วิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC สำหรับข้อมูลชุดเดียวกัน: กรณีศึกษาแบบจำลองคะแนนเครดิต

นางสาวเบญจพร เอี่ยมประโคน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2560

A PROPOSED METHOD TO COMPARE AREAS UNDER THE ROC CURVES FOR A SINGLE DATASET: A CASE STUDY OF CREDIT SCORING MODEL

Miss Benchaporn Iamprakhon

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Statistics
Department of Statistics
Faculty of Commerce and Accountancy
Chulalongkorn University
Academic Year 2017

หัวข้อวิทยานิพนธ์	วิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC สำหรับข้อมูลชุดเดียวกัน: กรณีศึกษาแบบจำลองคะแนนเครดิต
โดย	นางสาวเบญจพร เอี่ยมประโคน
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	อาจารย์ ดร. ณัตติฤดี เจริญรักษ์

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

....คณบดีคณะพาณิชยศาสตร์และการบัญชี

(รองศาสตราจารย์ ดร. พสุ เดชะรินทร์)

คณะกรรมการสอบวิทยานิพนธ์

....ประธานกรรมการ

(อาจารย์ ดร.อัครินทร์ ไพบูลย์พานิช)

....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(อาจารย์ ดร. ณัตติฤดี เจริญรักษ์)

....กรรมการ

(รองศาสตราจารย์ ดร.เสกสรร เกียรติสุไพบูลย์)

....กรรมการภายนอกมหาวิทยาลัย

(อาจารย์ ดร. อรุณี กำลัง)

เบญจพร เอี่ยมประโคน : วิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC สำหรับข้อมูลชุดเดียวกัน: กรณีศึกษาแบบจำลองคะแนนเครดิต (A PROPOSED METHOD TO COMPARE AREAS UNDER THE ROC CURVES FOR A SINGLE DATASET: A CASE STUDY OF CREDIT SCORING MODEL) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: อ. ดร. ณัตติฤดี เจริญรักษ์, 56 หน้า.

การเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปเป็นวิธีที่ใช้กันอย่างแพร่หลาย ซึ่งวิธีที่นิยมใช้ในการเปรียบเทียบ คือ วิธี Delong แต่วิธีนี้ควรนำมาใช้เมื่อผลต่างของพื้นที่ ROC มีการแจกแจงปกติ งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อหาวิธีเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปเมื่อผลต่างของพื้นที่ ROC ไม่มีการแจกแจงแบบปกติ โดยทำการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปและตัวแบบลดรูปให้มีการแจกแจงปกติแล้วนำไปทดสอบผลต่างด้วย Z-test (วิธี Transform)

นอกจากนี้ผู้วิจัยยังได้เปรียบเทียบประสิทธิภาพของการเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างวิธี Transform และวิธี Delong test โดยทำการจำลองข้อมูลจาก German credit ใน package caret ในโปรแกรม R เพื่อนำมาสร้างตัวแบบโลจิสติกสำหรับการเปรียบเทียบพื้นที่ใต้โค้งของทั้ง 2 วิธี และนำผลของทั้ง 2 วิธีในแต่ละขนาดตัวอย่างมาหาค่าสัมประสิทธิ์สหสัมพันธ์กับผลการวิเคราะห์จากวิธี Likelihood ratio test (LRT) ซึ่งเป็นอีกวิธีที่ใช้เปรียบเทียบตัวแบบโลจิสติก อนึ่งวิธี LRT นี้ไม่สามารถใช้เปรียบเทียบตัวแบบทางสถิติบางตัวแบบ เช่น Support Vector Machine แต่สามารถทำการเปรียบเทียบตัวแบบได้โดยการเปรียบเทียบพื้นที่ใต้โค้ง ROC ดังนั้นการเปรียบเทียบพื้นที่ใต้โค้ง ROC จึงเป็นวิธีหนึ่งในการช่วยเลือกตัวแบบที่เหมาะสม

ผลการวิจัยพบว่า ที่ขนาดตัวอย่าง 300 500 และ 1000 วิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC โดยวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC (วิธี Transform) มีค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ที่สอดคล้องกับวิธี LRT มากกว่าวิธี Delong อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 นอกจากนั้นพบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value มีค่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05

ภาควิชา สถิติ ลายมือชื่อนิสิต ....

