั้นตอน a และ b** วนไปเรื่อยๆ จนกระทั่งได้ $\left|\hat{\beta}_j^{(r+1)} - \hat{\beta}_j^{(r)}\right| \le 0.001$ ทุก j =0,1,2,..., p

2.4.4 วิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม (Maximum Likelihood Estimation using the EM algorithm)

ขั้นตอนวิธีอีเอ็มถูกคิดค้นโดย Dempster, Laird และ Rubin (1977) และในเวลาต่อมานั้น Aitkin (1981) นำขั้นตอนวิธีอีเอ็มมาประยุกต์กับการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดโดยใช้วิธีการวนซ้ำ ซึ่งมีขั้นตอนดังนี้

ขั้นตอน 0

ใช้ข้อมูลทั้งหมดเพื่อคำนวณค่า $\hat{\boldsymbol{\beta}}^{(0)}$ และ $\hat{\sigma}^{(0)}$ ด้วยวิธีภาวะน่าจะเป็นสูงสุด ซึ่งถือว่าข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 เสมือนเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 นั่นคือ ทำให้ฟังก์ชันภาวะน่าจะเป็นรวม $L = \prod_{i=1}^{n} f(\mathrm{Y}_i^*)$ มีค่าสูงสุด ซึ่งมีค่าเท่ากับการคำนวณค่า $\hat{\boldsymbol{\beta}}^{(0)}$ และ $\hat{\sigma}^{(0)}$ ด้วยวิธีกำลังสองต่ำที่สุด (เนื่องจากการคำนวณค่า $\hat{\boldsymbol{\beta}}^{(0)}$ และ $\hat{\sigma}^{(0)}$ ด้วยวิธีกำลังสองต่ำสุดของข้อมูลที่สมบูรณ์ที่มีการแจกแจงแบบปกติ จะเท่ากับการคำนวณค่า $\hat{\boldsymbol{\beta}}^{(0)}$ และ $\hat{\sigma}^{(0)}$ ด้วยวิธีภาวะน่าจะเป็นสูงสุด) ดังนั้น $\hat{\boldsymbol{\beta}}^{(0)} = \left((\mathbf{X})^{\mathrm{T}}\mathbf{X}\right)^{-1}(\mathbf{X})^{\mathrm{T}}\mathbf{Y}^*$

ขั้นตอน E (Expectation Step : E Step)

การแทนค่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ด้วยค่าคาดหวังแบบมีเงื่อนไข

$$E\left(Y_i^* \middle| Y_i^* > Y_c, \boldsymbol{\beta}, \sigma\right) = \hat{\mu}_i^{(r)} + \left(\hat{\sigma}^{(r)} \cdot h\left(\hat{z}_i^{(r)}\right)\right)$$

โดย $\hat{\mu}_i^{(r)} = \hat{\beta}_0^{(r)} + \hat{\beta}_1^{(r)} X_{1i} + \ldots + \hat{\beta}_p^{(r)} X_{pi}$

$$\hat{z}_i^{(r)} = \frac{\left(Y_c - \hat{\mu}_i^{(r)}\right)}{\hat{\sigma}^{(r)}}$$

$$f\left(\hat{z}_i^{(r)}\right) = \left(\frac{1}{\sqrt{2\pi}}\right) e^{\left(-\frac{\hat{z}_i^{2}}{2}\right)}$$

$$S\left(\hat{z}_i^{(r)}\right) = 1 - F\left(\hat{z}_i^{(r)}\right) = \int_{\hat{z}_i}^{\infty} f(t)\,dt$$

$$h\left(\hat{z}_i^{(r)}\right) = \frac{f\left(\hat{z}_i^{(r)}\right)}{S\left(\hat{z}_i^{(r)}\right)}$$

เมื่อ r คือ รอบที่ของการประมาณค่าพารามิเตอร์ , r = 0, 1, 2, 3, ...

r =0 คือ ขั้นตอน 0

และ $i = m+1,\ m+2,\ ...,\ n$

ขั้นตอน M (Maximization Step : M Step)

นำตัวแปรตามที่ได้จากการแทนค่าใน**ขั้นตอน** E มารวมกับตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 จะได้ข้อมูลที่มีความสมบูรณ์ $\left(\tilde{Y}_i\right)$, i = 1, 2, ..., n โดย

$$\tilde{Y}_i = \begin{cases} Y_i^* = Y_i & , Y_i \le Y_c \text{ (ถ้าเป็นตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 มีจำนวน } m \text{ ค่า)} \\ \hat{\mu}_i^{(r)} + \left(\hat{\sigma}^{(r)} \cdot h\left(\hat{z}_i^{(r)}\right)\right) & , Y_i > Y_c \text{ (ถ้าเป็นตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 มีจำนวน } n-m \text{ ค่า)} \end{cases}$$

จากนั้นนำข้อมูลสมบูรณ์ดังกล่าวมาประมาณค่า $\hat{\boldsymbol{\beta}}^{(r+1)}$ และ $\hat{\sigma}^{(r+1)}$ เพื่อทำให้ฟังก์ชันภาวะน่าจะเป็นรวม $L = \prod_{i=1}^{n} f(\tilde{Y}_i)$ มีค่าสูงสุด นั่นคือ หาค่า $\hat{\boldsymbol{\beta}}^{(r+1)}$ ด้วยวิธีกำลังสองต่ำที่สุด

จะได้ $\hat{\boldsymbol{\beta}}^{(r+1)} = \left((\mathbf{X})^{\mathrm{T}}\mathbf{X}\right)^{-1}(\mathbf{X})^{\mathrm{T}}\tilde{\mathbf{Y}}$

และ $$\hat{\sigma}^{(r+1)} = \sqrt{\frac{\sum_{i=1}^{m}\left(y_i - \hat{\mu}_i^{(r)}\right)^2 + \left[\hat{\sigma}^2\right]^{(r)} \sum_{m+1}^{n}\left[1 + \left(\hat{z}_i^{(r)}\ \mathrm{h}(z_i^{(r)})\right)\right]}{n}}$$

โดยที่ $\hat{\sigma}^{(0)}$ คือค่าประมาณของส่วนเบี่ยงเบนมาตรฐานในขั้นตอน 0

ทำ**ขั้นตอน E และ** M วนไปเรื่อยๆ จนกระทั่งได้ $\left|\hat{\beta}_j^{(r+1)} - \hat{\beta}_j^{(r)}\right| \le 0.001$ ทุก j =0,1,2,..., p

*****หมายเหตุ*****

วิธีของแชตเทอร์จีและแมคลีช กับวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม มีความแตกต่างดังต่อไปนี้

- **ขั้นตอน 0** วิธีของแชตเทอร์จีและแมคลีช คำนวณค่า $\hat{\boldsymbol{\beta}}^{(0)}$ โดยใช้เฉพาะข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ส่วนวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม ใช้ข้อมูลทั้งหมดเพื่อคำนวณค่า $\hat{\boldsymbol{\beta}}^{(0)}$ โดยถือว่าข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 เสมือนเป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1

- **ขั้นตอน a** วิธีของแชตเทอร์จีและแมคลีชมีการแทนค่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ด้วยค่า $\hat{\beta}_0^{(r)} + \hat{\beta}_1^{(r)} X_{1i} + \ldots + \hat{\beta}_p^{(r)} X_{pi}$ เมื่อ $\hat{\beta}_0^{(r)} + \hat{\beta}_1^{(r)} X_{1i} + \ldots + \hat{\beta}_p^{(r)} X_{pi} > Y_c$ และแทนค่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ด้วยค่า Y_c เหมือนเดิม เมื่อ $\hat{\beta}_0^{(r)} + \hat{\beta}_1^{(r)} X_{1i} + \ldots + \hat{\beta}_p^{(r)} X_{pi} \leq Y_c$

ส่วน**ขั้นตอน E** ของวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม มีการแทนค่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ด้วยค่าคาดหวังแบบมีเงื่อนไข $E\left(Y_i^* \middle| Y_i^* > Y_c, \boldsymbol{\beta}, \sigma\right)$

- ขั้นตอน E วิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม คำนวณทั้งค่า $\hat{\sigma}^{(r+1)}$ และ $\hat{\boldsymbol{\beta}}^{(r+1)}$ แต่**ขั้นตอน b** วิธีของแชตเทอร์จีและแมคลีชคำนวณแค่ค่า $\hat{\boldsymbol{\beta}}^{(r+1)}$ เท่านั้น

.

บทที่ 3
วิธีดำเนินการวิจัย

3.1 ขอบเขตของการวิจัย

เนื้อหาในการวิจัยครั้งนี้ได้จำลองข้อมูลเพื่อใช้เปรียบเทียบวิธีการประมาณค่าจากตัวแบบการถดถอยจำนวน 4 วิธี ได้แก่ วิธีกำลังสองต่ำสุด วิธีภาวะน่าจะเป็นสูงสุด วิธีของแชตเทอร์จีและแมคลีช และวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม สำหรับสัญลักษณ์ที่ใช้ในงานวิจัยนี้ ได้แก่

$\sigma_{x_1}^2$	คือ ความแปรปรวนของตัวแปรอิสระตัวที่ 1
$\sigma_{x_2}^2$	คือ ความแปรปรวนของตัวแปรอิสระตัวที่ 2
$\sigma_{x_1+x_2}^2$	คือ ความแปรปรวนรวมของตัวแปรอิสระตัวที่ 1 และ 2
σ_{ε}^2	คือ ความแปรปรวนของความคลาดเคลื่อน
$\sigma_{x_1}^2 : \sigma_{x_2}^2$	คือ อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2
$\sigma_{x_1+x_2}^2 : \sigma_{\varepsilon}^2$	คือ อัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อน
n	คือ ขนาดตัวอย่าง โดย n = 30, 50, 100
r	คือ เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา โดย r = 10%, 20%, 30%
OLS	คือ วิธีกำลังสองต่ำสุด
MLE	คือ วิธีภาวะน่าจะเป็นสูงสุด
CM	คือ วิธีของแชตเทอร์จีและแมคลีช
MLE_EM	คือ วิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม

3.1.1 ตัวแบบของค่าสังเกตของตัวแปรตามที่ใช้ในการศึกษาครั้งนี้

$$Y_i = e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\varepsilon_i} \text{ หรือ } \log(Y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

เมื่อ i = 1, 2, ..., n และตัวแปรอิสระแต่ละตัวไม่มีความสัมพันธ์กัน

เมื่อ Y_i คือ ค่าสังเกตของตัวแปรตาม ลำดับที่ i

X_{1i} คือ ตัวแปรอิสระตัวที่ 1 ลำดับที่ i เมื่อ $X_{1i} \overset{iid}{\sim} N(0,\sigma^2_{x_1})$

X_{2i} คือ ตัวแปรอิสระตัวที่ 2 ลำดับที่ i เมื่อ $X_{2i} \overset{iid}{\sim} N(0,\sigma^2_{x_2})$

ε_i คือ ความคลาดเคลื่อนลำดับที่ i เมื่อ $\varepsilon_i \overset{iid}{\sim} N(0,\sigma^2_{\varepsilon})$

β_p คือ ค่าที่แท้จริงของค่าสัมประสิทธิ์การถดถอยตัวที่ p โดย p =1, 2

β_0 คือ ระยะตัดแกน y ที่แท้จริง

3.1.2 เงื่อนไขที่ทำการศึกษา

3.1.2.1 n = 30, 50, 100; r = 10%, 20%, 30%; $\beta_0 = 0.3, \beta_1 = 1, \beta_2 = 1$

3.1.2.2 เงื่อนไขความแปรปรวนของตัวแปรอิสระและความคลาดเคลื่อนเป็นดังนี้

ตารางที่ 3. 1 $\sigma^2_{x_1} : \sigma^2_{x_2}$

$\sigma^2_{x_1} : \sigma^2_{x_2}$	$\sigma^2_{x_1}$	$\sigma^2_{x_2}$
1:5	0.015	0.075
1:2	0.030	0.060
1:1	0.045	0.045

ตารางที่ 3. 2 $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	$\sigma^2_{x_1+x_2}$	σ^2_{ε}
2:1	0.090	0.045
1:1	0.090	0.090
1:2	0.090	0.180

หรือเขียนเป็นตารางสรุปกรณีที่ต้องการศึกษา ดังนี้

ตารางที่ 3. 3 $\sigma^2_{x_1}, \sigma^2_{x_2}$ และ σ^2_{ε} ที่ต้องการศึกษา

กรณี	$\sigma^2_{x_1}$	$\sigma^2_{x_2}$	σ^2_{ε}
1	0.015	0.075	0.045
2	0.030	0.060	0.045
3	0.045	0.045	0.045
4	0.015	0.075	0.090
5	0.030	0.060	0.090
6	0.045	0.045	0.090
7	0.015	0.075	0.180
8	0.030	0.060	0.180
9	0.045	0.045	0.180

3.1.2.3 ใช้ R version 3.4.4 ในการเปรียบเทียบวิธีการประมาณค่าจากตัวแบบการถดถอยจำนวน 4 วิธี วิธีละ 81 สถานการณ์ และสถานการณ์ละ 10,000 รอบ

3.2 วิธีการดำเนินการวิจัย

3.2.1 การจำลองข้อมูล

3.2.1.1 สร้างตัวแปรอิสระ $X_{1i} \overset{iid}{\sim} N(0,\sigma^2_{x_1}),\ X_{2i} \overset{iid}{\sim} N(0,\sigma^2_{x_2})$ โดยที่ตัวแปรอิสระทั้งสองตัวไม่มีความสัมพันธ์กันและสร้างความคลาดเคลื่อน $\varepsilon_i \overset{iid}{\sim} N(0,\sigma^2_{\varepsilon})$ ซึ่งความแปรปรวนของตัวแปรอิสระทั้ง 2 ตัวและความแปรปรวนของความคลาดเคลื่อนใช้เงื่อนไขตามตารางสรุปกรณีที่ต้องการศึกษาจำนวน 9 กรณี และขนาดตัวอย่าง (n) มี 3 ขนาด ได้แก่ 30, 50, 100 และกำหนดให้ i =1, 2,..., n และ $\beta_0 = 0.3, \beta_1 = 1, \beta_2 = 1$

3.2.1.2 สร้างค่าจริงของตัวแปรตามที่มีความสัมพันธ์กับตัวแปรอิสระ คือ $Y_i' = e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}}$ ที่มีการแจกแจงแบบล็อกนอร์มอล

3.2.1.3 สร้างค่าสังเกตของตัวแปรตามที่มีความสัมพันธ์กับตัวแปรอิสระและความคลาดเคลื่อน คือ $Y_i = e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\varepsilon_i}$ ที่มีการแจกแจงแบบล็อกนอร์มอล

3.2.2 การสร้างตัวแปรตามที่ประกอบไปด้วยข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 และข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1

3.2.2.1 หาค่า Y_c ซึ่งเป็นค่าเปอร์เซ็นต์ไทล์ที่ 100-r ของตัวแปรสุ่มที่มีการแจกแจงเหมือนกับ Y_i เมื่อ r = 10%, 20%, 30%

3.2.2.2 แบ่ง Y_i ออกเป็น 2 ส่วน ได้แก่ Y_i ที่เป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ($Y_i \le Y_c$) และ Y_i เป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ($Y_i > Y_c$) จากนั้นสร้างชุดข้อมูลใหม่ของตัวแปรตาม $\left(Y_i^*\right)$ ที่ประกอบไปด้วยตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 และตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 มีเกณฑ์ในการสร้าง Y_i^* ดังนี้

- ถ้า Y_i เป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ($Y_i \le Y_c$) ให้ Y_i^* ยังมีค่าเท่ากับ Y_i เหมือนเดิม
- ถ้า Y_i เป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ($Y_i > Y_c$) ให้ Y_i^* มีค่าเท่ากับ Y_c

<u>**ตัวอย่างการสร้างตัวแปรตามที่ประกอบไปด้วยข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 และข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1**</u>

1) กำหนด n =10, r =20% และ $\beta_0 = 0.3, \beta_1 = 1, \beta_2 = 1$
2) สร้าง $X_{1i} \overset{iid}{\sim} N(0, 0.015), X_{2i} \overset{iid}{\sim} N(0, 0.075), \varepsilon_i \overset{iid}{\sim} N(0, 0.045)$
3) สร้าง $Y_i^{'} = e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}}$ และ $Y_i = e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\varepsilon_i}$
4) หาค่า Y_c ซึ่งเป็นค่าเปอร์เซ็นต์ไทล์ที่ 100-r = 100-20 = 80 ของตัวแปรสุ่มที่แจกแจงเหมือนกับ Y_i นั่นคือตัวแปรสุ่มที่มีการแจกแจง $LOGNOR(\mu = 0.3, \sigma^2 = 0.135)$ ดังนั้น Y_c = 1.839
5) แบ่ง Y_i ออกเป็น 2 ส่วน ได้แก่ Y_i ที่เป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ($Y_i \le Y_c$) ซึ่งได้แก่ $Y_1, Y_2, Y_4, Y_5, Y_6, Y_9, Y_{10}$ และ Y_i เป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ($Y_i > Y_c$) ซึ่งได้แก่ Y_3, Y_7, Y_8
6) สร้างชุดข้อมูลใหม่ของตัวแปรตาม $\left(Y_i^*\right)$ ที่ประกอบไปด้วยตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 และตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 มีเกณฑ์การสร้างดังนี้
   - ถ้า Y_i เป็นข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ซึ่งได้แก่ $Y_1, Y_2, Y_4, Y_5, Y_6, Y_9, Y_{10}$ จะกำหนดให้ Y_i^* เมื่อ i = 1, 2, 4, 5, 6, 9, 10 ยังมีค่าเท่ากับ Y_i ตัวเดิม หรือกล่าวว่า $Y_1^* = Y_1, Y_2^* = Y_2, Y_4^* = Y_4, Y_5^* = Y_5, Y_6^* = Y_6, Y_9^* = Y_9, Y_{10}^* = Y_{10}$

- ถ้า Y_i เป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 ซึ่งได้แก่ Y_3, Y_7, Y_8 จะกำหนดให้ค่า Y_i^* เมื่อ i =3, 7, 8 มีค่าเท่ากับ $Y_c = 1.839$ หรือกล่าวว่า $Y_3^* = Y_7^* = Y_8^* = 1.839$

ตารางที่ 3. 4 การสร้างตัวแปรตามที่ประกอบไปด้วยข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 และข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1

i	X_{1i}	X_{2i}	ε_i	Y_i'	Y_i	Y_i^*
1	-0.082	-0.416	-0.338	0.820	0.585	0.585
2	0.051	-0.427	-0.095	0.927	0.843	0.843
3	-0.245	0.191	0.434	1.279	1.974	1.839
4	-0.263	0.115	-0.141	1.164	1.011	1.011
5	-0.049	0.118	-0.053	1.447	1.372	1.372
6	-0.206	0.129	0.050	1.250	1.314	1.314
7	0.045	0.664	0.376	2.743	3.995	1.839
8	0.063	0.358	0.010	2.057	2.077	1.839
9	0.019	0.239	-0.001	1.747	1.745	1.745
10	-0.013	-0.332	-0.193	0.956	0.788	0.788

3.2.3 การประมาณค่าจากตัวแบบการถดถอย

หาค่า $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ ของ $\log(Y_i^*)$ โดยใช้วิธีกำลังสองต่ำสุด วิธีภาวะน่าจะเป็นสูงสุด วิธีของแชตเทอร์จีและแมคลีช และวิธีภาวะน่าจะเป็นสูงสุดด้วยขั้นตอนวิธีอีเอ็ม ต่อมาหาค่าประมาณของตัวแปรตาม ซึ่งคือ $\hat{Y}_i = e^{\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}}$ ของทั้ง 4 วิธี

3.2.4 การเปรียบเทียบวิธีการประมาณค่าจากตัวแบบการถดถอย

3.2.4.1 ค่าประมาณของตัวแปรตาม ใช้ค่า $AMSE(\hat{Y}_a)$ และค่า $RE\left(\hat{Y}_a\right)$ เป็นเกณฑ์เปรียบเทียบมีรายละเอียดดังต่อไปนี้

1) ค่า $AMSE\left(\hat{Y}_a\right)$ เป็นค่าเฉลี่ยของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนของตัวแปรตามของวิธี a จากการจำลองสถานการณ์ 10,000 รอบ โดยที่

$$AMSE\left(\hat{Y}_a\right) = \frac{\sum_{j=1}^{10000} MSE\left(\hat{Y}_a^{(j)}\right)}{10,000}$$

เมื่อ $MSE\left(\hat{Y}_a^{(j)}\right)$ คือ ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนของตัวแปรตาม จากการจำลองสถานการณ์รอบที่ j ของวิธี a โดยที่

$$MSE\left(\hat{Y}_a^{(j)}\right)=\frac{\sum_{i=1}^{n}\left(\hat{Y}_{i,a}-Y_i^{'}\right)^2}{n}$$

$\hat{Y}_{i,a}$ คือ ค่าประมาณของตัวแปรตามลำดับที่ i โดยวิธีการ a

$Y_i^{'}$ คือ ค่าที่แท้จริงของตัวแปรตามลำดับที่ i

j คือ รอบที่ของการจำลองสถานการณ์ โดย j = 1, 2, ..., 10000

a คือ วิธีประมาณค่าที่ใช้ในการศึกษา ซึ่ง a =OLS,CM,MLE,MLE_EM

<u>เกณฑ์ที่ใช้ในการเปรียบเทียบ</u>

วิธีที่มีค่า $AMSE\left(\hat{Y}_a\right)$ ต่ำสุดจะเป็นวิธีประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงสุด

2) $RE\left(\hat{Y}_a\right)$ เป็นค่าประสิทธิภาพสัมพัทธ์ของค่า $AMSE\left(\hat{Y}_a\right)$ โดยเป็นอัตราส่วนระหว่างค่า $AMSE\left(\hat{Y}_{MLE}\right)$ กับค่า $AMSE\left(\hat{Y}_a\right)$ หรือกล่าวว่าค่า $RE\left(\hat{Y}_a\right)$ เป็นอัตราส่วนระหว่างค่า $AMSE\left(\hat{Y}\right)$ ของวิธี MLE กับค่า $AMSE\left(\hat{Y}\right)$ ของวิธี a

การที่ใช้ค่า $RE\left(\hat{Y}_a\right)$ เนื่องจากต้องการเปรียบเทียบประสิทธิภาพของแต่ละวิธีให้เกิดความชัดเจนมากขึ้น และการศึกษาครั้งนี้เลือกวิธี MLE เป็นเกณฑ์ในการเปรียบเทียบ เพราะว่าวิธี MLE เป็นวิธีการประมาณค่าพื้นฐานที่มีประสิทธิภาพสูงในการประมาณค่า และเป็นวิธีที่ใช้ข้อมูลได้ครบถ้วนทั้งส่วนที่เป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 และข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 รวมถึงเป็นตัวประมาณที่ยืนยงภายใต้การเปลี่ยนแปลงที่เหมาะสม โดยที่

$$RE\left(\hat{Y}_a\right)=\frac{AMSE\left(\hat{Y}_{\text{MLE}}\right)}{AMSE\left(\hat{Y}_a\right)}$$

<u>**เกณฑ์ที่ใช้ในการเปรียบเทียบ**</u>

- วิธีที่มีค่า $RE\left(\hat{Y}_a\right)$ <u>มากกว่า 1</u> แสดงว่าวิธีนั้นมีประสิทธิภาพ<u>มากกว่า</u>วิธี MLE
- วิธีที่มีค่า $RE\left(\hat{Y}_a\right)$ <u>น้อยกว่า 1</u> แสดงว่าวิธีนั้นมีประสิทธิภาพ<u>ด้อยกว่า</u> วิธี MLE
- วิธีที่มีค่า $RE\left(\hat{Y}_a\right)$ <u>เท่ากับ 1</u> แสดงว่าวิธีนั้นมีประสิทธิภาพ<u>เทียบเท่า</u>กับวิธี MLE

3.2.4.2 ค่าประมาณของค่า $\beta_0, \beta_1, \beta_2$ ใช้ค่า $MSE\left(\hat{\beta}_{p,a}\right)$ และค่า $RE\left(\hat{\beta}_{p,a}\right)$ เป็นเกณฑ์ในการเปรียบเทียบ มีรายละเอียดดังต่อไปนี้

1) ค่า $MSE\left(\hat{\beta}_{p,a}\right)$ เป็นค่าเฉลี่ยของกำลังสองของความคลาดเคลื่อนของค่าสัมประสิทธิ์การถดถอยในตัวแบบการถดถอยตัวที่ p ของวิธี a จากการจำลองสถานการณ์ 10,000 รอบ เมื่อ

$$MSE\left(\hat{\beta}_{p,a}\right) = \frac{\sum_{j=1}^{10,000}\left(\hat{\beta}_{p,a}^{(j)} - \beta_p\right)^2}{10,000} \qquad ; p = 0, 1, 2$$

<u>**เกณฑ์ที่ใช้ในการเปรียบเทียบ**</u>

วิธีที่มีค่า $MSE\left(\hat{\beta}_{p,a}\right)$ ต่ำสุดจะเป็นวิธีประมาณค่าสัมประสิทธิ์การถดถอยที่มีประสิทธิภาพสูงสุด

2) ค่า $RE\left(\hat{\beta}_{p,a}\right)$ ซึ่งเป็นค่าประสิทธิภาพสัมพัทธ์ของค่า $MSE\left(\hat{\beta}_{p,a}\right)$ โดยเป็นอัตราส่วนระหว่าง ค่า $MSE\left(\hat{\beta}_{p,a}\right)$ ของวิธี MLE กับค่า $MSE\left(\hat{\beta}_{p,a}\right)$ ของวิธี a โดยที่

$$RE\left(\hat{\beta}_{p,a}\right) = \frac{MSE\left(\hat{\beta}_{p,\mathrm{MLE}}\right)}{MSE\left(\hat{\beta}_{p,a}\right)} \qquad ; p = 0, 1, 2$$