สาขาวิชา สถิติ ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

ปีการศึกษา 2560

5981559026 : MAJOR STATISTICS

KEYWORDS: ROC CURVE / CREDIT SCORING / LOGISTIC REGRESSION MODEL

BENCHAPORN IAMPRAKHON: A PROPOSED METHOD TO COMPARE AREAS UNDER THE ROC CURVES FOR A SINGLE DATASET: A CASE STUDY OF CREDIT SCORING MODEL. ADVISOR: NUTTIRUDEE CHAROENRUK, Ph.D., 56 pp.

Comparing areas under the ROC curves between full and reduced model has been widely used. The most commonly used method of comparison is Delong test. However, this method should be used when the difference of areas under the ROC curves is normally distributed. The purpose of this research is to find a method comparing areas under the ROC curves between full and reduced model when the difference of areas under the ROC curves is not normal distribution. In this research, the method is done by transforming the difference of areas under the ROC curves between full and reduced model to be normal distributed and then used Z-test to compare the difference (Transform method).

Additionally, we compared our Transform method with Delong test by simulated data from German credit in the Caret package in R program. To compare these two methods, we estimated logistic regression models and computed correlation coefficient between p-values from these two methods and p-values from the Likelihood ratio test (LRT) which is another method used to compare two models in the logistics model. LRT cannot be used to compare models in some statistical models such as Support Vector Machine. Nevertheless, we can compare the models by using area of ROC curve. Therefore, comparing areas under the ROC curves is one way to help us choose a proper model.

The research has revealed that at the different sample sizes including 300, 500 and 1,000, the transform method has higher correlation of p-values which correspond to LRT more than Delong test at significance level of 0.05. Moreover, The correlation of p-values are not significantly different at the significance level of 0.05.

Department:	Statistics	Student's Signature	....
Field of Study:	Statistics	Advisor's Signature	....
Academic Year:	2017		

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดีจากการดูแลและให้ความช่วยเหลือของ อาจารย์ ดร.ณัตติฤดี เจริญรักษ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ อาจารย์ ดร.อัครินทร์ ไพบูลย์พานิช ประธานกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.เสกสรร เกียรติสุไพบูลย์ และ อาจารย์ ดร.อรุณี กำลัง กรรมการสอบวิทยานิพนธ์ ที่ได้ให้คำปรึกษา ความรู้ คำแนะนำ ตลอดจนการเอาใจใส่ในการปรับปรุงงาน ตรวจสอบ แก้ไขข้อบกพร่องต่างๆ และส่งเสริมให้กำลังใจเป็นอย่างดีเสมอมา ผู้วิจัยขอกราบขอบระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณคณาจารย์ในภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัยทุกท่านที่ได้กรุณาประสิทธิ์ประสาทวิชาความรู้ทางคณิตศาสตร์และสถิติ ทำให้ผู้วิจัยสามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ให้เป็นประโยชน์สูงสุด และขอกราบขอบพระคุณบุคลากรทุกท่านในภาควิชาสถิติที่ได้อำนวยความสะดวกในด้านเอกสารและการประสานงานต่างๆ

สุดท้ายนี้ขอกราบขอบพระคุณ บิดา มารดา ญาติพี่น้องของผู้วิจัย ที่คอยให้กำลังใจและส่งเสริมสนับสนุนด้านการเรียนด้วยดีมาโดยตลอด รวมทั้งเพื่อนๆและทุกคนที่มีส่วนเกี่ยวข้องกับการวิจัยครั้งนี้ที่ได้ให้คำปรึกษา และกำลังใจตลอดระยะเวลาในการทำวิจัยได้เป็นอย่างดี

สารบัญ

หน้า

หน้า

สารบัญตาราง

สารบัญภาพ

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันงานวิจัยด้านต่างๆเกือบทุกแขนงทั้งทางด้านเศรษฐศาสตร์ สังคมศาสตร์ วิทยาศาสตร์ การแพทย์ การเงิน ได้นำเทคนิคการพยากรณ์มาใช้ด้วยกันหลายวิธี ตามลักษณะของข้อมูลและจุดประสงค์ของการพยากรณ์ วิธีการหนึ่งที่นิยมในปัจจุบัน คือ ตัวแบบการถดถอยโลจิสติก (logistic regression model) ซึ่งเป็นการวิเคราะห์ที่ตัวแปรตามเป็นข้อมูลเชิงคุณภาพ ส่วนตัวแปรต้นสามารถเป็นได้ทั้งข้อมูลเชิงคุณภาพและข้อมูลเชิงปริมาณ โดยตัวแบบการถดถอยโลจิสติกแบ่งเป็น 2 ประเภท คือ ตัวแบบการถดถอยโลจิสติกทวิ (binary logistic regression model) สำหรับกรณีที่ตัวแปรตามแบ่งเป็น 2 กลุ่ม และตัวแบบการถดถอยโลจิสติกพหุกลุ่ม (multinomial logistic regression model) สำหรับกรณีที่ตัวแปรตามที่มีหลายค่ามากกว่า 2 กลุ่ม