เมื่อ β_p คือ ค่าสัมประสิทธิ์การถดถอยตัวที่ p

$\hat{\beta}_{p,a}^{(j)}$ คือ ค่าประมาณของค่าสัมประสิทธิ์การถดถอยตัวที่ p ในการจำลองรอบที่ j โดยวิธี a เมื่อ p = 0, 1, 2 (เพื่อความสะดวกในการอธิบายในเรื่องเกณฑ์ที่ใช้ในการเปรียบเทียบนี้ กำหนดให้ ค่าสัมประสิทธิ์การถดถอยตัวที่ 0 คือ ระยะตัดแกน y)

j คือ รอบที่ของการจำลองสถานการณ์ โดย j = 1, 2, ..., 10000

a คือ วิธีประมาณค่าที่ใช้ในการศึกษา ซึ่ง a = OLS, CM, MLE, MLE_EM

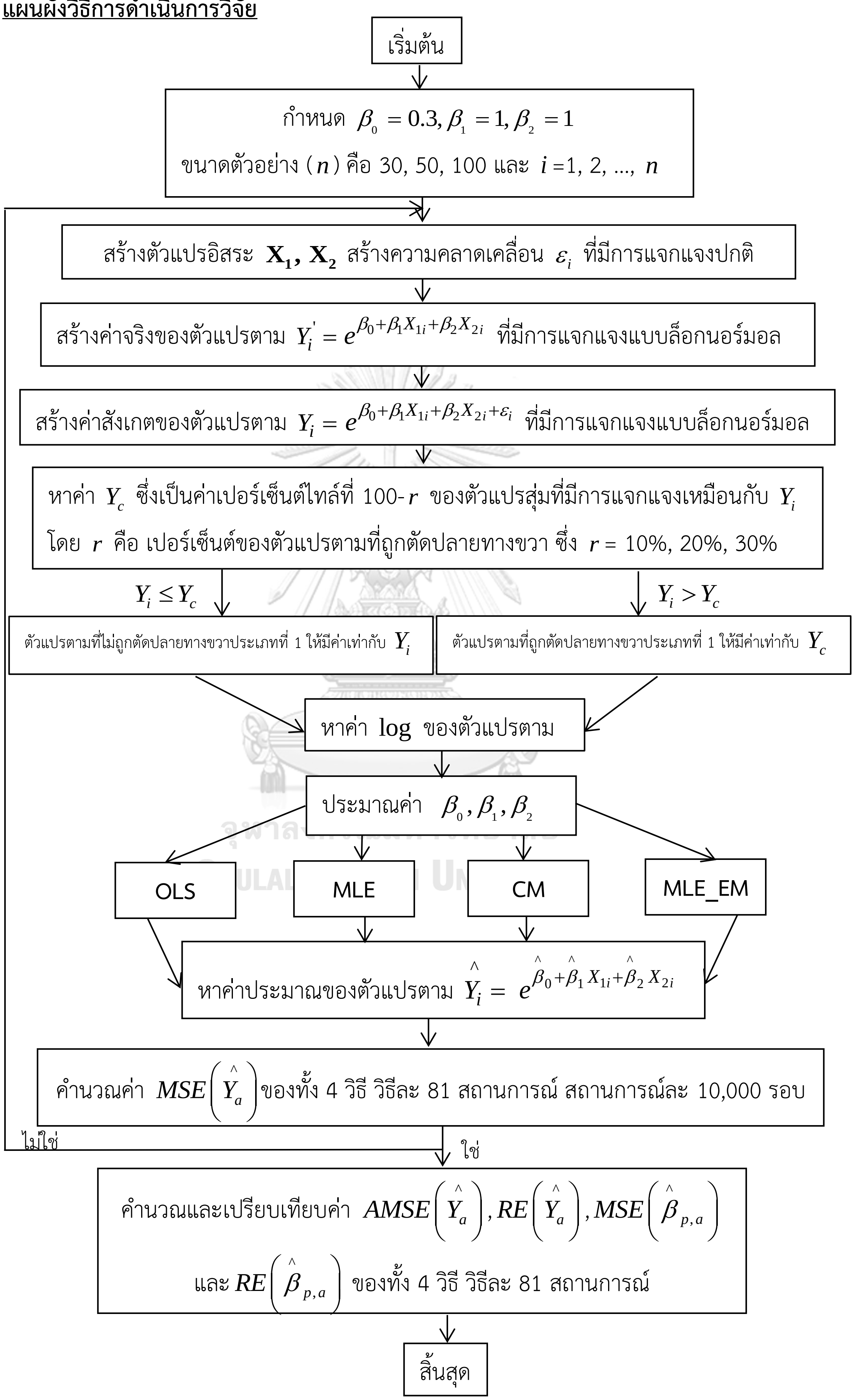
<u>เกณฑ์ที่ใช้ในการเปรียบเทียบ</u>

วิธีที่มีค่า $RE\left(\hat{\beta}_{p,a}\right)$ <u>มากกว่า 1</u> แสดงว่าวิธีนั้นมีประสิทธิภาพ<u>มากกว่า</u>วิธี MLE

วิธีที่มีค่า $RE\left(\hat{\beta}_{p,a}\right)$ <u>น้อยกว่า 1</u> แสดงว่าวิธีนั้นมีประสิทธิภาพ<u>ด้อยกว่า</u>วิธี MLE

วิธีที่มีค่า $RE\left(\hat{\beta}_{p,a}\right)$ <u>เท่ากับ 1</u> แสดงว่าวิธีนั้นมีประสิทธิภาพ<u>เทียบเท่า</u>กับวิธี MLE

แผนผังวิธีการดำเนินการวิจัย



บทที่ 4
ผลการวิจัย

สำหรับผลการวิจัยในครั้งนี้เป็นผลการเปรียบเทียบวิธีการประมาณค่าจากตัวแบบการถดถอย 4 วิธี ได้แก่ OLS, MLE, CM และ MLE_EM ซึ่งแบ่งผลการวิจัยเป็น 4 ส่วน ได้แก่

4.1 ผลการเปรียบเทียบวิธีการประมาณค่าตัวแปรตาม

4.2 ผลการเปรียบเทียบวิธีการประมาณค่า β_0

4.3 ผลการเปรียบเทียบวิธีการประมาณค่า β_1

4.4 ผลการเปรียบเทียบวิธีการประมาณค่า β_2

ก่อนจะเริ่มกล่าวถึงผลวิจัย มีข้อตกลงเบื้องต้น ดังนี้

- ความคลาดเคลื่อนกระจาย<u>มากกว่า</u>ตัวแปรอิสระ คือ $\sigma^2_{x_1+x_2} : \sigma^2_\varepsilon = 1:2$
- ความคลาดเคลื่อนกระจาย<u>น้อยกว่า</u>ตัวแปรอิสระ คือ $\sigma^2_{x_1+x_2} : \sigma^2_\varepsilon = 2:1$
- ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกัน คือ $\sigma^2_{x_1} : \sigma^2_{x_2} = 1:2, 1:5$
- ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกัน<u>น้อย</u> คือ $\sigma^2_{x_1} : \sigma^2_{x_2} = 1:2$
- ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกัน<u>มาก</u> คือ $\sigma^2_{x_1} : \sigma^2_{x_2} = 1:5$

ในการพิจารณาว่า วิธีไหนเป็นวิธีที่มีประสิทธิภาพสูงที่สุด จะพิจารณาจากค่า $RE\left(\hat{Y}_a\right)$, $RE\left(\hat{\beta}_{0,a}\right)$, $RE\left(\hat{\beta}_{1,a}\right)$ และ $RE\left(\hat{\beta}_{2,a}\right)$ เป็นหลัก โดยถ้าพิจารณาจากค่า $AMSE\left(\hat{Y}_a\right)$, $MSE\left(\hat{\beta}_{0,a}\right)$, $MSE\left(\hat{\beta}_{1,a}\right)$ และ $MSE\left(\hat{\beta}_{2,a}\right)$ ก็จะได้ผลสรุปที่เหมือนกัน

วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย คือ วิธีที่มีค่า RE สูงสุด หรือวิธีที่มีค่า RE น้อยกว่าค่า RE ที่มีค่ามากที่สุดในสถานการณ์เดียวกัน ไม่เกิน 0.01

วิธีที่มีประสิทธิภาพในการประมาณค่าจากตัวแบบการถดถอยด้อยกว่าวิธีที่มีประสิทธิภาพมากที่สุดเพียงเล็กน้อย คือ วิธีที่มีค่า RE น้อยกว่าค่า RE ที่มีค่ามากที่สุดในสถานการณ์เดียวกันอยู่ในช่วงที่มากกว่า 0.01 และไม่เกิน 0.05

4.1 ผลการเปรียบเทียบวิธีการประมาณค่าตัวแปรตาม

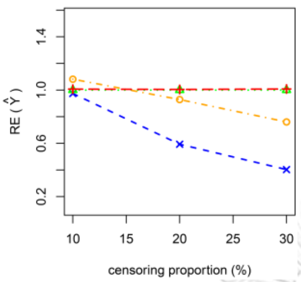
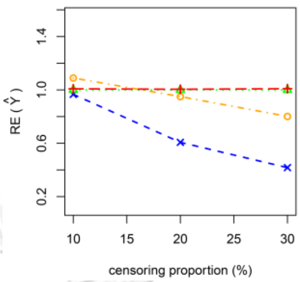
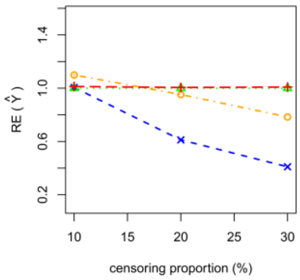
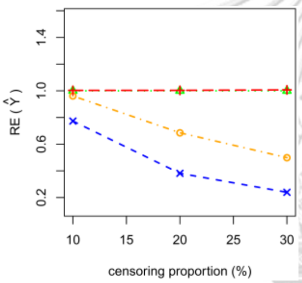
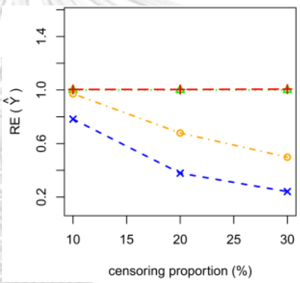
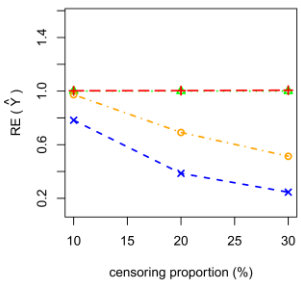
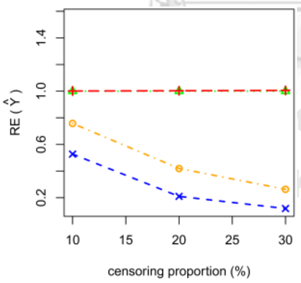
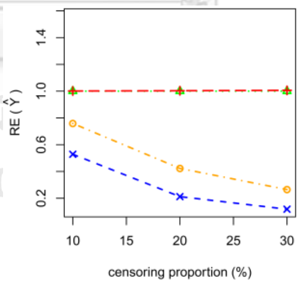
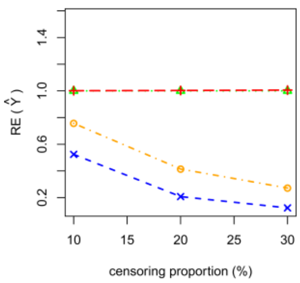
ตารางที่ 4. 1 ผลการเปรียบเทียบค่า $AMSE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 2:1$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$AMSE\left(\hat{Y}_a\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	0.01335	0.01342	**0.01221**	0.01325
			1:2	0.01366	0.01321	**0.01212**	0.01310
			1:1	0.01365	0.01331	**0.01230**	0.01321
		20%	1:5	0.02718	**0.01663**	0.01746	**0.01654**
			1:2	0.02712	**0.01647**	0.01733	**0.01639**
			1:1	0.02719	**0.01613**	0.01736	**0.01605**
		30%	1:5	0.05046	**0.02067**	0.02636	**0.02049**
			1:2	0.05097	**0.02128**	0.02655	**0.02109**
			1:1	0.05067	**0.02042**	0.02683	**0.02025**
	50	10%	1:5	0.00988	**0.00773**	0.00794	**0.00771**
			1:2	0.00985	**0.00771**	0.00794	**0.00768**
			1:1	0.01007	**0.00779**	0.00810	**0.00776**
		20%	1:5	0.02376	**0.00918**	0.01326	**0.00915**
			1:2	0.02399	**0.00907**	0.01337	**0.00904**
			1:1	0.02385	**0.00910**	0.01328	**0.00907**
		30%	1:5	0.04711	**0.01158**	0.02254	**0.01150**
			1:2	0.04715	**0.01136**	0.02281	**0.01129**
			1:1	0.04727	**0.01130**	0.02266	**0.01122**
	100	10%	1:5	0.00721	**0.00378**	0.00499	**0.00377**
			1:2	0.00724	**0.00383**	0.00504	**0.00382**
			1:1	0.00718	**0.00379**	0.00500	**0.00378**
		20%	1:5	0.02097	**0.00435**	0.01051	**0.00434**
			1:2	0.02089	**0.00443**	0.01048	**0.00442**
			1:1	0.02098	**0.00440**	0.01050	**0.00439**
		30%	1:5	0.04414	**0.00541**	0.01991	**0.00537**
			1:2	0.04469	**0.00528**	0.01992	**0.00524**
			1:1	0.04432	**0.00526**	0.02001	**0.00523**

ตารางที่ 4. 2 ผลการเปรียบเทียบค่า $RE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$RE\left(\hat{Y}_a\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	1.01	1.00	**1.10**	1.01
			1:2	0.97	1.00	**1.09**	1.01
			1:1	0.97	1.00	**1.08**	1.01
		20%	1:5	0.61	**1.00**	0.95	**1.01**
			1:2	0.61	**1.00**	0.95	**1.00**
			1:1	0.59	**1.00**	0.93	**1.00**
		30%	1:5	0.41	**1.00**	0.78	**1.01**
			1:2	0.42	**1.00**	0.80	**1.01**
			1:1	0.40	**1.00**	0.76	**1.01**
	50	10%	1:5	0.78	**1.00**	0.97	**1.00**
			1:2	0.78	**1.00**	0.97	**1.00**
			1:1	0.77	**1.00**	0.96	**1.00**
		20%	1:5	0.39	**1.00**	0.69	**1.00**
			1:2	0.38	**1.00**	0.68	**1.00**
			1:1	0.38	**1.00**	0.69	**1.00**
		30%	1:5	0.25	**1.00**	0.51	**1.01**
			1:2	0.24	**1.00**	0.50	**1.01**
			1:1	0.24	**1.00**	0.50	**1.01**
	100	10%	1:5	0.52	**1.00**	0.76	**1.00**
			1:2	0.53	**1.00**	0.76	**1.00**
			1:1	0.53	**1.00**	0.76	**1.00**
		20%	1:5	0.21	**1.00**	0.41	**1.00**
			1:2	0.21	**1.00**	0.42	**1.00**
			1:1	0.21	**1.00**	0.42	**1.00**
		30%	1:5	0.12	**1.00**	0.27	**1.01**
			1:2	0.12	**1.00**	0.26	**1.01**
			1:1	0.12	**1.00**	0.26	**1.01**

ตารางที่ 4. 3 กราฟเปรียบเทียบค่า $RE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 2:1$

$\sigma^2_{x_1} : \sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.1.1 ผลการเปรียบเทียบวิธีการประมาณค่าตัวแปรตามเมื่อพิจารณาจากค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$ กรณีที่ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

จากตารางที่ 4.1 - 4.3 พบว่า

วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่าตัวแปรตามที่มีประสิทธิภาพมากที่สุด ยกเว้นเมื่อตัวอย่างมีขนาดปานกลาง (n = 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี CM มีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี MLE และวิธี MLE_EM เล็กน้อย และเมื่อตัวอย่างมีขนาดเล็ก (n = 30) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี CM เป็นวิธีการประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงที่สุด

ปัจจัยที่มีผลต่อค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างใหญ่ขึ้น ทำให้ค่า $AMSE\left(\hat{Y}\right)$ ของทุกวิธีมีค่าลดลง นั่นคือ ทุกวิธีสามารถประมาณค่าของตัวแปรตามได้ใกล้เคียงกับค่าที่แท้จริงของตัวแปรตามมากขึ้น

เมื่อพิจารณาค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM และวิธี MLE พบว่ามี 2 กรณีเกิดขึ้นพอๆ กัน ได้แก่ กรณีแรก คือ ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ส่วนกรณีที่ 2 คือ $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM มีค่ามากกว่าวิธี MLE เท่ากับ 0.01 แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่าตัวแปรตามมากกว่าวิธี MLE แต่เนื่องจากต่างกันแค่เพียง 0.01 ซึ่งถือว่าแตกต่างกันน้อยมาก จึงสรุปได้ว่า วิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE และจากทั้งสองกรณี สามารถสรุปได้ว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าขนาดตัวอย่างจะเป็นเท่าไหร่ก็ตาม และเมื่อตัวอย่างมีขนาดใหญ่ขึ้น ค่า $RE\left(\hat{Y}\right)$ ของวิธี OLS และวิธี CM มีค่าลดลงมาก นั่นคือ เมื่อตัวอย่างใหญ่ขึ้น วิธี OLS และวิธี CM นั้นจะมีประสิทธิภาพในการประมาณค่าตัวแปรตามด้อยกว่าวิธี MLE เพิ่มขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ทำให้ค่า $AMSE\left(\hat{Y}\right)$ ของทุกวิธีมีค่ามากขึ้น นั่นคือ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น แล้วจะทำให้ทุกวิธีประมาณค่าตัวแปรตามได้ใกล้เคียงกับค่าที่แท้จริงของตัวแปรตามลดลง และเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น พบว่าค่า $AMSE\left(\hat{Y}\right)$ของวิธี OLS มีค่าเพิ่มขึ้นอย่างมาก รองลงมาคือวิธี CM นั่นคือ เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวามีผลต่อการประมาณค่าตัวแปรตามของวิธี OLS เป็นอย่างมาก

เมื่อพิจารณาค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM และวิธี MLE พบว่ามี 2 กรณีเกิดขึ้นพอๆ กัน ได้แก่ กรณีแรก คือ ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ส่วนกรณีที่ 2 คือ $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM มีค่ามากกว่าวิธี MLE เท่ากับ 0.01 แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่าตัวแปรตามมากกว่าวิธี MLE แต่เนื่องจากต่างกันแค่เพียง 0.01 ซึ่งถือว่าแตกต่างกันน้อยมาก จึงสรุปได้ว่า วิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE และจากทั้งสองกรณีจึงสรุปได้ว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าตัวแปรตามจะถูกตัดปลายทางขวาเป็นเท่าไหร่ก็ตาม และเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $RE\left(\hat{Y}\right)$ของวิธี OLS กับวิธี CM มีค่าลดลงมาก นั่นคือเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้นจะทำให้ประสิทธิภาพการประมาณค่าตัวแปรตามของวิธี OLS กับวิธี CM ลดลงมากขึ้นเมื่อเทียบกับวิธี MLE กล่าวคือ ยิ่งตัวแปรตามถูกตัดปลายทางขวามากขึ้น วิธี OLS กับวิธี CM จะประมาณค่าตัวแปรตามได้แย่กว่าวิธี MLE มากขึ้น

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

ถึงแม้ตัวแปรอิสระตัวที่ 1 จะมีการกระจายมากกว่าตัวแปรอิสระตัวที่ 2 มากขึ้น แต่ค่า $AMSE\left(\hat{Y}\right)$ ที่ขนาดตัวอย่างเดียวกัน ตัวแปรตามถูกตัดปลายทางขวาเท่ากันและอัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อนที่เท่ากัน มีค่าใกล้เคียงกันมาก เช่นเดียวกับค่า $RE\left(\hat{Y}\right)$ ดังนั้น อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่าตัวแปรตามน้อยมาก

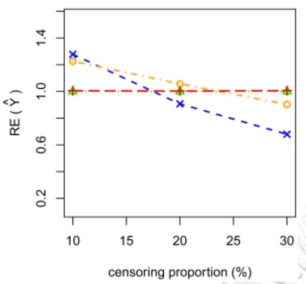
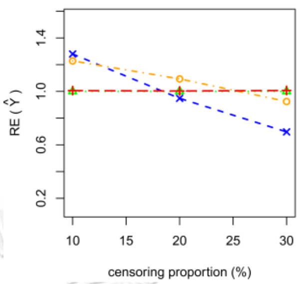
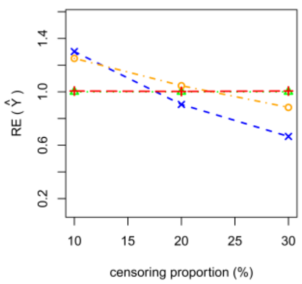
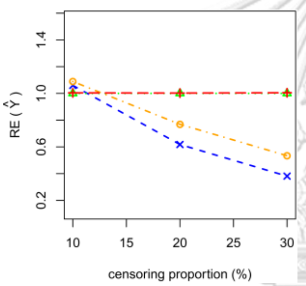
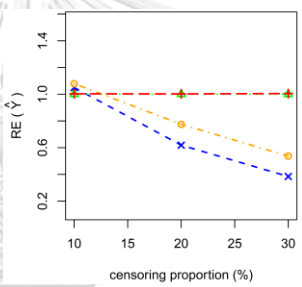
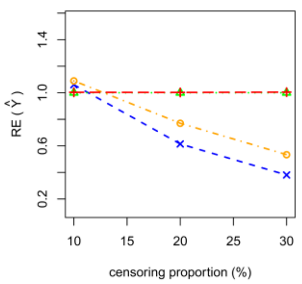
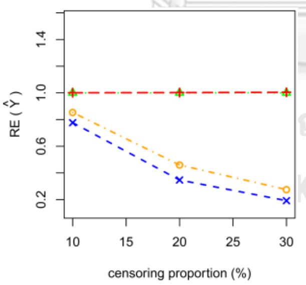
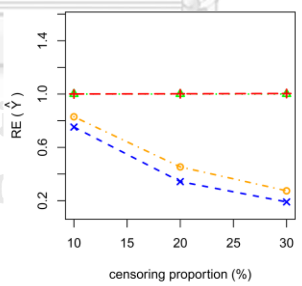
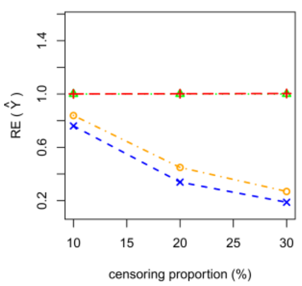
ตารางที่ 4. 4 ผลการเปรียบเทียบค่า $AMSE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$AMSE\left(\hat{Y}_a\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**0.02043**	0.02659	0.02126	0.02644
			1:2	**0.02034**	0.02604	0.02120	0.02588
			1:1	**0.02045**	0.02612	0.02130	0.02598
		20%	1:5	0.03334	0.03019	**0.02886**	0.03010
			1:2	0.03325	0.03152	**0.02883**	0.03143
			1:1	0.03321	0.03012	**0.02851**	0.03004
		30%	1:5	0.05799	**0.03861**	0.04369	**0.03837**
			1:2	0.05817	**0.04053**	0.04384	**0.04026**
			1:1	0.05749	**0.03906**	0.04328	**0.03882**
	50	10%	1:5	0.01425	0.01512	**0.01387**	0.01506
			1:2	0.01429	0.01501	**0.01391**	0.01496
			1:1	0.01418	0.01503	**0.01380**	0.01497
		20%	1:5	0.02823	**0.01734**	0.02252	**0.01730**
			1:2	0.02750	**0.01701**	0.02196	**0.01697**
			1:1	0.02812	**0.01738**	0.02260	**0.01733**
		30%	1:5	0.05282	**0.02011**	0.03767	**0.02001**
			1:2	0.05328	**0.02048**	0.03813	**0.02038**
			1:1	0.05347	**0.02038**	0.03817	**0.02028**
	100	10%	1:5	0.00970	**0.00738**	0.00879	**0.00737**
			1:2	0.00959	**0.00722**	0.00870	**0.00721**
			1:1	0.00944	**0.00733**	0.00859	**0.00732**
		20%	1:5	0.02418	**0.00819**	0.01823	**0.00817**
			1:2	0.02423	**0.00831**	0.01833	**0.00829**
			1:1	0.02399	**0.00831**	0.01810	**0.00830**
		30%	1:5	0.05017	**0.00942**	0.03499	**0.00938**
			1:2	0.05031	**0.00959**	0.03496	**0.00955**
			1:1	0.04990	**0.00957**	0.03485	**0.00953**

ตารางที่ 4. 5 ผลการเปรียบเทียบค่า $RE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:1$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$RE\left(\hat{Y}_a\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**1.30**	1.00	1.25	1.01
			1:2	**1.28**	1.00	1.23	1.01
			1:1	**1.28**	1.00	1.23	1.01
		20%	1:5	0.91	1.00	**1.05**	1.00
			1:2	0.95	1.00	**1.09**	1.00
			1:1	0.91	1.00	**1.06**	1.00
		30%	1:5	0.67	**1.00**	0.88	**1.01**
			1:2	0.70	**1.00**	0.92	**1.01**
			1:1	0.68	**1.00**	0.90	**1.01**
	50	10%	1:5	1.06	1.00	**1.09**	1.00
			1:2	1.05	1.00	**1.08**	1.00
			1:1	1.06	1.00	**1.09**	1.00
		20%	1:5	0.61	**1.00**	0.77	**1.00**
			1:2	0.62	**1.00**	0.77	**1.00**
			1:1	0.62	**1.00**	0.77	**1.00**
		30%	1:5	0.38	**1.00**	0.53	**1.00**
			1:2	0.38	**1.00**	0.54	**1.00**
			1:1	0.38	**1.00**	0.53	**1.00**
	100	10%	1:5	0.76	**1.00**	0.84	**1.00**
			1:2	0.75	**1.00**	0.83	**1.00**
			1:1	0.78	**1.00**	0.85	**1.00**
		20%	1:5	0.34	**1.00**	0.45	**1.00**
			1:2	0.34	**1.00**	0.45	**1.00**
			1:1	0.35	**1.00**	0.46	**1.00**
		30%	1:5	0.19	**1.00**	0.27	**1.00**
			1:2	0.19	**1.00**	0.27	**1.00**
			1:1	0.19	**1.00**	0.27	**1.00**

ตารางที่ 4. 6 กราฟเปรียบเทียบค่า $RE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:1$

$\sigma^2_{x_1} : \sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.1.2 ผลการเปรียบเทียบวิธีการประมาณค่าตัวแปรตามเมื่อพิจารณาจากค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$ กรณีที่ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:1$

จากตารางที่ 4.4 - 4.6 พบว่า

เมื่อตัวอย่างมีขนาดปานกลางถึงใหญ่ (n = 50, 100) ส่วนใหญ่ วิธี MLE และวิธี MLE_EM เป็นวิธีประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงที่สุด ยกเว้นเมื่อตัวอย่างมีขนาดปานกลาง (n = 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี CM เป็นวิธีประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงที่สุด และวิธี OLS มีประสิทธิภาพในการประมาณค่าตัวแปรตามที่ต่างจากวิธี CM เพียงเล็กน้อย

เมื่อขนาดตัวอย่างมีขนาดเล็ก (n = 30) วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงสุดกรณีที่ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) แต่ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) พบว่า วิธี CM เป็นวิธีที่มีประสิทธิภาพสูงที่สุดในการประมาณค่าตัวแปรตาม ส่วนในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตาม และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี OLS เพียงเล็กน้อย

ปัจจัยที่มีผลต่อค่า ค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$

ผลสรุปเหมือนกับตารางที่ 4.1 - 4.3 ยกเว้นปัจจัยด้านขนาดตัวอย่างและเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา ซึ่งพบว่า ส่วนใหญ่ ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE นั่นคือ วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าขนาดตัวอย่างหรือตัวแปรตามจะถูกตัดปลายทางขวาจะเป็นเท่าไหร่ก็ตาม

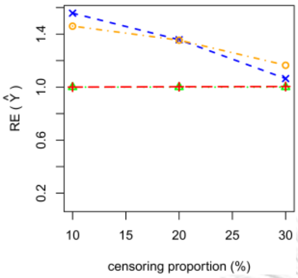
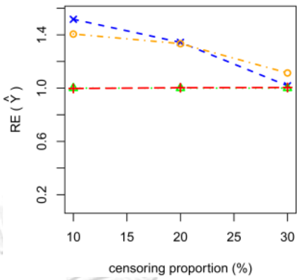
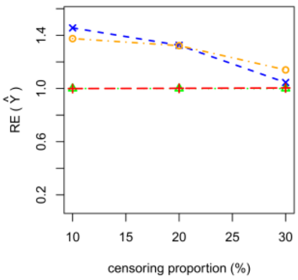
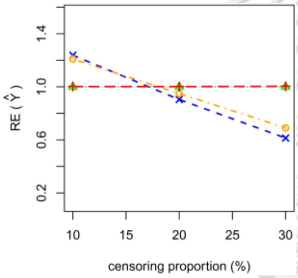
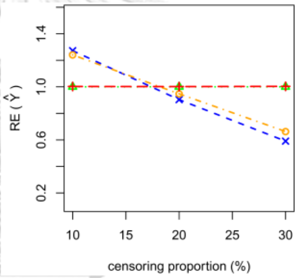
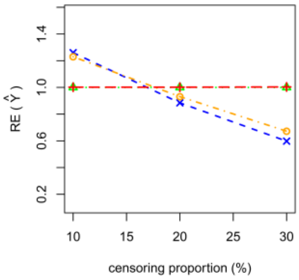
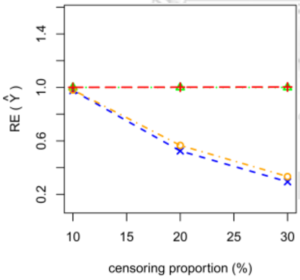
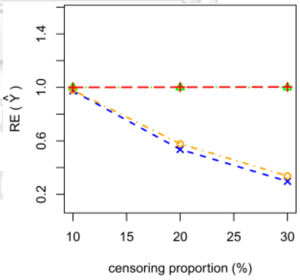
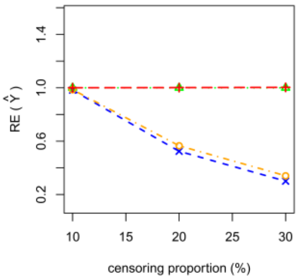
ตารางที่ 4. 7 ผลการเปรียบเทียบค่า $AMSE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$AMSE\left(\hat{Y}_a\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**0.03549**	0.05168	0.03757	0.05170
			1:2	**0.03542**	0.05371	0.03820	0.05386
			1:1	**0.03534**	0.05506	0.03770	0.05506
		20%	1:5	**0.04579**	0.06069	**0.04589**	0.06060
			1:2	**0.04559**	0.06126	**0.04592**	0.06108
			1:1	**0.04585**	0.06225	**0.04596**	0.06206
		30%	1:5	0.07192	0.07517	**0.06592**	0.07486
			1:2	0.07125	0.07250	**0.06508**	0.07218
			1:1	0.07082	0.07539	**0.06469**	0.07503
	50	10%	1:5	**0.02320**	0.02923	0.02377	0.02918
			1:2	**0.02327**	0.02962	0.02387	0.02954
			1:1	**0.02374**	0.02939	0.02429	0.02933
		20%	1:5	0.03708	**0.03279**	0.03522	**0.03273**
			1:2	0.03692	**0.03336**	0.03521	**0.03330**
			1:1	0.03662	**0.03312**	0.03489	**0.03306**
		30%	1:5	0.06550	**0.03914**	0.05824	**0.03900**
			1:2	0.06484	**0.03830**	0.05776	**0.03817**
			1:1	0.06549	**0.04020**	0.05821	**0.04005**
	100	10%	1:5	0.01457	**0.01432**	**0.01452**	**0.01431**
			1:2	0.01469	**0.01431**	0.01460	**0.01430**
			1:1	0.01445	**0.01411**	0.01438	**0.01411**
		20%	1:5	0.03009	**0.01577**	0.02788	**0.01574**
			1:2	0.02992	**0.01604**	0.02777	**0.01602**
			1:1	0.02976	**0.01562**	0.02758	**0.01560**
		30%	1:5	0.06009	**0.01804**	0.05298	**0.01798**
			1:2	0.06020	**0.01790**	0.05305	**0.01784**
			1:1	0.06018	**0.01771**	0.05309	**0.01766**

ตารางที่ 4. 8 ผลการเปรียบเทียบค่า $RE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:2$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$RE\left(\hat{Y}_a\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**1.46**	1.00	1.38	1.00
			1:2	**1.52**	1.00	1.41	1.00
			1:1	**1.56**	1.00	1.46	1.00
		20%	1:5	**1.33**	1.00	**1.32**	1.00
			1:2	**1.34**	1.00	**1.33**	1.00
			1:1	**1.36**	1.00	**1.35**	1.00
		30%	1:5	1.05	1.00	**1.14**	1.00
			1:2	1.02	1.00	**1.11**	1.00
			1:1	1.06	1.00	**1.17**	1.00
	50	10%	1:5	**1.26**	1.00	1.23	1.00
			1:2	**1.27**	1.00	1.24	1.00
			1:1	**1.24**	1.00	1.21	1.00
		20%	1:5	0.88	**1.00**	0.93	**1.00**
			1:2	0.90	**1.00**	0.95	**1.00**
			1:1	0.90	**1.00**	0.95	**1.00**
		30%	1:5	0.60	**1.00**	0.67	**1.00**
			1:2	0.59	**1.00**	0.66	**1.00**
			1:1	0.61	**1.00**	0.69	**1.00**
	100	10%	1:5	0.98	**1.00**	**0.99**	**1.00**
			1:2	0.97	**1.00**	0.98	**1.00**
			1:1	0.98	**1.00**	0.98	**1.00**
		20%	1:5	0.52	**1.00**	0.57	**1.00**
			1:2	0.54	**1.00**	0.58	**1.00**
			1:1	0.52	**1.00**	0.57	**1.00**
		30%	1:5	0.30	**1.00**	0.34	**1.00**
			1:2	0.30	**1.00**	0.34	**1.00**
			1:1	0.29	**1.00**	0.33	**1.00**

ตารางที่ 4. 9 กราฟเปรียบเทียบค่า $RE\left(\hat{Y}_a\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:2$

<table>
<tr><th>$\sigma^2_{x_1} : \sigma^2_{x_2}$
n</th><th>1:1</th><th>1:2</th><th>1:5</th></tr>
<tr><td>30</td><td></td><td></td><td></td></tr>
<tr><td>50</td><td></td><td></td><td></td></tr>
<tr><td>100</td><td></td><td></td><td></td></tr>
</table>

4.1.3 ผลการเปรียบเทียบวิธีการประมาณค่าตัวแปรตามเมื่อพิจารณาจากค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$ กรณีที่ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

จากตารางที่ 4.7 - 4.9 พบว่า

เมื่อตัวอย่างมีขนาดปานกลางถึงใหญ่ (n = 50, 100) พบว่า วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงที่สุดในทุกสถานการณ์ ยกเว้นในกรณีที่ตัวอย่างมีขนาดใหญ่ (n = 100) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) แม้วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตามยังคงเป็นวิธี MLE และวิธี MLE_EM แต่ในกรณีนี้วิธี CM และวิธี OLS จะมีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี วิธี MLE และวิธี MLE_EM เพียงเล็กน้อย และในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตาม และวิธี CM จะมีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี OLS เพียงเล็กน้อย

เมื่อขนาดตัวอย่างมีขนาดเล็ก (n = 30) ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) วิธี CM เป็นวิธีการประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงที่สุด แต่เมื่อตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) พบว่า วิธี OLS และวิธี CM จะเป็นวิธีที่มีประสิทธิภาพสูงที่สุดในการประมาณค่าตัวแปรตาม ส่วนกรณีตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตาม

ปัจจัยที่มีผลต่อค่า ค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$

ผลสรุปที่ได้เหมือนกับตารางที่ 4.1 - 4.3 ยกเว้นปัจจัยด้านขนาดตัวอย่างและเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา พบว่าส่วนใหญ่ ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับวิธี MLE นั่นคือ วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าขนาดตัวอย่างหรือเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวาจะเป็นเท่าไหร่ก็ตาม

4.1.4 สรุปผลการเปรียบเทียบวิธีการประมาณค่าตัวแปรตามเมื่อพิจารณาจากค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$

ตารางที่ 4. 10 วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่าตัวแปรตาม

n	r	$\sigma^2_{x_1+x_2} : \sigma^2_\varepsilon$	วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่าตัวแปรตาม
30	10%	2:1	CM*
		1:1	OLS*, CM
		1:2	OLS*
	20%	2:1	MLE*, MLE_EM*
		1:1	CM*
		1:2	OLS*, CM*
	30%	2:1, 1:1	MLE*, MLE_EM*
		1:2	CM*
50	10%	2:1	MLE*, MLE_EM*, CM
		1:1	CM*, OLS
		1:2	OLS*, CM
	20%, 30%	2:1,1:1,1:2	MLE*, MLE_EM*
100	10%	1:2	MLE*, MLE_EM*, OLS, CM
	20%, 30%	1:2	MLE*, MLE_EM*
	10%, 20%, 30%	2:1, 1:1	MLE*, MLE_EM*

หมายเหตุ * คือ วิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่าตัวแปรตาม

เมื่อขนาดตัวอย่างมีขนาดใหญ่ (n = 100) วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงที่สุดในทุกสถานการณ์ ยกเว้นในสถานการณ์ที่ความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระ และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) แม้ว่าวิธี MLE

และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตาม แต่สถานการณ์นี้ วิธี CM และวิธี OLS จะมีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย

เมื่อขนาดตัวอย่างมีขนาดปานกลาง (n = 50) วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตามในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก (r = 20%, 30%) ส่วนกรณีตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ผลสรุปเป็นดังนี้

- ความคลาดเคลื่อนกระจายน้อยกว่ากับตัวแปรอิสระ พบว่า วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตาม และวิธี CM นั้นมีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย
- ความคลาดเคลื่อนกระจายเท่ากันกับตัวแปรอิสระ พบว่า วิธี CM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตาม และวิธี OLS มีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี CM เพียงเล็กน้อย
- ความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระ พบว่า วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตาม และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามแตกต่างจากวิธี OLS เพียงเล็กน้อย

เมื่อตัวอย่างขนาดเล็ก (n = 30) พบว่าวิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีการประมาณค่าตัวแปรตามที่มีประสิทธิภาพสูงสุดในกรณีที่ความคลาดเคลื่อนกับตัวแปรอิสระกระจายเท่ากันและตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และในกรณีที่ความคลาดเคลื่อนกระจายน้อยกว่าตัวแปรอิสระและตัวแปรตามถูกตัดปลายทางขวาปานกลางและมาก (r = 20%, 30%) แต่ในกรณีที่ความคลาดเคลื่อนกระจายน้อยกว่าตัวแปรอิสระและตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) วิธี CM มีประสิทธิภาพสูงที่สุดในการประมาณค่าตัวแปรตาม เช่นเดียวกับในกรณีที่ความคลาดเคลื่อนกับตัวแปรอิสระกระจายเท่ากันและตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) และเมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระและตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ส่วนวิธี OLS มีประสิทธิภาพสูงสุดในการประมาณค่าตัวแปรตามเมื่อตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ความคลาดเคลื่อนกระจายมากกว่าหรือเท่ากับตัวแปรอิสระ ส่วนวิธี OLS และวิธี CM มีประสิทธิภาพสูงที่สุดในการประมาณค่าตัวแปรตามเมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระและตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%)

ปัจจัยที่มีผลต่อค่า $AMSE\left(\hat{Y}_a\right)$ และค่า $RE\left(\hat{Y}_a\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างใหญ่ขึ้น ทำให้ค่า $AMSE\left(\hat{Y}\right)$ ของทุกวิธีมีค่าลดลง นั่นคือ ทุกวิธีสามารถประมาณค่าของตัวแปรตามใกล้เคียงกับค่าที่แท้จริงของตัวแปรตามได้มากขึ้น

เมื่อความคลาดเคลื่อนกระจายน้อยกว่าตัวแปรอิสระ ส่วนใหญ่ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM จะมีค่ามากกว่าวิธี MLE เท่ากับ 0.01 แสดงว่า วิธี MLE_EM นั้นมีประสิทธิภาพในการประมาณค่าตัวแปรตามมากกว่าวิธี MLE แต่เนื่องจากต่างกันแค่เพียง 0.01 ถือว่าแตกต่างกันน้อยมาก ดังนั้นจึงสรุปได้ว่า วิธี MLE ยังคงเป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าขนาดตัวอย่างจะเป็นเท่าไหร่ก็ตาม ส่วนกรณีที่ความคลาดเคลื่อนกระจายมากกว่าหรือเท่ากับตัวแปรอิสระ ส่วนใหญ่ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM มีค่าเท่ากันกับวิธี MLE นั่นคือ วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าขนาดตัวอย่างจะเป็นเท่าไหร่ก็ตาม และเมื่อตัวอย่างมีขนาดใหญ่ขึ้น ค่า $RE\left(\hat{Y}\right)$ ของวิธี OLS กับวิธี CM มีค่าลดลงมาก นั่นคือเมื่อตัวอย่างมีขนาดใหญ่ขึ้นจะทำให้ประสิทธิภาพของการประมาณ ค่าตัวแปรตามของวิธี OLS และวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ ยิ่งตัวอย่างมีขนาดใหญ่ขึ้น วิธี OLS และวิธี CM จะประมาณค่าตัวแปรตามได้แย่กว่าวิธี MLE มากขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ทำให้ค่า $AMSE\left(\hat{Y}\right)$ ของทุกวิธีมีค่ามากขึ้น นั่นคือ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น จะทำให้ทุกวิธีประมาณค่าตัวแปรตามใกล้เคียงกับค่าแท้จริงของตัวแปรตามน้อยลง และเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $AMSE\left(\hat{Y}\right)$ ของวิธี OLS มีค่าเพิ่มขึ้นอย่างมาก รองลงมาคือ วิธี CM นั่นคือ เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวามีผลต่อการประมาณค่าตัวแปรตามของวิธี OLS เป็นอย่างมาก

ในกรณีที่ความคลาดเคลื่อนกระจายน้อยกว่าตัวแปรอิสระ พบว่า ส่วนใหญ่แล้ว ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM มีค่ามากกว่าค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE เท่ากับ 0.01 แสดงว่า วิธี MLE_EM นั้นมีประสิทธิภาพในการประมาณค่าตัวแปรตามมากกว่าวิธี MLE แต่เนื่องจากต่างกันแค่เพียง 0.01 ถือว่า

แตกต่างกันน้อยมาก ดังนั้นจึงสรุปได้ว่า วิธี MLE ยังคงเป็นวิธีที่มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าตัวแปรตามถูกตัดปลายทางขวาเป็นเท่าไหร่ก็ตาม ส่วนในกรณีความคลาดเคลื่อนกระจายมากกว่าหรือเท่ากับตัวแปรอิสระ พบว่า ส่วนใหญ่ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับค่า $RE\left(\hat{Y}\right)$ ของวิธี MLE นั่นคือวิธี MLE_EM มีประสิทธิภาพในการประมาณค่าตัวแปรตามเทียบเท่ากับวิธี MLE ไม่ว่าตัวแปรตามถูกตัดปลายทางขวาเป็นเท่าไหร่ก็ตาม แต่เมื่อตัวอย่างใหญ่ขึ้น ค่า $RE\left(\hat{Y}\right)$ ของวิธี OLS และวิธี CM จะมีค่าลดลงมาก นั่นคือ เมื่อตัวอย่างมีขนาดใหญ่ขึ้นจะทำให้ประสิทธิภาพของการประมาณมีค่าตัวแปรตามของวิธี OLS กับวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ ยิ่งตัวอย่างใหญ่ขึ้นวิธี OLS กับวิธี CM จะประมาณค่าตัวแปรตามได้แย่กว่าวิธี MLE มากขึ้น

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

ถึงแม้ว่าตัวแปรอิสระตัวที่ 1 จะมีการกระจายมากกว่าตัวแปรอิสระตัวที่ 2 มากขึ้น แต่ค่า $AMSE\left(\hat{Y}\right)$ ที่ขนาดตัวอย่างเดียวกัน เปอร์เซ็นต์การถูกตัดปลายทางขวาที่เท่ากันและอัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อนที่เท่ากันนั้นจะมีค่าใกล้เคียงมาก เช่นเดียวกับค่า $RE\left(\hat{Y}\right)$ ดังนั้น อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่าตัวแปรตามน้อยมาก

4. อัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อน

ในกรณีที่ความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระจะทำให้ค่า $AMSE\left(\hat{Y}\right)$ และค่า $RE\left(\hat{Y}\right)$ ของตัวแปรตามมีค่ามากกว่ากรณีที่ความคลาดเคลื่อนกระจายเท่ากันกับตัวแปรอิสระ และกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระตามลำดับ นั่นคือ อัตราส่วนการกระจายของตัวแปรอิสระต่อการกระจายของความคลาดเคลื่อนนั้นจะมีผลต่อการประมาณค่าตัวแปรตาม โดยเฉพาะอย่างยิ่งเมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้น จะทำให้ทุกวิธีสามารถประมาณค่าตัวแปรตามได้ใกล้เคียงกับค่าที่แท้จริงของตัวแปรตาม และเมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระ พบว่า วิธี OLS และวิธี CM นั้นจะมีประสิทธิภาพในการประมาณค่าตัวแปรตามได้ดีขึ้นเมื่อเทียบกับวิธี MLE

4.2 ผลการเปรียบเทียบวิธีการประมาณค่า β_0

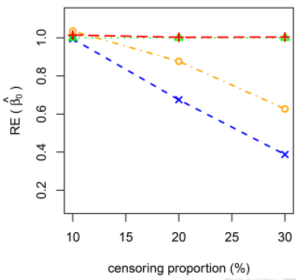
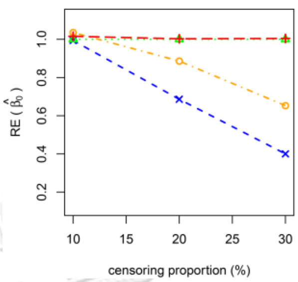
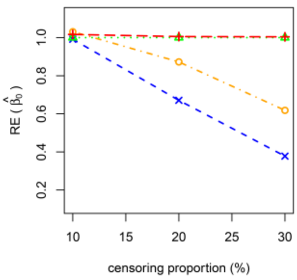
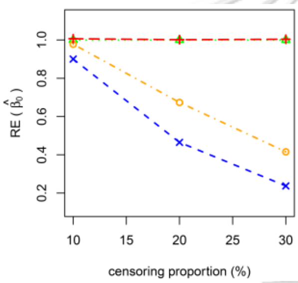
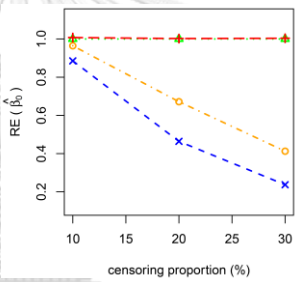
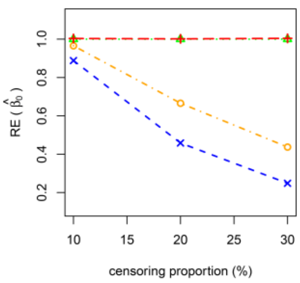
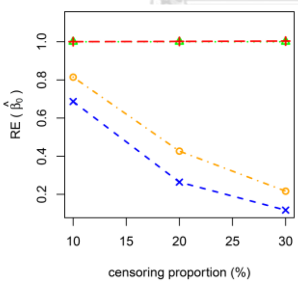
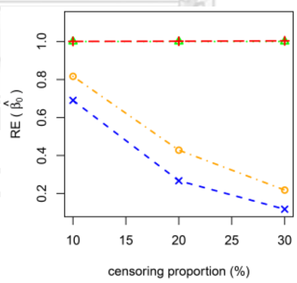
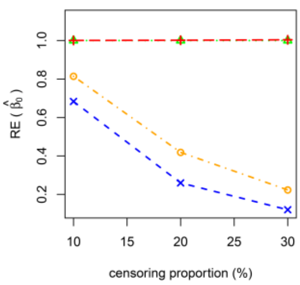
ตารางที่ 4. 11 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{0,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	0.00173	0.00171	**0.00166**	0.00169
			1:2	0.00170	0.00169	**0.00163**	0.00166
			1:1	0.00173	0.00172	**0.00166**	0.00169
		20%	1:5	0.00277	**0.00186**	0.00213	**0.00185**
			1:2	0.00278	**0.00191**	0.00216	**0.00191**
			1:1	0.00278	**0.00188**	0.00214	**0.00187**
		30%	1:5	0.00568	**0.00215**	0.00347	**0.00214**
			1:2	0.00563	**0.00225**	0.00345	**0.00224**
			1:1	0.00571	**0.00221**	0.00353	**0.00221**
	50	10%	1:5	0.00112	**0.00099**	0.00103	**0.00099**
			1:2	0.00110	**0.00097**	0.00101	**0.00096**
			1:1	0.00111	**0.00100**	0.00102	**0.00099**
		20%	1:5	0.00236	**0.00108**	0.00162	**0.00108**
			1:2	0.00230	**0.00107**	0.00159	**0.00107**
			1:1	0.00234	**0.00109**	0.00162	**0.00109**
		30%	1:5	0.00527	**0.00131**	0.00299	**0.00130**
			1:2	0.00537	**0.00127**	0.00308	**0.00127**
			1:1	0.00537	**0.00127**	0.00307	**0.00127**
	100	10%	1:5	0.00069	**0.00047**	0.00058	**0.00047**
			1:2	0.00071	**0.00049**	0.00060	**0.00049**
			1:1	0.00071	**0.00049**	0.00060	**0.00049**
		20%	1:5	0.00199	**0.00052**	0.00124	**0.00052**
			1:2	0.00201	**0.00054**	0.00125	**0.00053**
			1:1	0.00198	**0.00052**	0.00123	**0.00052**
		30%	1:5	0.00507	**0.00061**	0.00273	**0.00061**
			1:2	0.00511	**0.00060**	0.00275	**0.00060**
			1:1	0.00511	**0.00060**	0.00277	**0.00060**

ตารางที่ 4. 12 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$RE\left(\hat{\beta}_{0,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	0.99	1.00	**1.03**	1.02
			1:2	0.99	1.00	**1.04**	1.02
			1:1	0.99	1.00	**1.04**	1.01
		20%	1:5	0.67	**1.00**	0.87	**1.01**
			1:2	0.69	**1.00**	0.89	**1.00**
			1:1	0.67	**1.00**	0.88	**1.00**
		30%	1:5	0.38	**1.00**	0.62	**1.00**
			1:2	0.40	**1.00**	0.65	**1.00**
			1:1	0.39	**1.00**	0.63	**1.00**
	50	10%	1:5	0.89	**1.00**	0.97	**1.00**
			1:2	0.89	**1.00**	0.96	**1.01**
			1:1	0.90	**1.00**	0.98	**1.01**
		20%	1:5	0.46	**1.00**	0.67	**1.00**
			1:2	0.46	**1.00**	0.67	**1.00**
			1:1	0.46	**1.00**	0.67	**1.00**
		30%	1:5	0.25	**1.00**	0.44	**1.00**
			1:2	0.24	**1.00**	0.41	**1.00**
			1:1	0.24	**1.00**	0.42	**1.00**
	100	10%	1:5	0.68	**1.00**	0.81	**1.00**
			1:2	0.69	**1.00**	0.82	**1.00**
			1:1	0.69	**1.00**	0.81	**1.00**
		20%	1:5	0.26	**1.00**	0.42	**1.00**
			1:2	0.27	**1.00**	0.43	**1.00**
			1:1	0.26	**1.00**	0.43	**1.00**
		30%	1:5	0.12	**1.00**	0.22	**1.00**
			1:2	0.12	**1.00**	0.22	**1.00**
			1:1	0.12	**1.00**	0.22	**1.00**

ตารางที่ 4. 13 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

$\sigma^2_{x_1}:\sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.2.1 ผลการเปรียบเทียบวิธีการประมาณค่า β_0 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

จากตารางที่ 4.11 - 4.13 พบว่า

ส่วนใหญ่ วิธี MLE และวิธี MLE_EM เป็นวิธีประมาณค่า β_0 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีตัวอย่างมีขนาดปานกลาง (***n*** = 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (***r*** = 10%) แม้ว่าวิธี MLE และวิธี MLE_EM จะยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 แต่ในสถานการณ์นี้ วิธี CM มีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย ส่วนในกรณีที่ตัวอย่างมีขนาดเล็ก (***n*** = 30) และตัวแปรตามถูกตัดปลายทางขวาน้อย (***r*** = 10%) วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 คือ วิธี CM และอีก 3 วิธีที่เหลือต่างมีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจากวิธี CM เพียงเล็กน้อย

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างมีขนาดใหญ่ขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_0\right)$ ของทุกวิธีมีค่าลดลง นั่นคือ ทุกวิธีสามารถประมาณค่า β_0 ใกล้เคียงกับค่าที่แท้จริงของ β_0 มากขึ้นเมื่อตัวอย่างมีขนาดใหญ่ขึ้น

ส่วนใหญ่ค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี MLE_EM มีค่าเท่ากับค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี MLE แสดงว่าวิธี MLE_EM จะมีประสิทธิภาพในการประมาณค่า β_0 เท่ากับวิธี MLE แต่ในกรณีที่ขนาดตัวอย่างเล็กหรือปานกลาง (***n*** = 30, 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (***r*** = 10%) วิธี MLE_EM จะเป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_0 มากกว่าวิธี MLE เล็กน้อย และเมื่อตัวอย่างใหญ่ขึ้นค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี OLS กับวิธี CM มีค่าลดลงมาก นั่นคือเมื่อตัวอย่างมีขนาดใหญ่ขึ้นจะทำให้ประสิทธิภาพของการประมาณค่า β_0 ของวิธี OLS และวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ ยิ่งตัวอย่างใหญ่ขึ้น วิธี OLS และวิธี CM จะประมาณค่า β_0 ได้แย่กว่าวิธี MLE มากขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