ตัวแบบการถดถอยโลจิสติกทวิ (binary logistic regression model) เป็นตัวแบบถดถอยที่ตัวแปรตามมีค่าเพียง 2 ค่าคือ 0 และ 1 ส่วนตัวแปรอิสระอาจมีหนึ่งตัวแปรหรือหลายตัวแปรก็ได้ ซึ่งตัวแบบการถดถอยโลจิสติกทวิเป็นตัวแบบที่ได้รับความนิยมใช้กันมาก เช่น ในทางด้านการเงิน คือ การวิเคราะห์คะแนนเครดิตของลูกค้าว่าลูกค้าเป็นลูกหนี้ที่มีปัญหาหรือเป็นลูกหนี้ที่ดี ขั้นตอนที่สำคัญอีกขั้นตอนหนึ่งในการพยากรณ์ คือ การเลือกตัวแบบที่มีประสิทธิภาพ การวัดประสิทธิภาพของตัวแบบสามารถทำได้โดยวัดอัตราความถูกต้องของการพยากรณ์ซึ่งมีอยู่หลายวิธี โดยวิธีหนึ่งที่นิยมใช้ คือ การใช้ Receiver Operating Characteristic curve (ROC curve) ค่าที่ถูกใช้ในการบ่งบอกความถูกต้องของการพยากรณ์คือ พื้นที่ใต้โค้ง ROC หรือที่เรียกว่า AUC (Area under the ROC curve) ถ้าพื้นที่ใต้โค้ง ROC มีค่ามาก แสดงว่าตัวแบบนั้นมีความถูกต้องมากเช่นกัน

เส้นโค้ง ROC ถูกใช้ครั้งแรกในช่วงสงครามโลกครั้งที่สองสำหรับการวิเคราะห์สัญญาณเรดาร์ในปี ค.ศ. 1941 (Green & Swets, 1966) โดยการพล็อตกราฟระหว่างค่า Sensitivity และ 1-specificity โดย แกน y แทน Sensitivity และแกน x แทน 1-specificity ที่จุดตัดต่างๆเพื่อแบ่งผลลัพธ์ของการพยากรณ์ออกเป็นกลุ่มเหตุการณ์ที่สนใจและเหตุการณ์ที่ไม่ได้สนใจ (Hanley & McNeil, 1982) เช่น ในคะแนนเครดิต จะแบ่งออกเป็นกลุ่มคนที่ชำระหนี้ตรงตามเวลา และกลุ่มคนที่มีการผิดนัดชำระหนี้ DeLong (1988) ได้ทำการศึกษาเกี่ยวกับการหาพื้นที่ใต้โค้ง ROC และทำการ

เปรียบเทียบพื้นที่ใต้โค้ง ROC ของตัวแบบ แต่วิธีการเปรียบเทียบนี้ไม่ควรใช้เมื่อผลต่างของพื้นที่ใต้โค้ง ROC ของตัวแบบไม่เป็นการแจกแจงแบบปกติ (Demler, Pencina, & D'Agostino, 2012) แต่อย่างไรก็ตามวิธีการศึกษาถึงความแตกต่างระหว่างพื้นที่ใต้โค้ง ROC ว่ามีความแตกต่างอย่างมีนัยสำคัญหรือไม่นั้นยังมีอยู่จำนวนน้อย

ผู้วิจัยสนใจที่จะทำการศึกษาวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC จากตัวแบบการถดถอยโลจิสติก 2 ตัวแบบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ โดย 2 ตัวแบบคือ ตัวแบบเต็มรูป (full model) และ ตัวแบบลดรูป (reduced model) ซึ่งเป็นแบบจำลองที่ตัวแบบซ้อนกัน (nested model) กัน โดยข้อมูลในการสร้างตัวแบบมาจากข้อมูลชุดเดียวกัน และพื้นที่ใต้โค้ง ROC นั้นไม่มีการแจกแจงแบบปกติ ทำการเปรียบเทียบพื้นที่ใต้โค้งด้วยวิธี Delong test แบบปกติ กับวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ให้มีการแจกแจงแบบปกติแล้วนำไปทดสอบด้วย Z-test นอกจากนี้งานวิจัยนี้ได้ทำการทดสอบว่าขนาดตัวอย่างที่แตกต่างกันจะทำให้ผลการทดลองของทั้ง 2 วิธีแตกต่างกันหรือไม่ และตรวจสอบว่าขนาดตัวอย่างที่แตกต่างกันมีผลต่อประสิทธิภาพของการเปรียบเทียบพื้นที่ใต้โค้ง ROC ทั้งสองวิธีหรือไม่