ในกรณีที่เปอร์เซ็นต์การถูกตัดปลายทางขวาของตัวแปรตามเพิ่มขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_0\right)$ ของทุกวิธีมีค่ามากขึ้น นั่นคือ เมื่อเปอร์เซ็นต์การถูกตัดปลายทางขวาของตัวแปรตามมากขึ้น จะทำให้ทุกวิธีสามารถประมาณค่า β_0 ใกล้เคียงกับค่าที่แท้จริงของ β_0 น้อยลง โดยเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $MSE\left(\hat{\beta}_0\right)$ ของวิธี OLS จะมีค่าเพิ่มขึ้นอย่างมาก รองลงมา คือ วิธี CM นั่นคือ เปอร์เซ็นต์การถูกตัดปลายทางขวาของตัวแปรตามมีผลต่อการประมาณค่า β_0 ของวิธี OLS เป็นอย่างมาก

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM จะมีประสิทธิภาพในการประมาณค่า β_0 เท่ากับวิธี MLE แต่ในกรณีที่ขนาดตัวอย่างเท่ากับ 30 และ 50 และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) วิธี MLE_EM จะเป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_0 มากกว่าวิธี MLE เพียงเล็กน้อย และเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี OLS และวิธี CM มีค่าลดลงมาก นั่นคือ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น จะทำให้ประสิทธิภาพของการประมาณค่า β_0 ของวิธี OLS และวิธี CM ลดลงมากขึ้นเมื่อเทียบกับวิธี MLE กล่าวคือ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น วิธี OLS และวิธี CM จะประมาณค่า β_0 ได้แย่กว่าวิธี MLE มากขึ้น

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

ถึงแม้ตัวแปรอิสระตัวที่ 1 นั้นจะมีการกระจายมากกว่าตัวแปรอิสระตัวที่ 2 มากขึ้น แต่ว่าค่า $MSE\left(\hat{\beta}_0\right)$ ที่ขนาดตัวอย่างเดียวกัน ตัวแปรตามถูกตัดปลายทางขวาที่เท่ากันและอัตราส่วนของความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อนที่เท่ากันนั้นจะมีค่าใกล้เคียงมาก เช่นเดียวกับค่า $RE\left(\hat{\beta}_0\right)$ ดังนั้น อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่า β_0 น้อยมาก

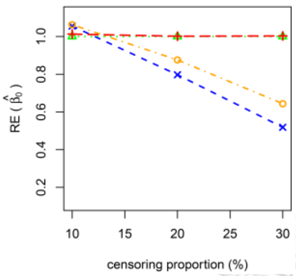
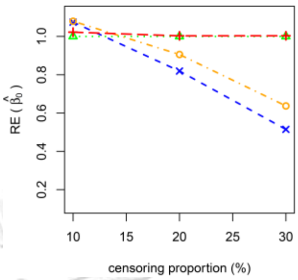
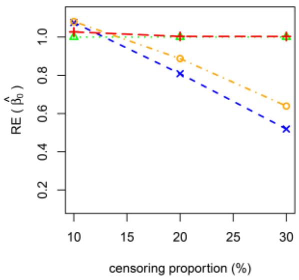
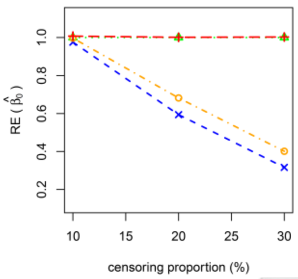
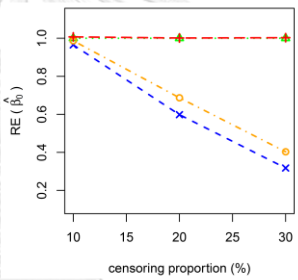
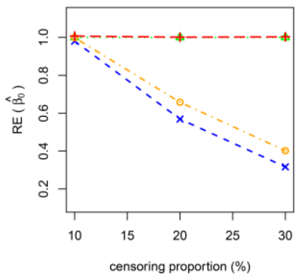
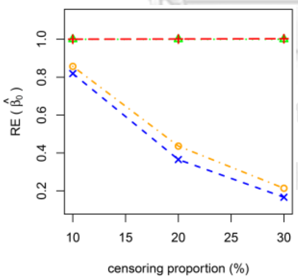
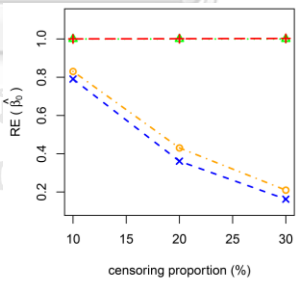
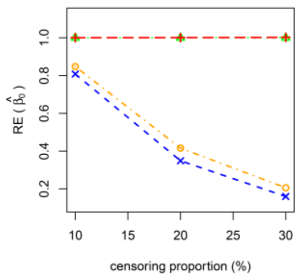
ตารางที่ 4. 14 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{0,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**0.00311**	0.00335	**0.00310**	0.00326
			1:2	**0.00311**	0.00334	**0.00309**	0.00327
			1:1	**0.00316**	0.00333	**0.00314**	0.00329
		20%	1:5	0.00456	**0.00368**	0.00415	**0.00367**
			1:2	0.00444	**0.00364**	0.00402	**0.00363**
			1:1	0.00461	**0.00367**	0.00419	**0.00367**
		30%	1:5	0.00817	**0.00424**	0.00664	**0.00423**
			1:2	0.00818	**0.00421**	0.00660	**0.00419**
			1:1	0.00823	**0.00427**	0.00663	**0.00425**
	50	10%	1:5	0.00203	**0.00199**	**0.00199**	**0.00198**
			1:2	0.00204	**0.00196**	**0.00199**	**0.00195**
			1:1	0.00202	**0.00197**	**0.00197**	**0.00195**
		20%	1:5	0.00360	**0.00205**	0.00311	**0.00205**
			1:2	0.00346	**0.00207**	0.00301	**0.00207**
			1:1	0.00352	**0.00209**	0.00307	**0.00209**
		30%	1:5	0.00743	**0.00235**	0.00585	**0.00234**
			1:2	0.00749	**0.00238**	0.00590	**0.00237**
			1:1	0.00755	**0.00239**	0.00595	**0.00238**
	100	10%	1:5	0.00120	**0.00097**	0.00114	**0.00097**
			1:2	0.00119	**0.00094**	0.00113	**0.00094**
			1:1	0.00118	**0.00096**	0.00112	**0.00096**
		20%	1:5	0.00292	**0.00102**	0.00245	**0.00102**
			1:2	0.00291	**0.00105**	0.00244	**0.00105**
			1:1	0.00286	**0.00104**	0.00239	**0.00104**
		30%	1:5	0.00702	**0.00112**	0.00545	**0.00112**
			1:2	0.00695	**0.00113**	0.00538	**0.00113**
			1:1	0.00702	**0.00117**	0.00546	**0.00116**

ตารางที่ 4. 15 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:1$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$RE\left(\hat{\beta}_{0,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**1.08**	1.00	**1.08**	1.03
			1:2	**1.07**	1.00	**1.08**	1.02
			1:1	**1.06**	1.00	**1.06**	1.01
		20%	1:5	0.81	**1.00**	0.89	**1.00**
			1:2	0.82	**1.00**	0.91	**1.00**
			1:1	0.80	**1.00**	0.88	**1.00**
		30%	1:5	0.52	**1.00**	0.64	**1.00**
			1:2	0.51	**1.00**	0.64	**1.00**
			1:1	0.52	**1.00**	0.64	**1.00**
	50	10%	1:5	0.98	**1.00**	**1.00**	**1.01**
			1:2	0.97	**1.00**	**0.99**	**1.01**
			1:1	0.98	**1.00**	**1.00**	**1.01**
		20%	1:5	0.57	**1.00**	0.66	**1.00**
			1:2	0.60	**1.00**	0.69	**1.00**
			1:1	0.59	**1.00**	0.68	**1.00**
		30%	1:5	0.32	**1.00**	0.40	**1.00**
			1:2	0.32	**1.00**	0.40	**1.00**
			1:1	0.32	**1.00**	0.40	**1.00**
	100	10%	1:5	0.81	**1.00**	0.85	**1.00**
			1:2	0.79	**1.00**	0.83	**1.00**
			1:1	0.82	**1.00**	0.86	**1.00**
		20%	1:5	0.35	**1.00**	0.42	**1.00**
			1:2	0.36	**1.00**	0.43	**1.00**
			1:1	0.37	**1.00**	0.44	**1.00**
		30%	1:5	0.16	**1.00**	0.21	**1.00**
			1:2	0.16	**1.00**	0.21	**1.00**
			1:1	0.17	**1.00**	0.21	**1.00**

ตารางที่ 4. 16 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

$\sigma^2_{x_1}:\sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.2.2 ผลการเปรียบเทียบวิธีการประมาณค่า β_0 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

จากตารางที่ 4.14 - 4.16 พบว่า

วิธี MLE และวิธี MLE_EM เป็นวิธีประมาณค่า β_0 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นในกรณีตัวอย่างมีขนาดปานกลาง (n = 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 และวิธี OLS จะมีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจากอีก 3 วิธี เพียงเล็กน้อย ส่วนในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 คือ วิธี CM และวิธี OLS

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$

ผลสรุปเหมือนกับตารางที่ 4.11 - 4.13

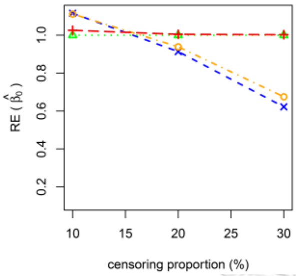
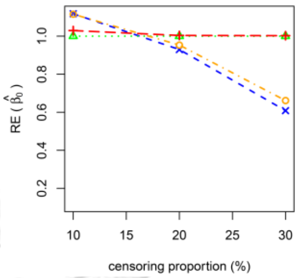
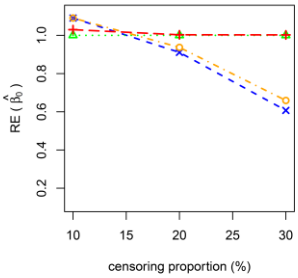
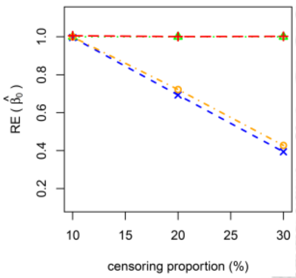
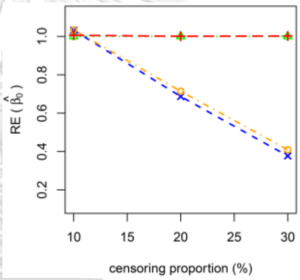
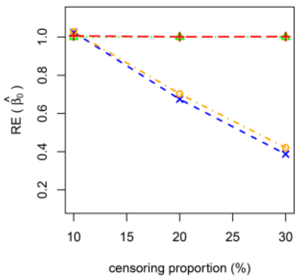
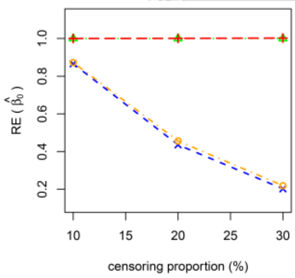
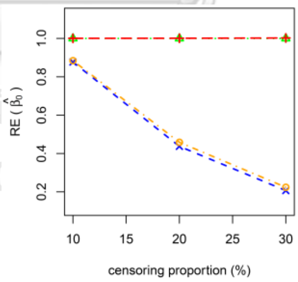
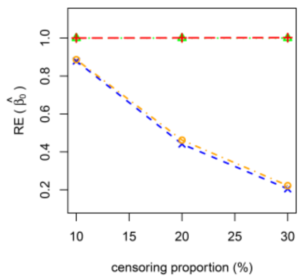
ตารางที่ 4. 17 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{0,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**0.00610**	0.00665	**0.00610**	0.00646
			1:2	**0.00604**	0.00675	**0.00605**	0.00656
			1:1	**0.00614**	0.00685	**0.00616**	0.00668
		20%	1:5	0.00777	**0.00708**	0.00756	**0.00706**
			1:2	0.00772	**0.00717**	0.00753	**0.00715**
			1:1	0.00791	**0.00722**	0.00769	**0.00719**
		30%	1:5	0.01337	**0.00812**	0.01231	**0.00810**
			1:2	0.01304	**0.00794**	0.01200	**0.00791**
			1:1	0.01309	**0.00815**	0.01206	**0.00812**
	50	10%	1:5	**0.00374**	0.00383	**0.00373**	0.00381
			1:2	**0.00375**	0.00387	**0.00375**	0.00385
			1:1	**0.00387**	0.00385	**0.00385**	0.00383
		20%	1:5	0.00596	**0.00402**	0.00572	**0.00401**
			1:2	0.00599	**0.00412**	0.00576	**0.00411**
			1:1	0.00596	**0.00413**	0.00572	**0.00413**
		30%	1:5	0.01190	**0.00461**	0.01095	**0.00460**
			1:2	0.01189	**0.00448**	0.01095	**0.00447**
			1:1	0.01197	**0.00471**	0.01103	**0.00470**
	100	10%	1:5	0.00217	**0.00191**	0.00215	**0.00191**
			1:2	0.00219	**0.00192**	0.00217	**0.00192**
			1:1	0.00219	**0.00190**	0.00218	**0.00190**
		20%	1:5	0.00457	**0.00202**	0.00436	**0.00201**
			1:2	0.00461	**0.00202**	0.00440	**0.00202**
			1:1	0.00465	**0.00202**	0.00444	**0.00202**
		30%	1:5	0.01080	**0.00222**	0.00997	**0.00221**
			1:2	0.01067	**0.00221**	0.00985	**0.00220**
			1:1	0.01074	**0.00217**	0.00991	**0.00217**

ตารางที่ 4. 18 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:2$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$RE\left(\hat{\beta}_{0,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**1.09**	1.00	**1.09**	1.03
			1:2	**1.12**	1.00	**1.12**	1.03
			1:1	**1.12**	1.00	**1.11**	1.03
		20%	1:5	0.91	**1.00**	0.94	**1.00**
			1:2	0.93	**1.00**	0.95	**1.00**
			1:1	0.91	**1.00**	0.94	**1.00**
		30%	1:5	0.61	**1.00**	0.66	**1.00**
			1:2	0.61	**1.00**	0.66	**1.00**
			1:1	0.62	**1.00**	0.68	**1.00**
	50	10%	1:5	**1.03**	1.00	**1.03**	1.01
			1:2	**1.03**	1.00	**1.03**	1.01
			1:1	**1.00**	1.00	**1.00**	1.01
		20%	1:5	0.67	**1.00**	0.70	**1.00**
			1:2	0.69	**1.00**	0.71	**1.00**
			1:1	0.69	**1.00**	0.72	**1.00**
		30%	1:5	0.39	**1.00**	0.42	**1.00**
			1:2	0.38	**1.00**	0.41	**1.00**
			1:1	0.39	**1.00**	0.43	**1.00**
	100	10%	1:5	0.88	**1.00**	0.89	**1.00**
			1:2	0.88	**1.00**	0.88	**1.00**
			1:1	0.87	**1.00**	0.87	**1.00**
		20%	1:5	0.44	**1.00**	0.46	**1.00**
			1:2	0.44	**1.00**	0.46	**1.00**
			1:1	0.44	**1.00**	0.46	**1.00**
		30%	1:5	0.21	**1.00**	0.22	**1.00**
			1:2	0.21	**1.00**	0.22	**1.00**
			1:1	0.20	**1.00**	0.22	**1.00**

ตารางที่ 4. 19 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{0,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

$\sigma^2_{x_1}:\sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.2.3 ผลการเปรียบเทียบวิธีการประมาณค่า β_0 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:2$

จากตารางที่ 4.17 - 4.19 พบว่า

วิธี MLE และวิธี MLE_EM เป็นวิธีประมาณค่า β_0 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี CM และวิธี OLS นั้นเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 ส่วนวิธี MLE และวิธีMLE_EM มีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจากวิธี CM และวิธี OLS เพียงเล็กน้อย ส่วนกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 คือ วิธี CM และวิธี OLS

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$

ผลสรุปเหมือนกับตารางที่ 4.11 - 4.13

4.2.4 การเปรียบเทียบวิธีการประมาณค่า β_0 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$

ตารางที่ 4. 20 วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_0

n	r	$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_0
30	10%	2:1	CM*, OLS, MLE, MLE_EM
		1:1, 1:2	OLS*, CM*
	20%, 30%	2:1, 1:1, 1:2	MLE*, MLE_EM*
50	10%	2:1	MLE*, MLE_EM*, CM
		1:1	MLE*, MLE_EM*, CM*, OLS
		1:2	OLS*, CM*, MLE, MLE_EM
	20%, 30%	2:1, 1:1, 1:2	MLE*, MLE_EM*
100	10%, 20%, 30%	2:1, 1:1, 1:2	MLE*, MLE_EM*

หมายเหตุ * คือ วิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_0

เมื่อตัวอย่างมีขนาดใหญ่ (n = 100) วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_0

เมื่อตัวอย่างมีขนาดเล็กและปานกลาง (n = 30, 50) ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลางและมาก (r = 20%, 30%) วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_0 แต่เมื่อตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) มีผลสรุปที่แตกต่างกันดังนี้

ตัวอย่างมีขนาดปานกลาง (n = 50)

- ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระ วิธี MLE และวิธี MLE_EM มีประสิทธิภาพมากสุดในการประมาณค่า β_0 และพบว่า วิธี CM มีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจาก วิธี MLE และวิธี MLE_EM เล็กน้อย

- ความคลาดเคลื่อนมีการกระจายเท่ากับตัวแปรอิสระ พบว่า วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_0 และวิธี OLS มีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจากอีก 3 วิธีเพียงเล็กน้อย
- ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ วิธี CM และวิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_0 และวิธี MLE กับวิธี MLE_EM จะมีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจากวิธี CM และวิธี OLS เล็กน้อย

ตัวอย่างมีขนาดเล็ก (n = 30)

- ความคลาดเคลื่อนกระจายน้อยกว่าตัวแปรอิสระ พบว่า วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 คือ วิธี CM และอีก 3 วิธีที่เหลือนั้นต่างมีประสิทธิภาพในการประมาณค่า β_0 แตกต่างจากวิธี CM เพียงเล็กน้อย
- ความคลาดเคลื่อนกระจายมากกว่าหรือเท่ากับตัวแปรอิสระ วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_0 คือ วิธี CM และวิธี OLS

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{0,a}\right)$ และค่า $RE\left(\hat{\beta}_{0,a}\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างมีขนาดใหญ่ขึ้น ทำให้ค่า $RE\left(\hat{\beta}_0\right)$ ของทุกวิธีมีค่าลดลง นั่นคือ ทุกวิธีสามารถประมาณค่า β_0 ใกล้เคียงกับค่าที่แท้จริงของ β_0 มากขึ้น

เมื่อตัวอย่างมีขนาดใหญ่ขึ้น พบว่าส่วนใหญ่ ค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_0 เท่ากับวิธี MLE แต่ในกรณีที่ขนาดตัวอย่างเท่ากับ 30 และ 50 และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) วิธี MLE_EM จะมีประสิทธิภาพในการประมาณค่า β_0 มากกว่าวิธี MLE เล็กน้อย และเมื่อตัวอย่างมีขนาดใหญ่ขึ้นค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี OLS และวิธี CM มีค่าลดลงมาก นั่นคือ เมื่อตัวอย่างมีขนาดใหญ่ขึ้นจะทำให้ประสิทธิภาพของการประมาณค่า β_0 ของวิธี OLS กับวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ ยิ่งตัวอย่างใหญ่ขึ้น วิธี OLS และวิธี CM จะประมาณค่า β_0 ได้แย่กว่าวิธี MLE มากขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

เมื่อตัวแปรตามถูกตัดปลายทางมากขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_0\right)$ ของทุกวิธีมีค่ามากขึ้น นั่นคือ เมื่อตัวแปรตามถูกตัดปลายทางมากขึ้น จะทำให้ทุกวิธีประมาณค่า β_0 ได้ใกล้เคียงกับค่าที่แท้จริงของ β_0 น้อยลง เมื่อตัวแปรตามถูกตัดปลายทางมากขึ้น ค่า $MSE\left(\hat{\beta}_0\right)$ ของวิธี OLS มีค่าเพิ่มขึ้นอย่างมาก รองลงมาคือวิธี CM นั่นคือ เปอร์เซ็นต์การถูกตัดปลายทางขวาของตัวแปรตามมีผลต่อการประมาณค่า β_0 ของวิธี OLS เป็นอย่างมาก

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_0 เท่ากับวิธี MLE แต่ในกรณีที่ขนาดตัวอย่างเท่ากับ 30, 50 และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_0 มากกว่าวิธี MLE เล็กน้อย และเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $RE\left(\hat{\beta}_0\right)$ ของวิธี OLS กับวิธี CM มีค่าลดลงมาก นั่นคือเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้นจะทำให้ประสิทธิภาพของการประมาณค่า β_0 ของวิธี OLS และวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ ยิ่งตัวแปรตามถูกตัดปลายทางขวามากขึ้น วิธี OLS กับวิธี CM จะประมาณค่า β_0 ได้แย่กว่าวิธี MLE มากขึ้น

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

ถึงแม้ตัวแปรอิสระตัวที่ 1 จะมีการกระจายมากกว่าตัวแปรอิสระตัวที่ 2 มากขึ้น แต่ค่า $MSE\left(\hat{\beta}_0\right)$ ที่ขนาดตัวอย่างเดียวกัน ตัวแปรตามถูกตัดปลายทางขวาเท่ากัน และอัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อนที่เท่ากัน จะมีค่าใกล้เคียงมาก เช่นเดียวกับค่า $RE\left(\hat{\beta}_0\right)$ ดังนั้น อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่า β_0 น้อยมาก

4. อัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อน

เมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระแล้วจะทำให้ค่า $MSE\left(\hat{\beta}_0\right)$ ของทุกวิธีมีค่ามากกว่ากรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ และกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระตามลำดับ นั่นคือ อัตราส่วนของการกระจายของตัวแปรอิสระต่อการกระจายของความคลาดเคลื่อนมีผลต่อการประมาณค่า β_0 โดยเฉพาะอย่างยิ่งเมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้นเท่าไหร่ จะทำให้ทุกวิธีสามารถประมาณค่า β_0 ได้ใกล้เคียงกับค่าที่แท้จริงของ β_0 ได้ลดลงเท่านั้น

4.3 ผลการเปรียบเทียบวิธีการประมาณค่า β_1

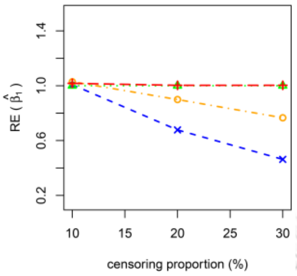
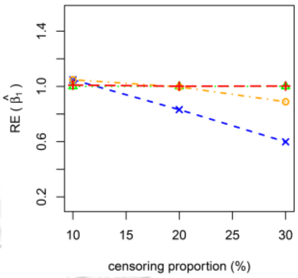
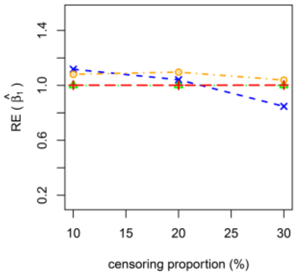
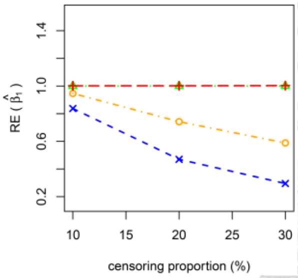
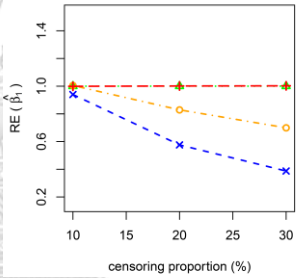
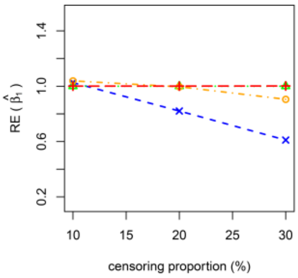
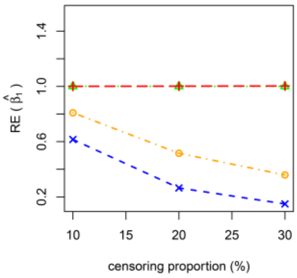
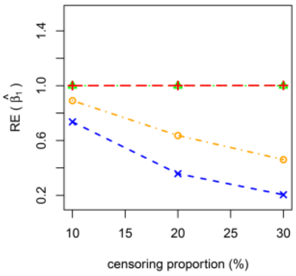
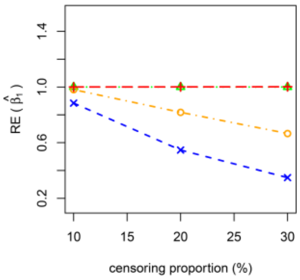
ตารางที่ 4. 21 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{1,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	**0.11134**	0.12440	0.11512	0.12421
			1:2	**0.06026**	0.06313	**0.06021**	0.06254
			1:1	0.04380	0.04413	**0.04288**	**0.04337**
		20%	1:5	0.13056	0.13592	**0.12392**	0.13575
			1:2	0.08495	**0.07065**	**0.07092**	**0.07056**
			1:1	0.06988	**0.04735**	0.05265	**0.04721**
		30%	1:5	0.17174	0.14551	**0.14011**	0.14522
			1:2	0.13360	**0.07999**	0.09002	**0.07979**
			1:1	0.11847	**0.05483**	0.07156	**0.05466**
	50	10%	1:5	**0.06719**	0.06892	**0.06638**	0.06887
			1:2	0.03810	**0.03581**	**0.03564**	**0.03584**
			1:1	0.02879	**0.02411**	0.02547	**0.02408**
		20%	1:5	0.09311	**0.07642**	**0.07674**	**0.07635**
			1:2	0.06548	**0.03771**	0.04549	**0.03766**
			1:1	0.05684	**0.02671**	0.03598	**0.02666**
		30%	1:5	0.13726	**0.08384**	0.09258	**0.08370**
			1:2	0.11432	**0.04429**	0.06331	**0.04419**
			1:1	0.10564	**0.03121**	0.05305	**0.03111**
	100	10%	1:5	0.03764	**0.03328**	0.03392	**0.03327**
			1:2	0.02340	**0.01722**	0.01933	**0.01722**
			1:1	0.01879	**0.01157**	0.01429	**0.01156**
		20%	1:5	0.06439	**0.03523**	0.04306	**0.03520**
			1:2	0.05171	**0.01851**	0.02908	**0.01849**
			1:1	0.04810	**0.01273**	0.02469	**0.01271**
		30%	1:5	0.11098	**0.03872**	0.05814	**0.03865**
			1:2	0.10226	**0.02090**	0.04537	**0.02086**
			1:1	0.09686	**0.01446**	0.04028	**0.01442**

ตารางที่ 4. 22 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$RE\left(\hat{\beta}_{1,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	**1.12**	1.00	1.08	1.00
			1:2	**1.05**	1.00	**1.05**	1.01
			1:1	1.01	1.00	**1.03**	**1.02**
		20%	1:5	1.04	1.00	**1.10**	1.00
			1:2	0.83	**1.00**	**1.00**	**1.00**
			1:1	0.68	**1.00**	0.90	**1.00**
		30%	1:5	0.85	1.00	**1.04**	1.00
			1:2	0.60	**1.00**	0.89	**1.00**
			1:1	0.46	**1.00**	0.77	**1.00**
	50	10%	1:5	**1.03**	1.00	**1.04**	1.00
			1:2	0.94	**1.00**	**1.00**	**1.00**
			1:1	0.84	**1.00**	0.95	**1.00**
		20%	1:5	0.82	**1.00**	**1.00**	**1.00**
			1:2	0.58	**1.00**	0.83	**1.00**
			1:1	0.47	**1.00**	0.74	**1.00**
		30%	1:5	0.61	**1.00**	0.91	**1.00**
			1:2	0.39	**1.00**	0.70	**1.00**
			1:1	0.30	**1.00**	0.59	**1.00**
	100	10%	1:5	0.88	**1.00**	0.98	**1.00**
			1:2	0.74	**1.00**	0.89	**1.00**
			1:1	0.62	**1.00**	0.81	**1.00**
		20%	1:5	0.55	**1.00**	0.82	**1.00**
			1:2	0.36	**1.00**	0.64	**1.00**
			1:1	0.26	**1.00**	0.52	**1.00**
		30%	1:5	0.35	**1.00**	0.67	**1.00**
			1:2	0.20	**1.00**	0.46	**1.00**
			1:1	0.15	**1.00**	0.36	**1.00**

ตารางที่ 4. 23 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 2:1$

$\sigma^2_{x_1} : \sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.3.1 ผลการเปรียบเทียบวิธีการประมาณค่า β_1 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 2:1$

จากตารางที่ 4.21 - 4.23 พบว่า

เมื่อขนาดตัวอย่างมีขนาดใหญ่ (n = 100)

โดยทั่วไป วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก แม้ในกรณีนี้ ทั้งวิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 แต่ในกรณีนี้ วิธี CM นั้นมีค่า $RE\left(\hat{\beta}_1\right)$ใกล้เคียงกับวิธี MLE และวิธี MLE_EM ดังนั้นจึงถือว่า วิธี CM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจาก วิธี MLE และวิธี MLE_EM เพียงเล็กน้อย

เมื่อขนาดตัวอย่างมีขนาดปานกลาง (n = 50)

ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก (r = 20%, 30%) พบว่า วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 ยกเว้นกรณีตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายที่แตกต่างกันมาก โดยพบว่า วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 นั้นคือ วิธี CM วิธี MLE และวิธี MLE_EM

ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ผลสรุปที่ได้เป็นดังนี้

- ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี MLE กับวิธี MLE_EM เพียงเล็กน้อย
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 คือ วิธี CM วิธี MLE และวิธี MLE_EM
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก วิธี CM และวิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 นอกจากนั้น พบว่า วิธี MLEและวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี CM และวิธี OLS เพียงเล็กน้อย

เมื่อขนาดตัวอย่างมีขนาดเล็ก (n = 30) ผลสรุปมีความหลากหลาย ดังนี้

- วิธี MLE และวิธี MLE_EM นั้นมีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อยหรือมีการกระจายเท่ากัน และในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
- วิธี CM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี MLE กับวิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี CM เพียงเล็กน้อย ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก
- วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 เมื่อตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
- วิธี CM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองกระจายต่างกันมาก
- วิธี CM และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี MLE กับวิธี OLS เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี CM กับวิธี MLE_EM เพียงเล็กน้อย เมื่อที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
- วิธี CM กับวิธี OLS มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี MLE กับวิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี CM และวิธี OLS เพียงเล็กน้อย ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
- วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี OLS เพียงเล็กน้อย กรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างมีขนาดใหญ่ขึ้นทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของทุกวิธีมีค่าลดลง นั่นคือทุกวิธีสามารถประมาณค่า β_1 ใกล้เคียงกับค่าที่แท้จริงของ β_1 มากขึ้น

นอกจากนั้น พบว่า ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี MLE_EM และวิธี MLE มีค่าเท่ากัน แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 เท่ากับวิธี MLE ยกเว้นเมื่อตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองกระจายเท่ากันหรือต่างกันน้อย พบว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 มากกว่าวิธี MLE เพียงเล็กน้อย และเมื่อตัวอย่างมีขนาดใหญ่ขึ้น ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี OLS กับวิธี CM มีค่าลดลงมาก นั่นคือ เมื่อตัวอย่างใหญ่ขึ้นจะทำให้ประสิทธิภาพของการประมาณค่า β_1 ของวิธี OLS กับวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ ยิ่งตัวอย่างมีขนาดใหญ่ขึ้น วิธี OLS และวิธี CM จะประมาณค่า β_1 ได้แย่กว่าวิธี MLE มากขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น จะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของทุกวิธีมีค่ามากขึ้น นั่นคือ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น จะทำให้ทุกวิธีประมาณค่า β_1 ใกล้เคียงกับค่า β_1 ที่แท้จริงน้อยลง ยกเว้นกรณีที่ขนาดตัวอย่างขนาดเล็ก (n = 30) ความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระ ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก ซึ่งเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของวิธี OLS และวิธี CM มีค่าลดลง นั่นคือ วิธี OLS และวิธี CM สามารถประมาณค่า β_1 ได้ใกล้เคียงกับค่าที่แท้จริงของ β_1 เพิ่มขึ้นเมื่อตัวแปรตามถูกตัดปลายทางขวาเพิ่มขึ้น

นอกจากนั้น พบว่า ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 เท่ากับวิธี MLE ยกเว้นกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากันหรือต่างกันน้อย พบว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 มากกว่าวิธี MLE เพียงเล็กน้อย

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $MSE\left(\hat{\beta}_1\right)$ ของวิธี OLS มีค่าเพิ่มขึ้นมาก รองลงมาคือวิธี CM นั่นคือ เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวามีผลต่อการประมาณค่า β_1 ของวิธี OLS เป็นอย่างมาก ยกเว้นกรณีที่ขนาดตัวอย่างมีขนาดเล็ก (n = 30) ความคลาดเคลื่อนมีการกระจายมากกว่าหรือเท่ากับตัวแปรอิสระ และกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระ ซึ่งเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลาย ทางขวามีผลต่อการประมาณ β_1 ของวิธี OLS เพียงเล็กน้อย

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น พบว่า ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี OLS และวิธี CM มีค่าลดลงมาก นั่นคือ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้นจะทำให้ทั้งวิธี OLS และวิธี CM ประมาณค่า β_1 ได้แย่กว่าวิธี MLE มากขึ้น

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

เมื่อตัวแปรอิสระตัวที่ 1 มีการกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากยิ่งขึ้นจะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของทุกวิธีเพิ่มขึ้น นั่นคือ อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่า β_1 โดยเมื่อตัวแปรอิสระตัวที่ 1 มีการกระจายแตกต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น จะทำให้ทุกวิธีสามารถประมาณค่า β_1 ได้ใกล้เคียงกับค่าที่แท้จริงของ β_1 ลดลง โดยเฉพาะในกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ นอกจากนั้น เมื่อตัวแปรอิสระตัวที่ 1 มีการกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น จะทำให้ค่า $RE\left(\hat{\beta}_1\right)$ ของทุกวิธีเพิ่มขึ้น แสดงว่าเมื่อตัวแปรอิสระตัวที่ 1 มีการกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น ทั้งวิธี OLS และวิธี CM จะสามารถประมาณค่า β_1 ได้ดีกว่าวิธี MLE มากขึ้นเท่านั้น

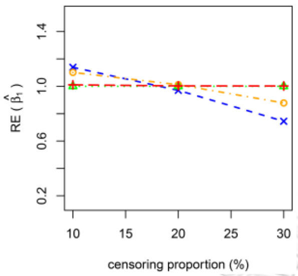
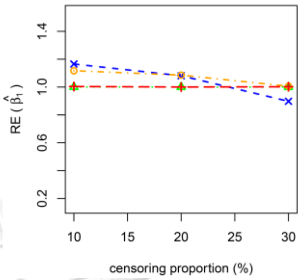
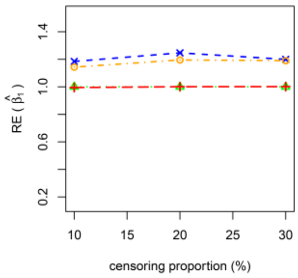
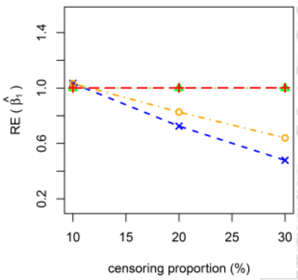
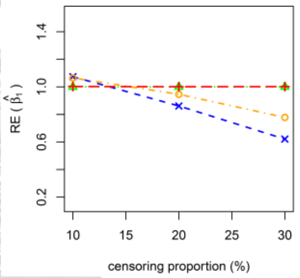
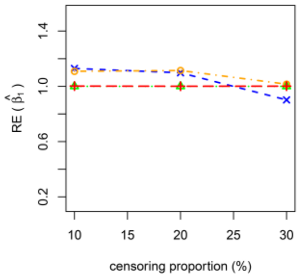
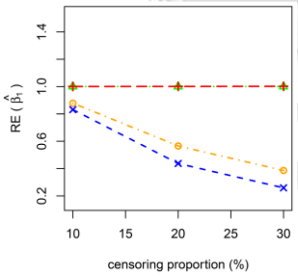
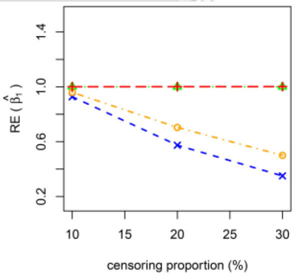
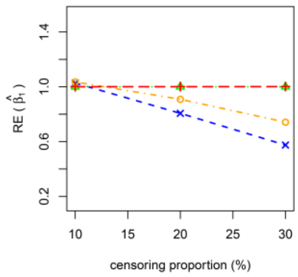
ตารางที่ 4. 24 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{1,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**0.20714**	0.24532	0.21469	0.24666
			1:2	**0.10724**	0.12497	0.11187	0.12451
			1:1	**0.07268**	0.08278	0.07516	0.08190
		20%	1:5	**0.21079**	0.26277	0.21993	0.26262
			1:2	**0.12414**	0.13420	**0.12362**	0.13418
			1:1	0.09347	**0.09068**	**0.08964**	**0.09042**
		30%	1:5	**0.23654**	0.28367	**0.23843**	0.28328
			1:2	0.16211	**0.14552**	**0.14464**	**0.14524**
			1:1	0.13926	**0.10369**	0.11804	**0.10346**
	50	10%	1:5	**0.12246**	0.13829	0.12472	0.13804
			1:2	**0.06420**	0.06893	**0.06473**	0.06873
			1:1	**0.04477**	0.04627	**0.04482**	0.04624
		20%	1:5	**0.13348**	0.14655	**0.13154**	0.14645
			1:2	0.08530	**0.07347**	0.07765	**0.07340**
			1:1	0.07136	**0.05173**	0.06258	**0.05168**
		30%	1:5	0.17484	0.15763	**0.15518**	0.15744
			1:2	0.13165	**0.08178**	0.10508	**0.08165**
			1:1	0.11720	**0.05609**	0.08757	**0.05598**
	100	10%	1:5	**0.06321**	0.06492	**0.06272**	0.06491
			1:2	0.03584	**0.03324**	0.03463	**0.03323**
			1:1	0.02670	**0.02220**	0.02526	**0.02219**
		20%	1:5	0.08508	**0.06859**	0.07549	**0.06855**
			1:2	0.06205	**0.03563**	0.05060	**0.03560**
			1:1	0.05448	**0.02383**	0.04213	**0.02381**
		30%	1:5	0.13012	**0.07475**	0.10078	**0.07467**
			1:2	0.11169	**0.03914**	0.07827	**0.03909**
			1:1	0.10302	**0.02672**	0.06916	**0.02667**

ตารางที่ 4. 25 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$RE\left(\hat{\beta}_{1,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**1.18**	1.00	1.14	0.99
			1:2	**1.17**	1.00	1.12	1.00
			1:1	**1.14**	1.00	1.10	1.01
		20%	1:5	**1.25**	1.00	1.19	1.00
			1:2	**1.08**	1.00	**1.09**	1.00
			1:1	0.97	**1.00**	**1.01**	**1.00**
		30%	1:5	**1.20**	1.00	**1.19**	1.00
			1:2	0.90	**1.00**	**1.01**	**1.00**
			1:1	0.74	**1.00**	0.88	**1.00**
	50	10%	1:5	**1.13**	1.00	1.11	1.00
			1:2	**1.07**	1.00	**1.06**	1.00
			1:1	**1.03**	1.00	**1.03**	1.00
		20%	1:5	**1.10**	1.00	**1.11**	1.00
			1:2	0.86	**1.00**	0.95	**1.00**
			1:1	0.72	**1.00**	0.83	**1.00**
		30%	1:5	0.90	1.00	**1.02**	1.00
			1:2	0.62	**1.00**	0.78	**1.00**
			1:1	0.48	**1.00**	0.64	**1.00**
	100	10%	1:5	**1.03**	1.00	**1.04**	1.00
			1:2	0.93	**1.00**	0.96	**1.00**
			1:1	0.83	**1.00**	0.88	**1.00**
		20%	1:5	0.81	**1.00**	0.91	**1.00**
			1:2	0.57	**1.00**	0.70	**1.00**
			1:1	0.44	**1.00**	0.57	**1.00**
		30%	1:5	0.57	**1.00**	0.74	**1.00**
			1:2	0.35	**1.00**	0.50	**1.00**
			1:1	0.26	**1.00**	0.39	**1.00**

ตารางที่ 4. 26 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:1$

$\sigma^2_{x_1} : \sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.3.2 ผลการเปรียบเทียบวิธีการประมาณค่า β_1 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:1$

จากตารางที่ 4.24 - 4.26 พบว่า

เมื่อตัวอย่างมีขนาดใหญ่ (n = 100)

โดยทั่วไป วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นในกรณีเหล่านี้ ได้แก่

- ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย พบว่า วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย
- ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก วิธี CM และวิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และพบว่า วิธี MLE กับวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี CM และวิธี OLS เพียงเล็กน้อย

เมื่อขนาดตัวอย่างมีขนาดเท่ากับ 30 และ 50 ผลสรุปมีความหลากหลาย ดังนี้

- วิธี MLE และวิธี MLE_EM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อยหรือมีการกระจายเท่ากัน และในกรณีที่ตัวอย่างนั้นมีขนาดปานกลาง (n = 50) ตัวแปรตามนั้นถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน และในกรณีที่ตัวอย่างมีขนาดเล็ก (n =30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองกระจายเท่ากัน
- วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจาก วิธี MLE และวิธี MLE_EM เพียงเล็กน้อย ในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
- วิธี CM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี MLE กับวิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี CM เพียงเล็กน้อย ในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก และตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%)

- วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวอย่างนั้นมีขนาดปานกลาง (n = 50) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ในกรณีที่ตัวอย่างขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองกระจายต่างกันน้อย ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองกระจายต่างกันมาก ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย
- วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และวิธี MLE กับวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี OLS และวิธี CM เพียงเล็กน้อยในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
- วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r =20%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก
- วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี OLS เพียงเล็กน้อย ในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมากและในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%)
- วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย และตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%)
- วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี OLS เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากอีก 3 วิธีเพียงเล็กน้อย เมื่อตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$

ผลสรุปเหมือนกับตารางที่ 4.21 - 4.23 ยกเว้น ปัจจัยด้านขนาดตัวอย่าง และเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา ซึ่งส่วนใหญ่ พบว่าค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 เท่ากับวิธี MLE ยกเว้นในกรณีที่ขนาดตัวอย่างเท่ากับ 30 ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน ซึ่งวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 มากกว่าวิธี MLE เล็กน้อย และในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก พบว่าวิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 น้อยกว่าวิธี MLE เพียงเล็กน้อย

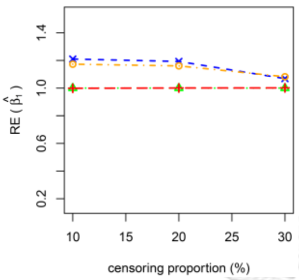
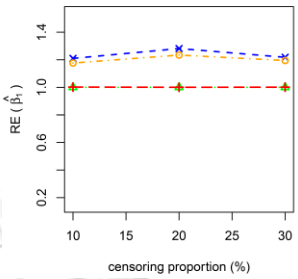
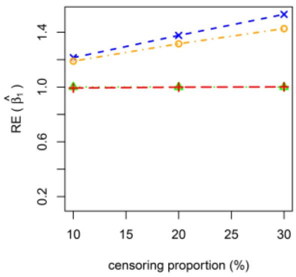
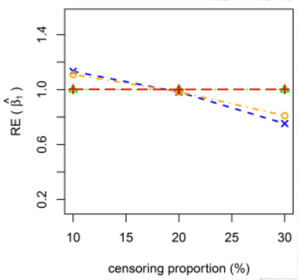
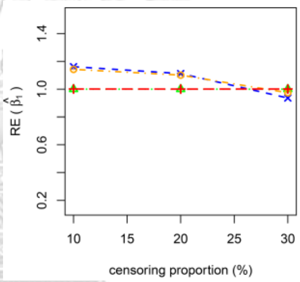
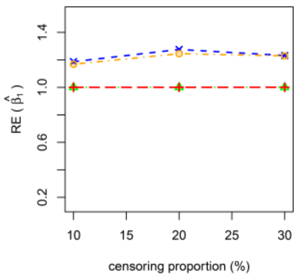
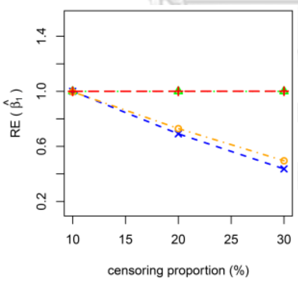
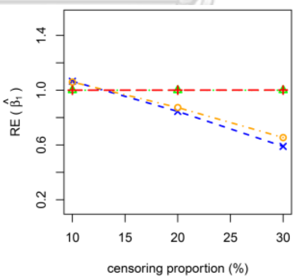
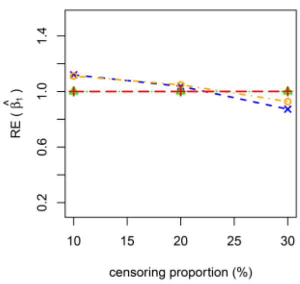
ตารางที่ 4. 27 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:2$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{1,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**0.40843**	0.49603	0.41705	0.49959
			1:2	**0.20041**	0.24249	0.20617	0.24201
			1:1	**0.13309**	0.16104	0.13729	0.16125
		20%	1:5	**0.36805**	0.50675	0.38518	0.50719
			1:2	**0.19821**	0.25398	0.20580	0.25388
			1:1	**0.14844**	0.17704	0.15255	0.17691
		30%	1:5	**0.36915**	0.56495	0.39600	0.56438
			1:2	**0.22801**	0.27769	0.23248	0.27736
			1:1	**0.18305**	0.19561	**0.18075**	0.19533
	50	10%	1:5	**0.22876**	0.27151	0.23238	0.27139
			1:2	**0.11826**	0.13735	0.12024	0.13718
			1:1	**0.08163**	0.09245	0.08333	0.09228
		20%	1:5	**0.22327**	0.28479	0.22876	0.28466
			1:2	**0.12878**	0.14321	**0.13005**	0.14312
			1:1	0.10035	**0.09851**	0.10016	**0.09845**
		30%	1:5	**0.24337**	0.29972	**0.24384**	0.29944
			1:2	0.16468	**0.15418**	0.15824	**0.15401**
			1:1	0.14138	**0.10638**	0.13126	**0.10624**
	100	10%	1:5	**0.11724**	0.13133	**0.11815**	0.13131
			1:2	**0.06123**	0.06523	**0.06155**	0.06521
			1:1	**0.04303**	**0.04306**	**0.04321**	**0.04305**
		20%	1:5	**0.12930**	0.13415	**0.12790**	0.13409
			1:2	0.08249	**0.06970**	0.07980	**0.06966**
			1:1	0.06600	**0.04563**	0.06255	**0.04560**
		30%	1:5	0.16231	**0.14169**	0.15272	**0.14157**
			1:2	0.12554	**0.07401**	0.11316	**0.07393**
			1:1	0.11583	**0.05054**	0.10196	**0.05048**

ตารางที่ 4. 28 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$RE\left(\hat{\beta}_{1,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**1.21**	1.00	1.19	0.99
			1:2	**1.21**	1.00	1.18	1.00
			1:1	**1.21**	1.00	1.17	1.00
		20%	1:5	**1.38**	1.00	1.32	1.00
			1:2	**1.28**	1.00	1.23	1.00
			1:1	**1.19**	1.00	1.16	1.00
		30%	1:5	**1.53**	1.00	1.43	1.00
			1:2	**1.22**	1.00	1.19	1.00
			1:1	**1.07**	1.00	**1.08**	1.00
	50	10%	1:5	**1.19**	1.00	1.17	1.00
			1:2	**1.16**	1.00	1.14	1.00
			1:1	**1.13**	1.00	1.11	1.00
		20%	1:5	**1.28**	1.00	1.24	1.00
			1:2	**1.11**	1.00	**1.10**	1.00
			1:1	0.98	**1.00**	0.98	**1.00**
		30%	1:5	**1.23**	1.00	**1.23**	1.00
			1:2	0.94	**1.00**	0.97	**1.00**
			1:1	0.75	**1.00**	0.81	**1.00**
	100	10%	1:5	**1.12**	1.00	**1.11**	1.00
			1:2	**1.07**	1.00	**1.06**	1.00
			1:1	**1.00**	**1.00**	**1.00**	**1.00**
		20%	1:5	**1.04**	1.00	**1.05**	1.00
			1:2	0.84	**1.00**	0.87	**1.00**
			1:1	0.69	**1.00**	0.73	**1.00**
		30%	1:5	0.87	**1.00**	0.93	**1.00**
			1:2	0.59	**1.00**	0.65	**1.00**
			1:1	0.44	**1.00**	0.50	**1.00**

ตารางที่ 4. 29 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{1,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

$\sigma^2_{x_1}:\sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.3.3 ผลการเปรียบเทียบวิธีการประมาณค่า β_1 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

จากตารางที่ 4.27 - 4.29 พบว่า

1) วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด ในกรณีที่
   - ตัวอย่างมีขนาดใหญ่ (n = 100) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%)
   - ตัวอย่างมีขนาดใหญ่ (n = 100) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r =20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากันหรือมีการกระจายแตกต่างกันน้อย
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากันหรือมีการกระจายแตกต่างกันน้อย
2) วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุดและวิธี CM กับวิธี OLS มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
3) ทุกวิธีมีประสิทธิภาพในการประมาณค่า β_1 เท่ากัน ในกรณีตัวอย่างมีขนาดใหญ่ (n =100) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองกระจายเท่ากัน
4) วิธี CM และวิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 ในกรณีที่
   - ตัวอย่างมีขนาดใหญ่ (n=100) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกัน
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) และตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน

5) วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก (r = 20%, 30%) ตัวแปรอิสระทั้งสองกระจายแตกต่างกันมาก

นอกจากกรณีเหล่านี้แล้ว วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี OLS เพียงเล็กน้อย

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$

ผลสรุปเหมือนกับตารางที่ 4.21 - 4.23 ยกเว้น ปัจจัยด้านขนาดตัวอย่าง และเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา ที่ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM จะมีประสิทธิภาพในการประมาณค่า β_1 เท่ากับวิธี MLE ยกเว้นในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน ซึ่งพบว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 มากกว่าวิธี MLE เพียงเล็กน้อย และในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก ซึ่งพบว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 น้อยกว่าวิธี MLE เพียงเล็กน้อย

4.3.4 สรุปผลการเปรียบเทียบวิธีการประมาณค่า β_1 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$

ตารางที่ 4. 30 วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1

n	r	$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	$\sigma^2_{x_1} : \sigma^2_{x_2}$	วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1
30	10%	2:1	1:1	CM*, MLE_EM*, OLS, MLE
			1:2	OLS*, CM*, MLE, MLE_EM
			1:5	OLS*, CM
		1:1,1:2	1:1,1:2,1:5	OLS*, CM
	20%	2:1	1:1	MLE*, MLE_EM*
			1:2	MLE*, MLE_EM*, CM*
			1:5	CM*
		1:1	1:1	MLE*, MLE_EM*, CM*, OLS
			1:2	OLS*, CM*
			1:5	OLS*
		1:2	1:1,1:2	OLS*
			1:5	OLS*, CM
	30%	2:1	1:1,1:2	MLE*, MLE_EM*
			1:5	CM*, MLE, MLE_EM
		1:1	1:1	MLE*, MLE_EM*
			1:2	MLE*, MLE_EM*, CM*
			1:5	OLS*, CM*
		1:2	1:1	OLS*, CM*
			1:2	OLS*
			1:5	OLS*, CM

n	r	$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	$\sigma^2_{x_1} : \sigma^2_{x_2}$	วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1
50	10%	2:1	1:1	MLE*, MLE_EM*, CM
			1:2	MLE*, MLE_EM*, CM*
			1:5	OLS*, CM*, MLE, MLE_EM
		1:1	1:1	OLS*, CM*, MLE, MLE_EM
			1:2	OLS*, CM*
			1:5	OLS*, CM
		1:2	1:1,1:2,1:5	OLS*, CM
	20%	2:1,1:1	1:1, 1:2	MLE*, MLE_EM*
		2:1	1:5	MLE*, MLE_EM*, CM*
		1:1	1:5	OLS*, CM*
		1:2	1:1	MLE*, MLE_EM*, OLS, CM
			1:2	OLS*, CM*
			1:5	OLS*, CM
	30%	2:1,1:1	1:1,1:2	MLE*, MLE_EM*
		2:1	1:5	MLE*, MLE_EM*
		1:1	1:5	CM*, MLE, MLE_EM
		1:2	1:1	MLE*, MLE_EM*
			1:2	MLE*, MLE_EM*, CM
			1:5	OLS*, CM*
100	10%	2:1	1:1,1:2	MLE*, MLE_EM*
			1:5	MLE*, MLE_EM*, CM
		1:1	1:1	MLE*, MLE_EM*
			1:2	MLE*, MLE_EM*, CM
			1:5	OLS*, CM*, MLE, MLE_EM

n	r	$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	$\sigma^2_{x_1}:\sigma^2_{x_2}$	วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1
100	10%	1:2	1:1	ทุกวิธี
			1:2, 1:5	OLS*, CM*
	20%	2:1,1:1	1:1,1:2,1:5	MLE*, MLE_EM*
		1:2	1:1,1:2	MLE*, MLE_EM*
			1:5	OLS*, CM*, MLE, MLE_EM
	30%	2:1,1:1,1:2	1:1,1:2,1:5	MLE*, MLE_EM*

หมายเหตุ * คือ วิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1

เมื่อขนาดตัวอย่างมีขนาดใหญ่ (n = 100)

โดยทั่วไป วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีดังต่อไปนี้

- กรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระ ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก และกรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ความคลาดเคลื่อนและตัวแปรอิสระมีการกระจายเท่ากัน ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย พบว่า วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย
- กรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ความคลาดเคลื่อนและตัวแปรอิสระมีการกระจายเท่ากัน ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก และในกรณีที่ตัวแปรตามนั้นถูกตัดปลายทางขวาปานกลาง (r = 20%) ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก พบว่า วิธี OLS และวิธี CM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด และวิธี MLE กับวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจาก วิธี OLS กับวิธี CM เพียงเล็กน้อย
- กรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่า ทุกวิธีมีประสิทธิภาพในการประมาณค่า β_1 เท่ากัน

- กรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ความคลาดเคลื่อนนั้นมีการกระจายมากกว่าตัวแปรอิสระ และตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกัน พบว่า วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1

เมื่อขนาดตัวอย่างมีขนาดลดลงเท่ากับ 30 และ 50 ผลสรุปมีความหลากหลาย จะเห็นว่าทุกปัจจัยล้วนส่งผลต่อประสิทธิภาพของการประมาณค่า β_1 ดังนั้นจึงจะทำการพิจารณาปัจจัยด้านอัตราส่วนของความแปรปรวนของตัวแปรอิสระต่อความแปรปรวนรวมของตัวแปรอิสระก่อน ดังนี้

<u>ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระ</u>

1) วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%)
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อยหรือมีการกระจายเท่ากัน
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อยหรือมีการกระจายเท่ากัน
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) และตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
2) วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และวิธี CM นั้นมีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย กรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
3) วิธี CM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และพบว่า วิธี MLE กับวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี CM เพียงเล็กน้อย ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก
4) วิธี CM วิธี MLE และวิธี MLE_EM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) และตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) และตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย

5) วิธี CM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองกระจายต่างกันมาก
6) วิธี CM และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี MLE กับวิธี OLS มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี CM กับวิธี MLE_EM เพียงเล็กน้อย เมื่อตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
7) วิธี CM กับวิธี OLS มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี MLE กับวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี CM และวิธี OLS เพียงเล็กน้อย กรณีที่ตัวอย่างขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก และในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
8) วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี OLS เพียงเล็กน้อย ในกรณีที่ตัวอย่างเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก

<u>**ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ**</u>

1) วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อยหรือมีการกระจายเท่ากัน
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
   - ตัวอย่างขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
2) วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจาก วิธี MLE และวิธี MLE_EM เพียงเล็กน้อย ในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
3) วิธี CM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และพบว่าวิธี MLE กับวิธี MLE_EM นั้นมีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี CM เล็กน้อย เมื่อตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองตัวมีการกระจายต่างกันมาก

4) วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
5) วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และพบว่า วิธี MLE กับวิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี OLS กับวิธี CM เล็กน้อย เมื่อตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน
6) วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 ในกรณีที่ตัวอย่างเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองกระจายต่างกันมาก
7) วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากวิธี OLS เพียงเล็กน้อย ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%)
8) วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 กรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย
9) วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_1 และวิธี OLS เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_1 ต่างจากอีก 3 วิธีเพียงเล็กน้อย ในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน

ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ

1) วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด ในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน

2) วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_1 ที่มีประสิทธิภาพสูงที่สุด และวิธี CM กับวิธี OLS มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจาก วิธี MLE และวิธี MLE_EM เล็กน้อย ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน

3) วิธี CM และวิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 ในกรณีที่
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก
   - ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันน้อย
   - ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน

4) วิธี OLS เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 กรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก (r = 20%, 30%) ตัวแปรอิสระทั้งสองกระจายแตกต่างกันมาก

นอกจากกรณีที่กล่าวมาข้างต้นแล้ว พบว่า วิธี OLS มีประสิทธิภาพสูงสุดในการประมาณค่า β_1 และวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 แตกต่างจากวิธี OLSเล็กน้อย

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{1,a}\right)$ และค่า $RE\left(\hat{\beta}_{1,a}\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างมีขนาดใหญ่ขึ้น จะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของทุกวิธีมีค่าลดลง นั่นคือ ทุกวิธีจะประมาณค่า β_1 ใกล้เคียงกับ β_1 ที่แท้จริงมากขึ้น

นอกจากนั้น พบว่า ส่วนใหญ่ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 เท่ากับวิธี MLE และเมื่อตัวอย่างมีขนาดใหญ่ขึ้น ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี OLS กับวิธี CM มีค่าลดลงมาก นั่นคือ เมื่อตัวอย่างมีขนาดใหญ่ขึ้นจะทำให้ประสิทธิภาพของการประมาณค่า β_1 ของวิธี OLS กับวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ ยิ่งตัวอย่างมีขนาดใหญ่ขึ้น วิธี OLS กับวิธี CM จะประมาณค่า β_1 แย่กว่าวิธี MLE มากขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้นจะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของทุกวิธีมีค่ามากขึ้น นั่นคือ หากตัวแปรตามถูกตัดปลายทางขวามากขึ้น จะทำให้ทุกวิธีประมาณค่า β_1 ได้ใกล้เคียงกับค่า β_1 ที่แท้จริงน้อยลง ยกเว้นกรณีที่ขนาดตัวอย่างขนาดเล็ก (n = 30) ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ตัวแปรอิสระทั้งสองมีการกระจายแตกต่างกันมาก พบว่า วิธี OLS กับวิธี CM สามารถประมาณค่าสัมประสิทธิ์หน้าตัวแปรอิสระตัวที่ 1 ได้ใกล้เคียงกับค่า β_1 ที่แท้จริงเพิ่มขึ้นเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น

นอกจากนั้น พบว่า ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_1 เท่ากับวิธี MLE และเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น จะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของวิธี OLS และ CM มีค่าลดลง นั่นคือ วิธี OLS และวิธี CM สามารถประมาณค่า β_1 ได้ใกล้เคียงกับค่า β_1 ที่แท้จริงเพิ่มขึ้น

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $MSE\left(\hat{\beta}_1\right)$ ของวิธี OLS มีค่าเพิ่มขึ้นมากรองลงมาคือ วิธี CM นั่นคือ เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวามีผลต่อการประมาณค่า β_1 ของวิธี OLS เป็นอย่างมาก ยกเว้นกรณีที่ขนาดตัวอย่างขนาดเล็ก (n = 30) ความคลาดเคลื่อนมีการกระจายมากกว่าหรือเท่ากับตัวแปรอิสระ และกรณีที่กรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ซึ่งเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวามีผลต่อการประมาณ β_1 ของวิธี OLS เพียงเล็กน้อย

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี OLS และวิธี CM จะมีค่าลดลงมาก นั่นคือ เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น จะทำให้ประสิทธิภาพของการประมาณค่า β_1 ของวิธี OLS และวิธี CM เมื่อเทียบกับวิธี MLE ลดลงมากขึ้น กล่าวคือ วิธี OLS กับวิธี CM จะประมาณค่า β_1 ได้แย่ขึ้นเมื่อเทียบกับวิธี MLE

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

เมื่อตัวแปรอิสระตัวที่ 1 นั้นมีการกระจายแตกต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น จะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ ของทุกวิธีเพิ่มขึ้น โดยเฉพาะอย่างยิ่งกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระแล้วจะเห็นการเพิ่มขึ้นของค่า $MSE\left(\hat{\beta}_1\right)$ ชัดเจนมากยิ่งขึ้น แสดงให้เห็นว่า อัตราส่วน

ความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่า β_1 โดยเมื่อตัวแปรอิสระตัวที่ 1 มีการกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น จะทำให้ทุกวิธีสามารถประมาณค่า β_1 ได้ใกล้เคียงกับ β_1 ที่แท้จริงลดลง โดยเฉพาะอย่างยิ่งในกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ นอกจากนั้นเมื่อตัวแปรอิสระตัวที่ 1 มีกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น จะทำให้ค่า $RE\left(\hat{\beta}_1\right)$ ของทุกวิธีเพิ่มขึ้น แสดงว่า เมื่อตัวแปรอิสระตัวที่ 1 กระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น วิธี OLS และวิธี CM สามารถประมาณค่า β_1 ได้ดีขึ้นเมื่อเทียบกับวิธี MLE

4. อัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อน

เมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมีการจะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ มีค่ามากกว่ากรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากับตัวแปรอิสระ และกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระตามลำดับ นั่นคือ อัตราส่วนของการกระจายของตัวแปรอิสระต่อการกระจายของความคลาดเคลื่อนมีผลต่อการประมาณค่า β_1 โดยเฉพาะ เมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้นเท่าไหร่ จะทำให้ทุกวิธีสามารถประมาณค่า β_1 ได้ใกล้เคียงกับค่า β_1 ที่แท้จริงได้น้อยลงเท่านั้น

เมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ทำให้ค่า $RE\left(\hat{\beta}_1\right)$ ของวิธี OLS กับวิธี CM มีค่ามากกว่ากรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ และกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระตามลำดับ นั่นคือ ยิ่งความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้นเท่าไหร่ จะทำให้วิธี OLS กับวิธี CM มีประสิทธิภาพในการประมาณค่า β_1 ได้ดีขึ้นเมื่อเทียบกับวิธี MLE และเมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระจะทำให้ค่า $MSE\left(\hat{\beta}_1\right)$ และค่า $RE\left(\hat{\beta}_1\right)$ ที่ได้จากวิธี OLS กับวิธี CM มีค่าใกล้เคียงกันมากขึ้น มากกว่าในกรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ และในกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระตามลำดับ นั่นคือ ยิ่งความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้นแล้วจะทำให้วิธี OLS และวิธี CM สามารถประมาณค่า β_1 ได้ใกล้เคียงกันมากยิ่งขึ้น ยกเว้นกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%)

4.4 ผลการเปรียบเทียบวิธีการประมาณค่า β_2

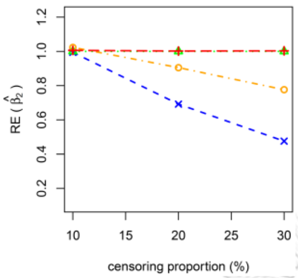
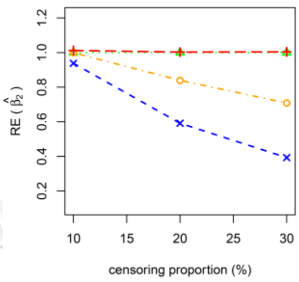
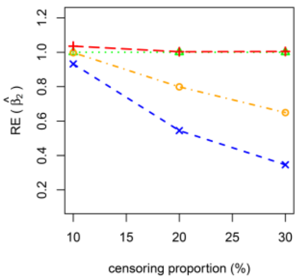
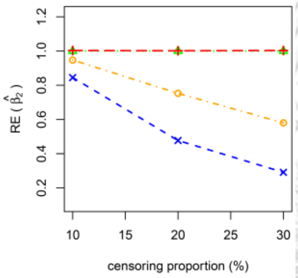
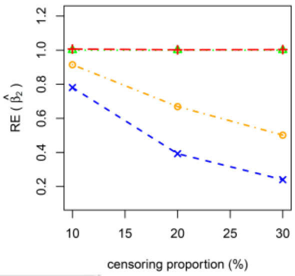
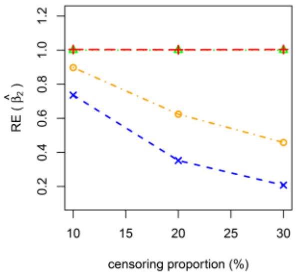
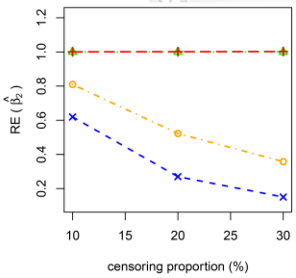
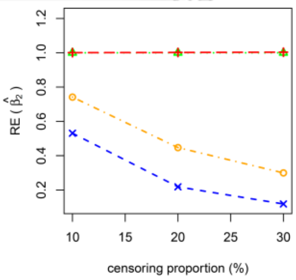
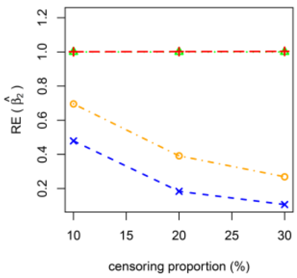
ตารางที่ 4. 31 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{2,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	0.02844	0.02651	0.02657	**0.02560**
			1:2	0.03440	**0.03226**	**0.03229**	**0.03186**
			1:1	0.04341	0.04316	**0.04218**	**0.04291**
		20%	1:5	0.05767	**0.03141**	0.03932	**0.03131**
			1:2	0.06153	**0.03639**	0.04335	**0.03629**
			1:1	0.07033	**0.04866**	0.05376	**0.04856**
		30%	1:5	0.10649	**0.03685**	0.05675	**0.03669**
			1:2	0.11082	**0.04353**	0.06146	**0.04334**
			1:1	0.11690	**0.05561**	0.07164	**0.05543**
	50	10%	1:5	0.02050	**0.01509**	0.01680	**0.01503**
			1:2	0.02336	**0.01824**	0.01994	**0.01812**
			1:1	0.02948	**0.02491**	0.02629	**0.02484**
		20%	1:5	0.04934	**0.01737**	0.02782	**0.01733**
			1:2	0.05288	**0.02076**	0.03103	**0.02072**
			1:1	0.05723	**0.02730**	0.03627	**0.02726**
		30%	1:5	0.09999	**0.02077**	0.04531	**0.02069**
			1:2	0.10204	**0.02451**	0.04884	**0.02442**
			1:1	0.10545	**0.03071**	0.05293	**0.03062**
	100	10%	1:5	0.01523	**0.00729**	0.01048	**0.00728**
			1:2	0.01665	**0.00884**	0.01192	**0.00883**
			1:1	0.01887	**0.01168**	0.01443	**0.01168**
		20%	1:5	0.04460	**0.00818**	0.02090	**0.00816**
			1:2	0.04580	**0.01000**	0.02236	**0.00998**
			1:1	0.04815	**0.01299**	0.02488	**0.01298**
		30%	1:5	0.09463	**0.01003**	0.03732	**0.01000**
			1:2	0.09547	**0.01131**	0.03776	**0.01127**
			1:1	0.09755	**0.01467**	0.04099	**0.01463**

ตารางที่ 4. 32 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 2:1$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$RE\left(\hat{\beta}_{2,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
2:1	30	10%	1:5	0.93	1.00	1.00	**1.04**
			1:2	0.94	**1.00**	**1.00**	**1.01**
			1:1	0.99	1.00	**1.02**	**1.01**
		20%	1:5	0.54	**1.00**	0.80	**1.00**
			1:2	0.59	**1.00**	0.84	**1.00**
			1:1	0.69	**1.00**	0.91	**1.00**
		30%	1:5	0.35	**1.00**	0.65	**1.00**
			1:2	0.39	**1.00**	0.71	**1.00**
			1:1	0.48	**1.00**	0.78	**1.00**
	50	10%	1:5	0.74	**1.00**	0.90	**1.00**
			1:2	0.78	**1.00**	0.91	**1.01**
			1:1	0.84	**1.00**	0.95	**1.00**
		20%	1:5	0.35	**1.00**	0.62	**1.00**
			1:2	0.39	**1.00**	0.67	**1.00**
			1:1	0.48	**1.00**	0.75	**1.00**
		30%	1:5	0.21	**1.00**	0.46	**1.00**
			1:2	0.24	**1.00**	0.50	**1.00**
			1:1	0.29	**1.00**	0.58	**1.00**
	100	10%	1:5	0.48	**1.00**	0.70	**1.00**
			1:2	0.53	**1.00**	0.74	**1.00**
			1:1	0.62	**1.00**	0.81	**1.00**
		20%	1:5	0.18	**1.00**	0.39	**1.00**
			1:2	0.22	**1.00**	0.45	**1.00**
			1:1	0.27	**1.00**	0.52	**1.00**
		30%	1:5	0.11	**1.00**	0.27	**1.00**
			1:2	0.12	**1.00**	0.30	**1.00**
			1:1	0.15	**1.00**	0.36	**1.00**

ตารางที่ 4. 33 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

<table>
<tr><th>$\sigma^2_{x_1}:\sigma^2_{x_2}$ / n</th><th>1:1</th><th>1:2</th><th>1:5</th></tr>
<tr><td>30</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>50</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>100</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

4.4.1 ผลการเปรียบเทียบวิธีการประมาณค่า β_2 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=2:1$

จากตารางที่ 4.31 - 4.33 พบว่า

โดยทั่วไปวิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงสุด ยกเว้นในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ซึ่งอัตราส่วนระหว่างตัวแปรอิสระทั้งสองจะเป็นตัวบ่งบอกถึงวิธีการที่มีประสิทธิภาพสูงที่สุดในประมาณค่า β_2 ดังนี้

- ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่า วิธี CM และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และวิธี OLS กับวิธี MLE เป็นวิธีที่มีประสิทธิภาพแตกต่างจากวิธี CM และวิธี MLE_EM เล็กน้อย
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย พบว่า วิธีที่มีประสิทธิภาพสูงสุดในการประมาณ ค่า β_2 คือ วิธี CM วิธี MLE และวิธี MLE_EM
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก พบว่า วิธี MLE_EM มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และวิธี MLE กับวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจาก วิธี MLE_EM เพียงเล็กน้อย

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างใหญ่ขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ ของทุกวิธีมีค่าลดลง นั่นคือ ทุกวิธีนั้นสามารถประมาณค่า β_2 ใกล้เคียงกับ β_2 ที่แท้จริงมากขึ้น

ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี MLE_EM มีค่าเท่ากับค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_2 เท่ากับวิธี MLE ยกเว้นในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และในกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองกระจายต่างกันน้อย พบว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_2 มากกว่าวิธี MLE เล็กน้อย และเมื่อตัวอย่างมีขนาดใหญ่ขึ้น ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี OLS และวิธี CM มีค่าลดลงเป็นอย่างมาก แสดงว่า วิธี OLS กับวิธี CM จะประมาณค่า β_2 ได้แย่กว่าวิธี MLE มากขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ ของทุกวิธีมีค่ามากขึ้น นั่นคือ ทุกวิธีประมาณค่า β_2 ใกล้เคียงกับค่า β_2 ที่แท้จริงน้อยลง

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_2 เท่ากับวิธี MLE ยกเว้นในกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ซึ่งพบว่าวิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_2 มากกว่าวิธี MLE เล็กน้อย และเมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น พบว่าค่า $MSE\left(\hat{\beta}_2\right)$ ของวิธี OLS มีค่าเพิ่มขึ้นมาก รองลงมาคือวิธี CM นั่นคือ เปอร์เซ็นต์การถูกตัดปลายทางขวาของตัวแปรตามมีผลต่อการประมาณค่า β_2 ของวิธี OLS เป็นอย่างมาก

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

เมื่อตัวแปรอิสระตัวที่ 1 กระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ ของทุกวิธีลดลง โดยเฉพาะอย่างยิ่งเมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ จะเห็นการลดลงของค่า $MSE\left(\hat{\beta}_2\right)$ ชัดเจนมากยิ่งขึ้น นั่นคือ อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่า β_2 โดยเมื่อตัวแปรอิสระ ตัวที่ 1 มีกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น จะทำให้ทุกวิธีสามารถประมาณค่า β_2 ได้ใกล้เคียงกับ β_2 ที่แท้จริงเพิ่มขึ้น นอกจากนั้นเมื่อตัวแปรอิสระตัวที่ 1 มีกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น จะทำให้ค่า $RE\left(\hat{\beta}_2\right)$ ของทุกวิธีลดลงมากขึ้น แสดงว่า เมื่อตัวแปรอิสระทั้งสองตัวกระจายต่างกันมากขึ้น ทั้งวิธี OLS และวิธี CM จะประมาณค่า β_2 ได้ใกล้เคียงกับวิธี MLE ลดลงเท่านั้น

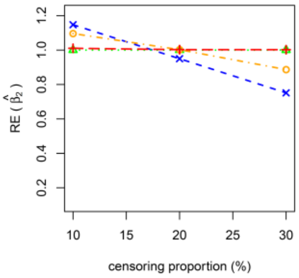
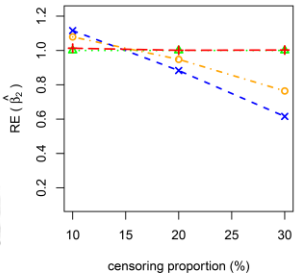
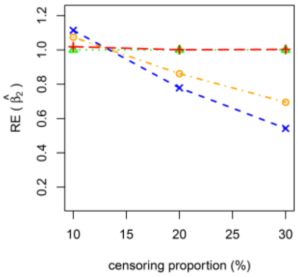
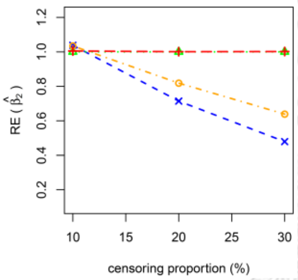
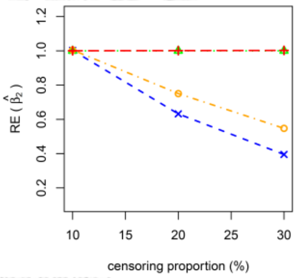
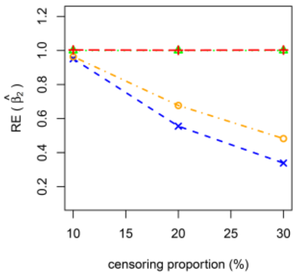
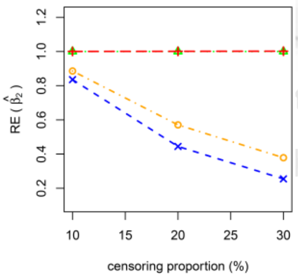
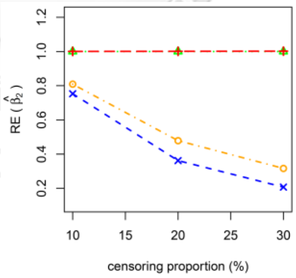
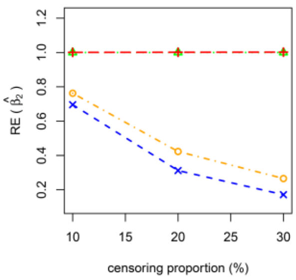
ตารางที่ 4. 34 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_\varepsilon=1:1$

$\sigma^2_{x_1+x_2}:\sigma^2_\varepsilon$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{2,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**0.04663**	0.05195	0.04837	0.05100
			1:2	**0.05707**	0.06368	0.05903	0.06286
			1:1	**0.07414**	0.08505	0.07760	0.08418
		20%	1:5	0.07236	**0.05633**	0.06539	**0.05629**
			1:2	0.08077	**0.07132**	0.07523	**0.07124**
			1:1	0.09448	**0.08975**	**0.08957**	**0.08953**
		30%	1:5	0.12077	**0.06551**	0.09422	**0.06533**
			1:2	0.12878	**0.07927**	0.10376	**0.07908**
			1:1	0.13689	**0.10287**	0.11602	**0.10264**
	50	10%	1:5	0.02978	**0.02832**	0.02938	**0.02824**
			1:2	**0.03562**	**0.03571**	**0.03554**	**0.03570**
			1:1	**0.04578**	0.04752	**0.04594**	0.04728
		20%	1:5	0.05822	**0.03232**	0.04774	**0.03228**
			1:2	0.06204	**0.03921**	0.05224	**0.03916**
			1:1	0.07177	**0.05127**	0.06263	**0.05121**
		30%	1:5	0.10521	**0.03562**	0.07379	**0.03553**
			1:2	0.10981	**0.04338**	0.07925	**0.04328**
			1:1	0.11720	**0.05617**	0.08793	**0.05606**
	100	10%	1:5	0.01986	**0.01381**	0.01811	**0.01381**
			1:2	0.02246	**0.01691**	0.02090	**0.01690**
			1:1	0.02662	**0.02223**	0.02510	**0.02222**
		20%	1:5	0.04841	**0.01510**	0.03568	**0.01508**
			1:2	0.05170	**0.01871**	0.03908	**0.01868**
			1:1	0.05512	**0.02449**	0.04292	**0.02447**
		30%	1:5	0.09835	**0.01683**	0.06338	**0.01680**
			1:2	0.10071	**0.02082**	0.06584	**0.02078**
			1:1	0.10338	**0.02630**	0.06951	**0.02626**

ตารางที่ 4. 35 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$RE\left(\hat{\beta}_{2,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:1	30	10%	1:5	**1.11**	1.00	1.07	1.02
			1:2	**1.12**	1.00	1.08	1.01
			1:1	**1.15**	1.00	1.10	1.01
		20%	1:5	0.78	**1.00**	0.86	**1.00**
			1:2	0.88	**1.00**	0.95	**1.00**
			1:1	0.95	**1.00**	**1.00**	**1.00**
		30%	1:5	0.54	**1.00**	0.70	**1.00**
			1:2	0.62	**1.00**	0.76	**1.00**
			1:1	0.75	**1.00**	0.89	**1.00**
	50	10%	1:5	0.95	**1.00**	0.96	**1.00**
			1:2	**1.00**	**1.00**	**1.00**	**1.00**
			1:1	**1.04**	1.00	**1.03**	1.00
		20%	1:5	0.56	**1.00**	0.68	**1.00**
			1:2	0.63	**1.00**	0.75	**1.00**
			1:1	0.71	**1.00**	0.82	**1.00**
		30%	1:5	0.34	**1.00**	0.48	**1.00**
			1:2	0.40	**1.00**	0.55	**1.00**
			1:1	0.48	**1.00**	0.64	**1.00**
	100	10%	1:5	0.70	**1.00**	0.76	**1.00**
			1:2	0.75	**1.00**	0.81	**1.00**
			1:1	0.84	**1.00**	0.89	**1.00**
		20%	1:5	0.31	**1.00**	0.42	**1.00**
			1:2	0.36	**1.00**	0.48	**1.00**
			1:1	0.44	**1.00**	0.57	**1.00**
		30%	1:5	0.17	**1.00**	0.27	**1.00**
			1:2	0.21	**1.00**	0.32	**1.00**
			1:1	0.25	**1.00**	0.38	**1.00**

ตารางที่ 4. 36 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:1$

$\sigma^2_{x_1}:\sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.4.2 ผลการเปรียบเทียบวิธีการประมาณค่า β_2 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:1$

จากตารางที่ 4.34 - 4.36 พบว่า

เมื่อตัวอย่างมีขนาดใหญ่ (n = 100)

วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด

เมื่อตัวอย่างมีขนาดปานกลาง (n = 50)

ส่วนใหญ่ วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่าอัตราส่วนระหว่างตัวแปรอิสระทั้งสองเป็นตัวบ่งบอกถึงวิธีการที่มีประสิทธิภาพสูงที่สุดในประมาณค่า β_2 ดังนี้

- ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่า วิธี OLS และวิธี CM มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และวิธี MLE กับวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจากวิธี OLS และวิธี CM เพียงเล็กน้อย
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย ทุกวิธีต่างมีประสิทธิภาพในการประมาณค่า β_2 เท่ากัน
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และทั้งวิธี OLS กับวิธี CM มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจาก วิธี MLE และวิธี MLE_EM เพียงเล็กน้อย

เมื่อตัวอย่างมีขนาดเล็ก (n = 30) ผลสรุปมีความหลากหลาย ดังนี้

<u>ตัวแปรตามถูกตัดปลายทางขวามาก</u> (r = 30%)

วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด

<u>ตัวแปรตามถูกตัดปลายทางขวาปานกลาง</u> (r = 20%)

- ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่า วิธี CM วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี OLS มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากอีก 3 วิธีเพียงเล็กน้อย
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย พบว่า วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจาก วิธี MLE และวิธี MLE_EM เพียงเล็กน้อย

- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก พบว่า วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด

ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%)

วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากวิธี OLS เพียงเล็กน้อย

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$

ผลสรุปเหมือนกับตารางที่ 4.31 - 4.33 ยกเว้น ปัจจัยด้านขนาดตัวอย่าง และเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา ซึ่งส่วนใหญ่ ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_2 เท่ากับวิธี MLE ยกเว้นกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ซึ่งพบว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 มากกว่าวิธี MLE เพียงเล็กน้อย

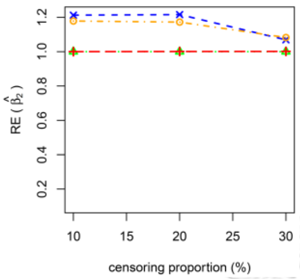
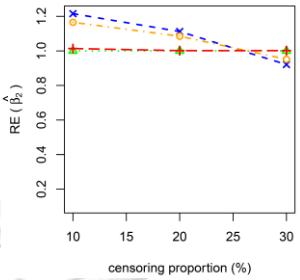
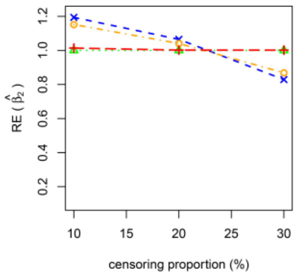
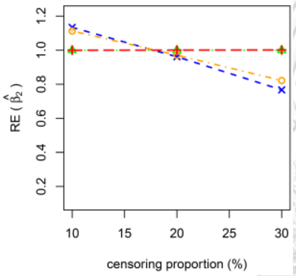
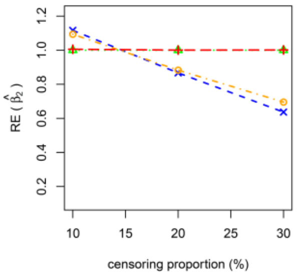
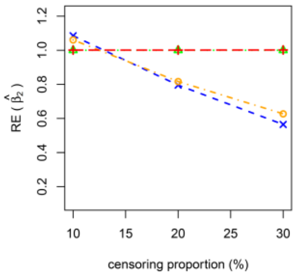
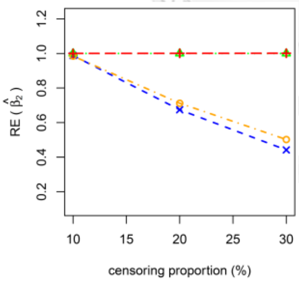
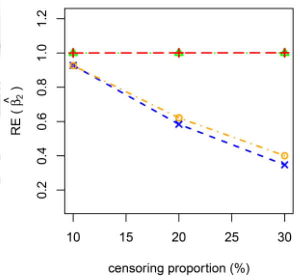
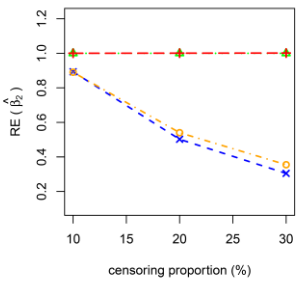
ตารางที่ 4. 37 ผลการเปรียบเทียบค่า $MSE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

$\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1}:\sigma^2_{x_2}$	$MSE\left(\hat{\beta}_{2,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**0.08230**	0.09826	0.08530	0.09692
			1:2	**0.10557**	0.12831	0.11008	0.12657
			1:1	**0.13590**	0.16476	0.13982	0.16476
		20%	1:5	**0.10394**	0.11068	0.10636	0.11047
			1:2	**0.12249**	0.13625	0.12558	0.13603
			1:1	**0.14883**	0.18094	0.15438	0.18075
		30%	1:5	0.14564	**0.12073**	0.13898	**0.12053**
			1:2	0.15743	**0.14501**	0.15239	**0.14478**
			1:1	**0.17641**	0.18856	**0.17418**	0.18828
	50	10%	1:5	**0.05000**	0.05425	0.05115	0.05420
			1:2	**0.06122**	0.06835	0.06248	0.06796
			1:1	**0.08077**	0.09161	0.08233	0.09162
		20%	1:5	0.07654	**0.06088**	0.07455	**0.06083**
			1:2	0.08466	**0.07340**	0.08313	**0.07334**
			1:1	0.10105	**0.09715**	0.10004	**0.09709**
		30%	1:5	0.11944	**0.06744**	0.10754	**0.06733**
			1:2	0.12722	**0.08101**	0.11650	**0.08089**
			1:1	0.14107	**0.10820**	0.13165	**0.10805**
	100	10%	1:5	0.02961	**0.02645**	0.02970	**0.02644**
			1:2	0.03546	**0.03290**	0.03551	**0.03290**
			1:1	**0.04365**	**0.04313**	0.04379	**0.04312**
		20%	1:5	0.05653	**0.02838**	0.05253	**0.02836**
			1:2	0.06125	**0.03576**	0.05753	**0.03574**
			1:1	0.06756	**0.04555**	0.06397	**0.04552**
		30%	1:5	0.10456	**0.03186**	0.08978	**0.03181**
			1:2	0.11038	**0.03831**	0.09567	**0.03826**
			1:1	0.11452	**0.05057**	0.10072	**0.05051**

ตารางที่ 4. 38 ผลการเปรียบเทียบค่า $RE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:2$

$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	n	r	$\sigma^2_{x_1} : \sigma^2_{x_2}$	$RE\left(\hat{\beta}_{2,a}\right)$			
				a = OLS	MLE	CM	MLE_EM
1:2	30	10%	1:5	**1.19**	1.00	1.15	1.01
			1:2	**1.22**	1.00	1.17	1.01
			1:1	**1.21**	1.00	1.18	1.00
		20%	1:5	**1.06**	1.00	1.04	1.00
			1:2	**1.11**	1.00	1.09	1.00
			1:1	**1.22**	1.00	1.17	1.00
		30%	1:5	0.83	**1.00**	0.87	**1.00**
			1:2	0.92	**1.00**	0.95	**1.00**
			1:1	**1.07**	1.00	**1.08**	1.00
	50	10%	1:5	**1.08**	1.00	1.06	1.00
			1:2	**1.12**	1.00	1.09	1.01
			1:1	**1.13**	1.00	1.11	1.00
		20%	1:5	0.80	**1.00**	0.82	**1.00**
			1:2	0.87	**1.00**	0.88	**1.00**
			1:1	0.96	**1.00**	0.97	**1.00**
		30%	1:5	0.56	**1.00**	0.63	**1.00**
			1:2	0.64	**1.00**	0.70	**1.00**
			1:1	0.77	**1.00**	0.82	**1.00**
	100	10%	1:5	0.89	**1.00**	0.89	**1.00**
			1:2	0.93	**1.00**	0.93	**1.00**
			1:1	**0.99**	**1.00**	0.98	**1.00**
		20%	1:5	0.50	**1.00**	0.54	**1.00**
			1:2	0.58	**1.00**	0.62	**1.00**
			1:1	0.67	**1.00**	0.71	**1.00**
		30%	1:5	0.30	**1.00**	0.35	**1.00**
			1:2	0.35	**1.00**	0.40	**1.00**
			1:1	0.44	**1.00**	0.50	**1.00**

ตารางที่ 4. 39 กราฟเปรียบเทียบค่า $RE\left(\hat{\beta}_{2,a}\right)$ เมื่อ $\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon} = 1:2$

$\sigma^2_{x_1} : \sigma^2_{x_2}$ / n	1:1	1:2	1:5
30			
50			
100			

4.4.3 ผลการเปรียบเทียบวิธีการประมาณค่า β_2 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$ กรณีที่ $\sigma^2_{x_1+x_2}:\sigma^2_{\varepsilon}=1:2$

จากตารางที่ 4.37 - 4.39 พบว่า

เมื่อตัวอย่างมีขนาดใหญ่ (n = 100)

วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีที่ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองกระจายเท่ากัน ซึ่งพบว่า วิธี OLS วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM มีประสิทธิภาพต่างจากอีก 3 วิธีเพียงเล็กน้อย

เมื่อตัวอย่างมีขนาดปานกลาง (n = 50)

<u>ตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก</u> (r = 20%, 30%)

วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีกระจายเท่ากัน ซึ่งแม้วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 แต่วิธี OLS และวิธี CM จะมีค่า $RE\left(\hat{\beta}_2\right)$ ที่แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย ดังนั้นจึงถือว่า วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย

<u>ตัวแปรตามถูกตัดปลายทางขวาน้อย</u> (r = 10%)

วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากวิธี OLS เพียงเล็กน้อย

เมื่อตัวอย่างมีขนาดเล็ก (n=30) ผลสรุปเป็นดังนี้

<u>ตัวแปรตามถูกตัดปลายทางขวามาก</u> (r = 30%)

- ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่า วิธี OLS และวิธี CM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย พบว่า วิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย

- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก พบว่า วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด

<u>ตัวแปรตามถูกตัดปลายทางขวาน้อยถึงปานกลาง</u> (r = 10%, 20%)

วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากวิธี OLS เพียงเล็กน้อย

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$

ผลสรุปเหมือนกับตารางที่ 4.31 - 4.33 ยกเว้นปัจจัยในด้านขนาดตัวอย่าง และปัจจัยในด้านเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา ซึ่งพบว่า ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE แสดงว่า วิธี MLE_EM นั้นมีประสิทธิภาพในการประมาณค่า β_2 เท่ากับวิธี MLE ยกเว้นกรณีที่ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกัน และกรณีที่ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย ซึ่งพบว่า วิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 มากกว่าวิธี MLE เพียงเล็กน้อย

4.4.4 สรุปผลการเปรียบเทียบวิธีการประมาณค่า β_2 เมื่อพิจารณาจากค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$

ตารางที่ 4. 40 วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2

n	r	$\sigma^2_{x_1+x_2} : \sigma^2_\varepsilon$	$\sigma^2_{x_1} : \sigma^2_{x_2}$	วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2
30	2:1	10%	1:1	MLE_EM*, CM*, MLE, OLS
			1:2	MLE*, MLE_EM*,CM*
			1:5	MLE_EM*, CM, MLE
		20%, 30%	1:1, 1:2, 1:5	MLE*, MLE_EM*

n	r	$\sigma^2_{x_1+x_2} : \sigma^2_{\varepsilon}$	$\sigma^2_{x_1} : \sigma^2_{x_2}$	วิธีที่มีประสิทธิภาพมากที่สุดและวิธีที่มีประสิทธิภาพเทียบเคียงกับวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2
30	1:1	10%	1:1, 1:2, 1:5	OLS*, CM
		20%	1:1	MLE*, MLE_EM*, CM*, OLS
			1:2	MLE*, MLE_EM*, CM
			1:5	MLE*, MLE_EM*
		30%	1:1, 1:2, 1:5	MLE*, MLE_EM*
	1:2	10%, 20%	1:1, 1:2, 1:5	OLS*, CM
		30%	1:1	OLS*, CM*
			1:2	MLE*, MLE_EM*, CM
			1:5	MLE*, MLE_EM*
50	2:1	10%, 20%, 30%	1:1, 1:2, 1:5	MLE*, MLE_EM*
	1:1	10%	1:1	CM*, OLS*, MLE, MLE_EM
			1:2	ทุกวิธี*
			1:5	MLE*, MLE_EM*, CM, OLS
		20%, 30%	1:1, 1:2, 1:5	MLE*, MLE_EM*
	1:2	10%	1:1, 1:2, 1:5	OLS*, CM
		20%	1:1	MLE*, MLE_EM*, CM, OLS
			1:2, 1:5	MLE*, MLE_EM*
		30%	1:1, 1:2, 1:5	MLE*, MLE_EM*
100	2:1, 1:1	10%, 20%, 30%	1:1, 1:2, 1:5	MLE*, MLE_EM*
	1:2	10%	1:1	MLE*, MLE_EM*, OLS*, CM
			1:2, 1:5	MLE*, MLE_EM*
		20%, 30%	1:1, 1:2, 1:5	MLE*, MLE_EM*

หมายเหตุ * คือ วิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2

เมื่อตัวอย่างมีขนาดใหญ่ (n = 100)

ส่วนใหญ่ พบว่า วิธี MLE และวิธี MLE_EM เป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน ซึ่งพบว่า วิธี OLS วิธี MLE_EM และวิธี MLE เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากอีก 3 วิธีเพียงเล็กน้อย

เมื่อตัวอย่างมีขนาดปานกลาง (n = 50)

ส่วนใหญ่ วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด ยกเว้นกรณีดังนี้

- กรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน ซึ่งวิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และพบว่า วิธี MLE กับวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจากวิธี OLS และวิธี CM เพียงเล็กน้อย
- กรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย พบว่าทุกวิธีต่างมีประสิทธิภาพในการประมาณค่า β_2 เท่ากัน
- กรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก และกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่ากับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่าวิธี MLE และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และวิธี OLS กับวิธี CM มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย
- กรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่ากับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) วิธี OLS มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM มีประสิทธิภาพในการประมาณค่าต่างจากวิธี OLS เพียงเล็กน้อย

เมื่อตัวอย่างมีขนาดเล็ก (n = 30) ผลสรุปมีความหลากหลาย ดังนี้

<u>ตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก</u> (r = 20%, 30%)

วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีการประมาณค่า β_2 ที่มีประสิทธิภาพสูงที่สุด ยกเว้น

- ในกรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย และในกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย พบว่า วิธี MLE และวิธี MLE_EM ยังคงเป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 แต่ในกรณีนี้วิธี CM มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจากวิธี MLE และวิธี MLE_EM เพียงเล็กน้อย
- ในกรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่า วิธี CM วิธี MLE และวิธี MLE_EM มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี OLS มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากอีก 3 วิธีเพียงเล็กน้อย
- ในกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่ากับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) ตัวแปรอิสระทั้งสองกระจายเท่ากัน พบว่า วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2
- ในกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่ากับตัวแปรอิสระ ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) พบว่า วิธี OLS เป็นวิธีที่มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่า β_2 ต่างจากวิธี OLS เพียงเล็กน้อย

<u>ตัวแปรตามถูกตัดปลายทางขวาน้อย</u> (r = 10%)

ในกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าหรือเท่ากับตัวแปรอิสระ พบว่า วิธี OLS มีประสิทธิภาพมากที่สุดในการประมาณค่า β_2 และวิธี CM มีประสิทธิภาพต่างจากวิธี OLS เล็กน้อย

ในกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระ อัตราส่วนระหว่างตัวแปรอิสระทั้งสองจะเป็นตัวบ่งบอกถึงวิธีการที่มีประสิทธิภาพสูงที่สุดในประมาณค่า β_2 ดังนี้

- ตัวแปรอิสระทั้งสองมีการกระจายเท่ากัน พบว่า วิธี CM และวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และวิธี OLS กับวิธี MLE มีประสิทธิภาพในการประมาณค่า β_2 แตกต่างจากวิธี CM และวิธี MLE_EM เพียงเล็กน้อย
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย พบว่า วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่า β_2 คือ วิธี CM วิธี MLE และวิธี MLE_EM
- ตัวแปรอิสระทั้งสองมีการกระจายต่างกันมาก พบว่า วิธี MLE_EM นั้นมีประสิทธิภาพสูงสุดในการประมาณค่า β_2 และวิธี MLE กับวิธี CM มีประสิทธิภาพที่แตกต่างจากวิธี MLE_EM เพียงเล็กน้อย

ปัจจัยที่มีผลต่อค่า $MSE\left(\hat{\beta}_{2,a}\right)$ และค่า $RE\left(\hat{\beta}_{2,a}\right)$

1. ขนาดตัวอย่าง

เมื่อตัวอย่างใหญ่ขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ ของทุกวิธีมีค่าลดลง นั่นคือ ทุกวิธีประมาณค่า β_2 ใกล้เคียงกับ β_2 ที่แท้จริงมากขึ้น

และพบว่า ส่วนใหญ่แล้ว ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี MLE_EM จะมีค่าเท่ากับวิธี MLE นั่นคือ วิธี MLE_EM นั้นจะมีประสิทธิภาพในการประมาณค่า β_2 เทียบเท่ากับวิธี MLE แต่ยกเว้นในกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าหรือเท่ากับตัวแปรอิสระ ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และในกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าหรือมากกว่าตัวแปรอิสระ ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายที่แตกต่างกันน้อย และกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ตัวอย่างมีขนาดเล็ก (n = 30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และตัวแปรอิสระทั้งสองมีการกระจายต่างกัน ซึ่งพบว่า วิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_2 มากกว่าวิธี MLE เพียงเล็กน้อย และเมื่อตัวอย่างมีขนาดใหญ่ขึ้น ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี OLS และวิธี CM มีค่าลดลงมาก นั่นคือ วิธี OLS และวิธี CM จะประมาณค่า β_2 ได้แย่กว่าวิธี MLE มากขึ้น

2. เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา

เมื่อตัวแปรตามถูกตัดปลายทางขวาเพิ่มขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ ของทุกวิธีมากขึ้น นั่นคือ ทุกวิธีจะประมาณค่า β_2 ได้ใกล้เคียงกับค่า β_2 ที่แท้จริงน้อยลง

นอกจากนั้น พบว่า ส่วนใหญ่ ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี MLE_EM มีค่าเท่ากับวิธี MLE นั่นคือ วิธี MLE_EM นั้นจะมีประสิทธิภาพในการประมาณค่า β_2 เท่ากับวิธี MLE ไม่ว่าตัวแปรตามถูกตัดปลายทางขวาเป็นเท่าไหร่ก็ตาม ยกเว้นในกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าหรือเท่ากับตัวแปรอิสระ ตัวอย่างมีขนาดเล็ก (n=30) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) และในกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าหรือมากกว่าตัวแปรอิสระ ตัวอย่างมีขนาดปานกลาง (n = 50) ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกันน้อย และในกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ตัวอย่างขนาดเล็ก (n = 30)

ตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) ตัวแปรอิสระทั้งสองมีการกระจายต่างกัน ซึ่งพบว่าวิธี MLE_EM มีประสิทธิภาพในการประมาณค่า β_2 มากกว่าวิธี MLE เพียงเล็กน้อย

เมื่อตัวแปรตามถูกตัดปลายทางขวามากขึ้น ค่า $MSE\left(\hat{\beta}_2\right)$ ของวิธี OLS มีค่าเพิ่มขึ้นมาก รองลงมาคือวิธี CM นั่นคือ เปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวามีผลต่อการประมาณค่า β_2 ของวิธี OLS เป็นอย่างมาก

3. อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2

เมื่อตัวแปรอิสระตัวที่ 1 กระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น ทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ ของทุกวิธีลดลง โดยเฉพาะอย่างยิ่งกรณีที่ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ ซึ่งจะเห็นการลดลงของค่า $MSE\left(\hat{\beta}_2\right)$ ที่ชัดเจนมากยิ่งขึ้น นั่นคือ อัตราส่วนความแปรปรวนของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลต่อการประมาณค่า β_2 โดยเฉพาะอย่างยิ่งเมื่อตัวแปรอิสระตัวที่ 1 มีกระจายต่างกับตัวแปรอิสระตัวที่ 2 มากขึ้น ซึ่งจะทำให้ทุกวิธีสามารถประมาณค่า β_2 ได้ใกล้เคียงกับ β_2 ที่แท้จริงเพิ่มขึ้น

4. อัตราส่วนความแปรปรวนรวมของตัวแปรอิสระต่อความแปรปรวนของความคลาดเคลื่อน

เมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระจะทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ มีค่ามากกว่ากรณีที่ความคลาดเคลื่อน มีการกระจายเท่ากันกับตัวแปรอิสระ และกรณีที่ความคลาดเคลื่อนกระจายน้อยกว่าตัวแปรอิสระ ตามลำดับ นั่นคือ อัตราส่วนของการกระจายของตัวแปรอิสระต่อการกระจายของความคลาดเคลื่อนมีผลต่อการประมาณค่า β_2 โดยยิ่งความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้นเท่าไหร่ จะทำให้ ทุกวิธีสามารถประมาณค่า β_2 ได้ใกล้เคียงกับค่า β_2 ที่แท้จริงได้น้อยลงเท่านั้น

เมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระจะทำให้ค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี OLS และวิธี CM มีค่ามากกว่ากรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากันกับตัวแปรอิสระ และกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระ ตามลำดับ นั่นคือ ยิ่งความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้นเท่าไหร่ จะทำให้วิธี OLS และวิธี CM มีประสิทธิภาพในการประมาณค่า β_2 มากกว่าวิธี MLE กับวิธี MLE_EM เพิ่มขึ้นเท่านั้น ยกเว้นกรณีที่ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) และตัวอย่างมีขนาดใหญ่ (n = 100) ซึ่งค่า $RE\left(\hat{\beta}_2\right)$ ของวิธี CM

มีค่าใกล้เคียงกันมาก หรือกล่าวได้ว่า แม้ความคลาดเคลื่อนจะกระจายมากกว่าตัวแปรอิสระมากขึ้น แต่วิธี CM จะมีประสิทธิภาพในการประมาณค่า β_2 เหมือนเดิม และพบว่าเมื่อความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระจะทำให้ค่า $MSE\left(\hat{\beta}_2\right)$ และ $RE\left(\hat{\beta}_2\right)$ ที่ได้จากวิธี OLS และวิธี CM มีค่าใกล้เคียงกันมาก มากกว่ากรณีที่ความคลาดเคลื่อนมีการกระจายเท่ากับตัวแปรอิสระ และกรณีที่ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระตามลำดับ นั่นคือ ยิ่งความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระมากขึ้นแล้วจะทำให้วิธี OLS และวิธี CM ประมาณค่า β_2 ได้ใกล้เคียงกันมากยิ่งขึ้น

บทที่ 5
การสรุปและอภิปรายผลการวิจัย และข้อเสนอแนะ

ในการประยุกต์ จะให้ความสนใจในการประมาณค่าตัวแปรตามมากกว่า ดังนั้นจึงสรุปผลจากการเปรียบเทียบค่า $RE\left(\hat{Y}_a\right)$ เป็นหลัก

5.1 การสรุปและอภิปรายผลการวิจัย

ตารางที่ 5. 2 วิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย

n \ r	10%			20%			30%		
30	1:2	1:1	2:1	1:2	1:1	2:1	1:2	1:1	2:1
50	1:2	1:1	2:1	1:2	1:1	2:1	1:2	1:1	2:1
100	1:2	1:1	2:1	1:2	1:1	2:1	1:2	1:1	2:1

OLS   CM   MLE และ MLE_EM   OLS และ CM

วิธี OLS เป็นวิธีที่คำนวณง่าย แต่วิธี OLS ไม่ได้ให้ความสำคัญกับข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 1 เพราะในการคำนวณจะถือว่าตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 เสมือนเป็นตัวแปรตามที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ดังนั้นค่าประมาณของตัวแปรตามที่ได้จะต่ำกว่าค่าแท้จริงของตัวแปรตาม ส่วนวิธี CM เป็นวิธีการที่นำวิธี OLS มาประยุกต์ใช้และมีการคำนวณที่ซับซ้อนมากกว่าวิธี OLS รวมถึงมีการให้ความสำคัญกับตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 มากกว่าวิธี OLS เพราะได้มีการจัดการกับตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ก่อนนำไปใช้ประมาณค่าพารามิเตอร์ แต่การจัดการกับตัวแปรตามไม่ได้มีความซับซ้อนในการคำนวณมากเท่ากับวิธี MLE และวิธี MLE_EM ดังนั้นถือว่า CM ใช้ข้อมูลในส่วนของตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ยังไม่ครบถ้วนสมบูรณ์ ส่วนวิธี MLE และ วิธี MLE_EM ได้ให้ความสำคัญและใช้ข้อมูล

ทั้งสองส่วน ได้แก่ ข้อมูลถูกตัดปลายทางขวาประเภทที่ 1 และข้อมูลที่ไม่ถูกตัดปลายทางขวาประเภทที่ 1 ได้ครบถ้วนสมบูรณ์ โดยวิธี MLE ถือเป็นวิธีที่มีประสิทธิภาพที่สูงแต่มีความซับซ้อนในการคำนวณ ส่วนวิธี MLE_EM เป็นวิธีที่ได้ประยุกต์ใช้วิธี MLE และมีการคำนวณง่ายกว่าวิธี MLE

เมื่อตัวอย่างมีขนาดใหญ่ (n = 100) ทุกวิธีสามารถประมาณค่าตัวแปรตามได้ใกล้เคียงกับค่าแท้จริงของตัวแปรตามมากขึ้น แต่วิธี OLS กับวิธี CM มีประสิทธิภาพในการประมาณต่ำกว่าวิธี MLE และวิธี MLE_EM และเนื่องจากวิธี MLE_EM มีการคำนวณง่ายกว่าวิธี MLE และในบางกรณีจะพบว่าวิธี MLE_EM นั้นจะมีประสิทธิภาพสูงกว่าวิธี MLE เล็กน้อย ดังนั้น ผู้วิจัยจึงเสนอให้ประมาณค่าจากตัวแบบการถดถอยด้วยวิธี MLE_EM แต่กรณีที่ความคลาดเคลื่อนกระจายมากกว่าตัวแปรอิสระและตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี OLS และวิธี CM เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าจากตัวแบบการถดถอยใกล้เคียงกับวิธี MLE และวิธี MLE_EM ซึ่งจากกรณีนี้จะเห็นได้ว่า อัตราส่วนการกระจายของตัวแปรอิสระต่อการกระจายของความคลาดเคลื่อนนั้นมีผลต่อการประมาณค่าจากตัวแบบการถดถอย โดยเฉพาะอย่างยิ่งเมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระมากขึ้น จะทำให้ทุกวิธีมีประสิทธิภาพในการประมาณค่าจากตัวแบบการถดถอยแย่ลง แต่วิธี CM และวิธี OLS จะเป็นวิธีที่มีประสิทธิภาพในการประมาณดีขึ้นเมื่อเทียบกับวิธี MLE และวิธี MLE_EM นอกจากนั้น พบว่า วิธี OLS เป็นวิธีที่มีประสิทธิภาพในการประมาณค่าจากตัวแบบการถดถอยใกล้เคียงกับวิธี CM มากยิ่งขึ้น และเมื่อตัวแปรตามถูกตัดปลายทางขวาน้อยลงทำให้ทุกวิธีประมาณค่าตัวแปรตามได้ใกล้เคียงกับค่าที่แท้จริงของตัวแปรตามมากยิ่งขึ้น หรือกล่าวได้ว่าทุกวิธีล้วนมีประสิทธิภาพในการประมาณค่าจากตัวแบบการถดถอยดีขึ้น และยิ่งตัวแปรตามถูกตัดปลายทางขวาน้อยลง ทั้งวิธี OLS และวิธี CM จะมีประสิทธิภาพในการประมาณค่าจากตัวแบบการถดถอยดีขึ้นเมื่อเทียบกับวิธี MLE และวิธี MLE_EM

เมื่อตัวอย่างมีขนาดปานกลาง (n = 50) และตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก (r = 20%, 30%) พบว่า วิธี MLE และวิธี MLE_EM ยังคงมีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย ส่วนเมื่อตัวอย่างมีขนาดปานกลาง (n = 50) และตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า

- วิธี MLE และวิธี MLE_EM มีประสิทธิภาพสูงสุดเมื่อความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระ

- วิธี CM มีประสิทธิภาพสูงสุดเมื่อความคลาดเคลื่อนมีการกระจายเท่ากับตัวแปรอิสระ
- วิธี OLS มีประสิทธิภาพสูงสุดเมื่อความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ

เมื่อตัวอย่างมีขนาดเล็ก (n = 30) พบว่า ทั้งอัตราส่วนการกระจายของความคลาดเคลื่อนต่อการกระจายของตัวแปรอิสระ และเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวาล้วนมีผลต่อการประมาณค่าจากตัวแบบการถดถอย โดยสามารถทำการพิจารณาได้จากอัตราส่วนการกระจายของความคลาดเคลื่อนต่อการกระจายของตัวแปรอิสระเป็นอันดับแรกก่อน จากนั้นจึงพิจารณาเปอร์เซ็นต์ของตัวแปรตามที่ถูกตัดปลายทางขวา ดังนั้นผลสรุปจึงเป็นดังนี้

- ความคลาดเคลื่อนมีการกระจายมากกว่าตัวแปรอิสระ เมื่อตัวแปรตามถูกตัดปลายทางขวาน้อยถึงปานกลาง (r = 10%, 20%) วิธี OLS มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอยและเมื่อตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) วิธี CM มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย
- ความคลาดเคลื่อนมีการกระจายน้อยกว่าตัวแปรอิสระ เมื่อตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี CM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย และเมื่อตัวแปรตามถูกตัดปลายทางขวาปานกลางถึงมาก (r = 20%, 30%) พบว่า วิธี MLEและวิธี MLE_EM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย
- ความคลาดเคลื่อนมีการกระจายเท่ากับตัวแปรอิสระ ในกรณีที่ตัวแปรตามถูกตัดปลายทางขวามาก (r = 30%) พบว่า วิธี MLE และวิธี MLE_EM มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย ส่วนกรณีที่ตัวแปรตามถูกตัดปลายทางขวาปานกลาง (r = 20%) พบว่า วิธี CM เป็นวิธีที่มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย แต่หากตัวแปรตามถูกตัดปลายทางขวาน้อย (r = 10%) พบว่า วิธี OLS มีประสิทธิภาพสูงสุดในการประมาณค่าจากตัวแบบการถดถอย

นอกจากนั้นพบว่า อัตราส่วนการกระจายของตัวแปรอิสระตัวที่ 1 ต่อตัวแปรอิสระตัวที่ 2 มีผลน้อยมากต่อการประมาณค่าจากตัวแบบการถดถอย

5.2 ข้อเสนอแนะ

การวิจัยครั้งนี้ได้ทำการศึกษาภายใต้ขอบเขตของข้อมูลที่ตัวแปรตามที่ถูกตัดปลายทางขวาประเภทที่ 1 ที่มาจากการแจกแจงแบบล็อกนอร์มอล โดยตัวแปรตามถูกตัดปลายทางขวาเท่ากับ 10%, 20% และ 30% ขนาดตัวอย่างเท่ากับ 30, 50, 100 ตัวแปรอิสระ 2 ตัว ไม่มีความสัมพันธ์กัน ดังนั้นเพื่อเป็นแนวทางในการทำการศึกษาเพิ่มเติมนอกเหนือจากขอบเขตของการวิจัยในครั้งนี้ สามารถทำการศึกษากรณีอื่นๆ ดังนี้

1) ตัวแปรตามมีการแจกแจงเป็นแบบอื่น เช่น แกมมา เอ็กซ์โปเนนเชียล เป็นต้น
2) ตัวแปรตามเป็นข้อมูลแบบอื่น เช่น ข้อมูลที่ถูกตัดปลายทางแบบสุ่ม ข้อมูลที่ถูกตัดแบบช่วง ข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 เป็นต้น
3) ตัวแปรอิสระมีมากกว่า 2 ตัว
4) ตัวแปรอิสระมีความสัมพันธ์กันในระดับต่างๆ
5) เพิ่มระดับความแปรปรวนระหว่างตัวแปรอิสระรวมกับความคลาดเคลื่อน

บรรณานุกรม

ภาษาอังกฤษ

Aitkin, M. 1981. A Note on the Regression Analysis of Censored Data. Technometrics 23 (May 1981) : 161-163.

Chatterjee, S., and McLeish, D.L. 1986. Fitting Linear Regression to Censored Data by Least Squares and Maximum Likelihood Methods. Communications in Statistics - Theory and Methods 15 (February 1986) : 3227-3243.

Dempster, A.P., Laird, N.M., and Rubin, D.B. 1977. Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society Series B (Methodological) 39 (January 1977) : 1-38.

Gauss, C.F. 1809. Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium. [Online]. Retrieved from : https://archive.org/details/bub_gb_ORUOAAAAQAAJ [2017, November 11].

Jöreskog, K.G. 2002. Censored Variables and Censored Regression. [Online]. Retrieved from : http://www.ssicentral.com/lisrel/techdocs/censor.pdf [2017, November 11].

Klein, J.P., and Moeschberger, M.L. 2003. Survival Analysis Techniques for Censored and Truncated Data. Second Edition. New York : Springer-Verlag New York, Inc.

Tobin, J. 1958. Estimation of Relationships for Limited Dependent Variables. Econometrica 26 (January 1958) : 24-36.

ภาษาไทย

เกศินี พรมธ. 2555. การวิเคราะห์เพื่อปรับปรุงการใช้พลังงานของห้องสะอาดด้วยวิธีการเชิงสถิติ กรณีศึกษาอุตสาหกรรมผลิต วอยซ์ คอยล์ มอเตอร์. วารสารวิจัยพลังงาน 9 (กันยายน – ธันวาคม 2555) : 24-35.

จำเนียน จำนงค์รักษ์. 2539. การพยากรณ์ในสมการถดถอยเชิงเส้นพหุเมื่อค่าตัวแปรตามถูกตัดปลายทางขวา. วิทยานิพนธ์ปริญญามหาบัณฑิต, สาขาวิชาการประกันภัย ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ธีระพร วีระถาวร. 2536. การอนุมานเชิงสถิติขั้นกลาง: โครงสร้างและความหมาย. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ธีระพร วีระถาวร. 2537. ความน่าจะเป็นกับการประยุกต์. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร : สำนักพิมพ์อักษรกราฟฟิค.

บังอร กุมพล. 2539. การวิเคราะห์การถดถอยเชิงเส้นพหุ เมื่อตัวแปรตามมีค่าถูกตัดทิ้งทางขวากรณีค่าตัดทิ้งประเภทที่ 1. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ สาขาวิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

สุพล ดุรงค์วัฒนา. 2558. Regression Models: Analytics-based Approach. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร : บริษัทแดเน็กซ์ อินเตอร์คอร์ปอเรชั่น จำกัด.

ภาคผนวก

รายละเอียดของโปรแกรมที่ใช้ในงานวิจัย

ในการวิจัยครั้งนี้ได้ใช้โปรแกรม R เวอร์ชัน 3.4.4 ในการจำลองข้อมูลและการประมาณจากตัวแบบการถดถอยจำนวน 4 วิธี ซึ่งมีคำสั่งดังต่อไปนี้

```
library(censReg)
Variance<-data.frame(var_x1=rep(c(0.015,0.030,0.045),3), var_x2=rep(c(0.075,0.060,
0.045),3), var_error=c(rep(0.045,3),rep(0.090,3),rep(0.180,3)))
N<-c(30,50,100)
R<-c(0.9,0.8,0.7)
beta0<-0.3
beta1<-1
beta2<-1

nloop<-10000
for( i in 1:nrow(Variance)){
for( n in N){
for( r in R){
A<-c()
B<-c()
C<-c()
NN<-c()
RR<-c()
TOTAL_betahat0_OLS<-c()
TOTAL_betahat1_OLS<-c()
TOTAL_betahat2_OLS<-c()
TOTAL_betahat0_MLE<-c()
TOTAL_betahat1_MLE<-c()
TOTAL_betahat2_MLE<-c()
TOTAL_betahat0_CM<-c()
TOTAL_betahat1_CM<-c()
TOTAL_betahat2_CM<-c()
```

```
TOTAL_betahat0_MLE_EM<-c()
TOTAL_betahat1_MLE_EM<-c()
TOTAL_betahat2_MLE_EM<-c()
TOTAL_MSE_OLS<-c()
TOTAL_MSE_CM<-c()
TOTAL_MSE_MLE<-c()
TOTAL_MSE_MLE_EM<-c()

for ( k in 1:nloop){
a<-Variance[i,1]
b<-Variance[i,2]
c<-Variance[i,3]

###Generate Data###
function_generate<-function(a,b,c,n,r){
x1<-rnorm(n,0,sqrt(a))
x2<-rnorm(n,0,sqrt(b))
error<-rnorm(n,0,sqrt(c))
ytrue<-exp(0.3+x1+x2)
yobs<-exp(0.3+x1+x2+error)
cen<-qlnorm(r,0.3,sqrt(a+b+c))
logcen<-log(cen)
ystar<-apply(cbind(yobs,cen),1,min)
logystar<-log(ystar)
specifycen<-as.numeric(yobs>cen)
datafull<-data.frame(x1=x1,x2=x2,ytrue=ytrue,specifycen=specifycen,
logcen=rep(logcen,n),logystar=logystar)
return(datafull)
}
```

```
myfun<-function(a,b,c,n,r){
out<-function_generate(a,b,c,n,r)
while(sum(out$specifycen==0)<3){out<-myfun(a,b,c,n,r)}
out
}
data1<-myfun(a,b,c,n,r)

###OLS###
OLS<-lm(logystar ~ x1 + x2,data=data1)

###MSE_OLS###
betahat0_OLS<-as.numeric(coef(OLS)[1])
betahat1_OLS<-as.numeric(coef(OLS)[2])
betahat2_OLS<-as.numeric(coef(OLS)[3])
logyhat_OLS<-betahat0_OLS+(betahat1_OLS*data1[1])+(betahat2_OLS*data1[2])
yhat_OLS<-exp(logyhat_OLS)
mse_OLS<-mean((yhat_OLS-data1[3])^2)
TOTAL_betahat0_OLS<-c(TOTAL_betahat0_OLS,betahat0_OLS)
TOTAL_betahat1_OLS<-c(TOTAL_betahat1_OLS,betahat1_OLS)
TOTAL_betahat2_OLS<-c(TOTAL_betahat2_OLS,betahat2_OLS)
TOTAL_MSE_OLS<-c(TOTAL_MSE_OLS,mse_OLS)

###MLE###
MLE<-function(data1){
if(sum(data1$specifycen==0)<n){
mle<-censReg(logystar ~ x1 + x2,data=data1 , left = -Inf, right = data1[1,5])
return(mle)}
else{
mle<-lm(logystar ~ x1 + x2,data=data1)
return(mle)}
}
```

```
MLE<-MLE(data1)

###MSE_MLE###
betahat0_MLE<-as.numeric(coef(MLE)[1])
betahat1_MLE<-as.numeric(coef(MLE)[2])
betahat2_MLE<-as.numeric(coef(MLE)[3])
logyhat_MLE<-betahat0_MLE+(betahat1_MLE*data1[1])+(betahat2_MLE*data1[2])
yhat_MLE<-exp(logyhat_MLE)
mse_MLE<-mean((yhat_MLE-data1[3])^2)
TOTAL_betahat0_MLE<-c(TOTAL_betahat0_MLE,betahat0_MLE)
TOTAL_betahat1_MLE<-c(TOTAL_betahat1_MLE,betahat1_MLE)
TOTAL_betahat2_MLE<-c(TOTAL_betahat2_MLE,betahat2_MLE)
TOTAL_MSE_MLE<-c(TOTAL_MSE_MLE,mse_MLE)

###CM####
function_new_betahat_CM<-function(old_betahat_CM){
logyhat_censor_CM<-old_betahat_CM[1]+(old_betahat_CM[2]*data_censor[1])
+(old_betahat_CM[3]*data_censor[2])

###logyhat_censor_CM or logcen###
newlogy_censor_CM<-apply(cbind(data_censor[5],logyhat_censor_CM),1,max)
data2<-data.frame(data_censor[,1:5],logystar=newlogy_censor_CM)
data3<-rbind(data_nocensor,data2)
CM<-lm(logystar ~ x1+x2,data=data3)
new_betahat_CM<-as.numeric(coef(CM))
return(new_betahat_CM)
}

data_nocensor<-subset(data1, specifycen ==0)
data_censor<-subset(data1, specifycen ==1)
CM_begin<-lm(logystar ~ x1 + x2,data=data_nocensor)
```

```
old_betahat_CM<-coef(CM_begin)
new_betahat_CM<-function_new_betahat_CM(old_betahat_CM)
max_ab_diff_CM<-max(abs(old_betahat_CM-new_betahat_CM))

while(max_ab_diff_CM>0.001){
old_betahat_CM<-new_betahat_CM
new_betahat_CM<-function_new_betahat_CM(old_betahat_CM)
max_ab_diff_CM<-max(abs(old_betahat_CM-new_betahat_CM))
print(new_betahat_CM)
}

###MSE_CM###
betahat0_CM<-new_betahat_CM[1]
betahat1_CM<-new_betahat_CM[2]
betahat2_CM<-new_betahat_CM[3]
logyhat_CM<-betahat0_CM+(betahat1_CM*data1[1])+(betahat2_CM*data1[2])
yhat_CM<-exp(logyhat_CM)
mse_CM<-mean((yhat_CM-data1[3])^2)
TOTAL_betahat0_CM<-c(TOTAL_betahat0_CM,betahat0_CM)
TOTAL_betahat1_CM<-c(TOTAL_betahat1_CM,betahat1_CM)
TOTAL_betahat2_CM<-c(TOTAL_betahat2_CM,betahat2_CM)
TOTAL_MSE_CM<-c(TOTAL_MSE_CM,mse_CM)

###MLE_EM###
function_new_MLE_EM<-function(old_betahat_MLE_EM,old_sigmahat_MLE_EM){
### E-Step###
muhat_censore<-old_betahat_MLE_EM[1]+(old_betahat_MLE_EM[2]*data_censor[1])
+(old_betahat_MLE_EM[3]*data_censor[2])
muhat_nocensore<-old_betahat_MLE_EM[1]+(old_betahat_MLE_EM[2]*
data_nocensor[1])+(old_betahat_MLE_EM[3]*data_nocensor[2])
zhat<-as.matrix((data_censor[5]-muhat_censore)/old_sigmahat_MLE_EM)
```

```
density_funtion<-(1/(sqrt(2*pi)))*exp(-(zhat^2)/2)
survival_function<-pnorm(zhat, mean = 0, sd = 1, lower.tail = FALSE)
hazard_function<-(density_funtion/survival_function)
log_yhat_MLE_EM_censor<-muhat_censore+(old_sigmahat_MLE_EM*hazard_function)
data4<-data.frame(data_censor[,1:5],logystar = log_yhat_MLE_EM_censor)
colnames(data4)<-c('x1','x2','ytrue','specifycen','logcen','logystar')
data5<-rbind(data_nocensor,data4)
###M-Step ###
EM<-lm(logystar ~ x1 + x2,data=data5)
new_betahat_MLE_EM<-as.numeric(coef(EM))
new_sigmahat_MLE_EM<-sqrt((sum((data_nocensor[6]-muhat_nocensore)^2)
+((old_sigmahat_MLE_EM^2)*sum(1+(zhat*hazard_function))))/n)
new<-c(new_betahat_MLE_EM,new_sigmahat_MLE_EM)
return(new)
}

MLE_EM_begin<-lm(logystar ~ x1 + x2,data=data1)
old_betahat_MLE_EM<-coef(MLE_EM_begin)
old_sigmahat_MLE_EM<-summary(MLE_EM_begin)$sigma
NEW<-function_new_MLE_EM(old_betahat_MLE_EM,old_sigmahat_MLE_EM)
new_betahat_MLE_EM<-NEW[1:3]
new_sigmahat_MLE_EM<-NEW[4]
max_ab_diff_MLE_EM<-max(abs(old_betahat_MLE_EM-new_betahat_MLE_EM))

while(max_ab_diff_MLE_EM>0.001){
old_betahat_MLE_EM<-new_betahat_MLE_EM
old_sigmahat_MLE_EM<-new_sigmahat_MLE_EM
new_betahat_sigmahat<-function_new_MLE_EM(old_betahat_MLE_EM,
old_sigmahat_MLE_EM)
new_betahat_MLE_EM<-new_betahat_sigmahat[1:3]
new_sigmahat_MLE_EM<-new_betahat_sigmahat[4]
```

```
max_ab_diff_MLE_EM<-max(abs(old_betahat_MLE_EM-new_betahat_MLE_EM))
print(new_betahat_sigmahat)
}

###MSE_MLE_EM###
betahat0_MLE_EM<-new_betahat_sigmahat[1]
betahat1_MLE_EM<-new_betahat_sigmahat[2]
betahat2_MLE_EM<-new_betahat_sigmahat[3]
logyhat_MLE_EM<-betahat0_MLE_EM+(betahat1_MLE_EM*data1[1])+
(betahat2_MLE_EM*data1[2])
yhat_MLE_EM<-exp(logyhat_MLE_EM)
mse_MLE_EM<-mean((yhat_MLE_EM-data1[3])^2)
TOTAL_betahat0_MLE_EM<-c(TOTAL_betahat0_MLE_EM,betahat0_MLE_EM)
TOTAL_betahat1_MLE_EM<-c(TOTAL_betahat1_MLE_EM,betahat1_MLE_EM)
TOTAL_betahat2_MLE_EM<-c(TOTAL_betahat2_MLE_EM,betahat2_MLE_EM)
TOTAL_MSE_MLE_EM<-c(TOTAL_MSE_MLE_EM,mse_MLE_EM)
A<-c(A,a)
B<-c(B,b)
C<-c(C,c)
NN<-c(NN,n)
RR<-c(RR,(1-r)*100)

###COUNT TIME###
pie(c(k,nloop-k),c(k,nloop-k),radius=1,main=paste(a,b,c,n,(1-r)*100,sep=','),clockwise=T)
} # end loop total

DATA<-data.frame(A,B,C,NN,RR,TOTAL_betahat0_OLS,TOTAL_betahat1_OLS,
TOTAL_betahat2_OLS,TOTAL_MSE_OLS,TOTAL_betahat0_MLE,TOTAL_betahat1_MLE,
TOTAL_betahat2_MLE,TOTAL_MSE_MLE,TOTAL_betahat0_CM,TOTAL_betahat1_CM,
TOTAL_betahat2_CM,TOTAL_MSE_CM,TOTAL_betahat0_MLE_EM,TOTAL_betahat1_
MLE_EM,TOTAL_betahat2_MLE_EM,TOTAL_MSE_MLE_EM)
```

```
colnames(DATA)<-c('variance x1','variance x2','variance error','n','r','betahat0 OLS',
'betahat1 OLS','betahat2 OLS','MSE OLS','betahat0 MLE','betahat1 MLE','betahat2 MLE',
'MSE MLE','betahat0 CM','betahat1 CM','betahat2 CM','MSE CM','betahat0 MLE_EM',
'betahat1 MLE_EM','betahat2 MLE_EM','MSE MLE_EM')
write.csv(DATA,paste('output1','var x1 =',a,',var x2 =',b,',var error =',c,',n =',n,',r =',
(1-r)*100,'.csv',sep=' '),row.names=F)
}}}
###Find MSE(betahat0),  MSE(betahat1),  MSE(betahat2), AMSE(yhat) ###
DATA<-c()

for( i in 1:nrow(Variance)){
for( n in N){
for( r in R){
a<-Variance[i,1]
b<-Variance[i,2]
c<-Variance[i,3]

total_data<-read.csv(paste('output1','var x1 =',a,',var x2 =',b,',var error =',c,',n =',n,',
r =',(1-r)*100,'.csv',sep=' '))
data_OLS<-data.frame(total_data[6:9],"OLS")
data_MLE<-data.frame(total_data[10:13],"MLE")
data_CM<-data.frame(total_data[14:17],"CM")
data_MLE_EM<-data.frame(total_data[18:21],"MLE_EM")

function_dat<-function(data){
variance_x1<-a
variance_x2<-b
variance_error<-c
variance_x<-a+b
variance_x1.variance_x2<-a/b
variance_x.variance_error<-(a+b)/c
```

```
method<-data[1,5]
bias_betahat0<-mean(as.matrix(data[1]))-beta0
bias_betahat1<-mean(as.matrix(data[2]))-beta1
bias_betahat2<-mean(as.matrix(data[3]))-beta2
var_betahat0<-as.numeric(var(data[1]))
var_betahat1<-as.numeric(var(data[2]))
var_betahat2<-as.numeric(var(data[3]))
mse_betahat0<-var_betahat0+((bias_betahat0)^2)
mse_betahat1<-var_betahat1+((bias_betahat1)^2)
mse_betahat2<-var_betahat2+((bias_betahat2)^2)
AMSE_yhat<-mean(as.matrix(data[4]))
dat<-data.frame(n,100*(1-r),variance_x1,variance_x2,variance_error,variance_x,
variance_x1.variance_x2,variance_x.variance_error,method,mse_betahat0,
mse_betahat1,mse_betahat2,AMSE_yhat)
colnames(dat)<-
c('n','r','variance_x1','variance_x2','variance_error','variance_x','variance_x1.variance_x2','
variance_x.variance_error','method','mse_betahat0','mse_betahat1','mse_betahat2',
'AMSE_yhat')
return(dat)
}
DATA<-rbind(DATA,function_dat(data_OLS),function_dat(data_MLE),
function_dat(data_CM),function_dat(data_MLE_EM))
}}}
write.csv(DATA,paste('output2','.csv'),row.names = FALSE)

###Find RE(betahat0), RE(betahat1), RE(betahat2), RE(yhat)###
total_data<-read.csv('output2 .csv')
data_OLS<-subset(total_data,method=="OLS")
data_MLE<-subset(total_data,method=="MLE")
data_CM<-subset(total_data,method=="CM")
data_MLE_EM<-subset(total_data,method=="MLE_EM")
```

```
function_RE<-function(data,data_MLE){
dat<-c('n','r','variance_x1','variance_x2','variance_error','variance_x',
'variance_x1.variance_x2','variance_x.variance_error','method','mse_betahat0',
'mse_betahat1','mse_betahat2','AMSE_yhat')

A<-c()
for(k in 10:13){
a<-data_MLE[k]/data[k]
A<-c(A,a)}
dat<-data.frame(data[dat][1:9],A)
colnames(dat)<-c('n','r','variance_x1','variance_x2','variance_error','variance_x',
'variance_x1.variance_x2','variance_x.variance_error','method','RE_mse_betahat0',
'RE_mse_betahat1','RE_mse_betahat2','RE_AMSE_yhat')
return(dat)
}

complete_data<-rbind(function_RE(data_OLS,data_MLE),function_RE(data_MLE,
data_MLE),function_RE(data_CM,data_MLE),function_RE(data_MLE_EM,data_MLE))
write.csv(complete_data,paste('output3','.csv',sep=''),row.names = FALSE)
```

ประวัติผู้เขียน

ชื่อ-สกุล	นางสาวศิวพร ทิพย์พันธุ์
วัน เดือน ปี เกิด	30 พฤษภาคม 2536
สถานที่เกิด	โรงพยาบาลจังหวัดน่าน
วุฒิการศึกษา	B.Ed.
ที่อยู่ปัจจุบัน	จงสิริอพาร์ทเม้นท์ ซอยอร่ามศรี ถนนพญาไท แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพมหานคร 10400