1.2 วัตถุประสงค์การวิจัย

(1) เพื่อหาวิธีเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่

(2) เพื่อเปรียบเทียบวิธี Delong test กับวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ให้มีการแจกแจงแบบปกติแล้วนำไปทดสอบด้วย Z-test ว่าทั้ง 2 วิธีมีประสิทธิภาพแตกต่างกันอย่างไร

(3) เพื่อศึกษาว่าขนาดตัวอย่างมีผลต่อประสิทธิภาพของวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ให้มีการแจกแจงแบบปกติแล้วนำไปทดสอบด้วย Z-test หรือไม่

1.3 สมมติฐานการวิจัย

การทดสอบความแตกต่างของพื้นที่ใต้โค้ง ROC ด้วยวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ให้มีการแจกแจงแบบปกติจะให้ประสิทธิภาพการทดสอบดีกว่าวิธี Delong test

1.4 ขอบเขตงานวิจัย

(1) ทำการจำลองข้อมูลที่นำมาสร้างตัวแบบโลจิสติกเพื่อทำนายการผิดนัดชำระหนี้ ซึ่งข้อมูลประกอบไปด้วยตัวแปรต้นคือ อายุ เพศ เงินฝากธนาคาร ประเภทที่อยู่อาศัย และระยะเวลาในการทำงาน ตัวแปรตามคือ การผิดนัดชำระหนี้ ซึ่งทำการจำลองจากข้อมูล German credit ใน package caret ในโปรแกรม R

(2) กำหนดขนาดตัวอย่างในการศึกษาครั้งนี้คือ 300, 500 และ 1,000

(3) ทำการสร้างแบบจำลองเพื่อเปรียบเทียบความแตกต่างของพื้นที่ใต้โค้ง ROC

(4) ทำการแปลงข้อมูลความแตกต่างของพื้นที่ใต้โค้ง ROC ให้มีการแจกแจงแบบปกติ ด้วยวิธี Yeo-Johnson transformation แล้วนำไปทดสอบความแตกต่างของพื้นที่ใต้โค้ง ROC โดยวิธี Z-test

1.5 ประโยชน์ที่คาดว่าจะได้รับ

ผลการศึกษาจะแสดงให้เห็นถึงประสิทธิภาพของวิธีเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูป โดยใช้ข้อมูลชุดเดียวกัน และประสิทธิภาพของวิธีเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปที่ขนาดตัวอย่างแตกต่างกัน

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การจำลองข้อมูล Bootstrap

การจำลองข้อมูลเป็นวิธีการที่สร้างข้อมูลขึ้นมาเพื่อประมาณการแจกแจงของประชากรจากข้อมูลตัวอย่างที่มีอยู่ ข้อมูลที่สร้างขึ้นมาใหม่จะมีลักษณะใกล้เคียงกับการแจกแจงของประชากรจริง วิธีการจำลองข้อมูลมีอยู่ด้วยกันหลายวิธี โดยงานวิจัยนี้จะใช้วิธีการจำลองข้อมูลด้วยวิธี Bootstrap (Efron, 1979) เนื่องจากเป็นวิธีการจำลองข้อมูลที่ได้รับความนิยมและมีขั้นตอนการจำลองที่ไม่ซับซ้อน (Davison & Hinkley, 1997) การจำลองข้อมูลด้วยวิธี Bootstrap แบ่งได้เป็น 2 ประเภทคือ

(1) Parametric bootstrap เป็นการจำลองข้อมูลโดยที่รู้การแจกแจงแต่ไม่รู้ค่าพารามิเตอร์ $F_Y(y;.)$ โดยวิธี parametric bootstrap จะทำการประมาณค่าพารามิเตอร์ ($\hat{\theta}$) จากข้อมูลตัวอย่าง โดยวิธีที่ได้รับความนิยมในการประมาณค่าพารามิเตอร์คือ Maximum likelihood estimate เมื่อได้ค่าพารามิเตอร์จะทำการจำลองข้อมูลจากการแจกแจงที่รู้ด้วยค่าพารามิเตอร์ที่ประมาณได้ $F_Y(y;\hat{\theta})$ เช่น ข้อมูล (X) มีการแจกแจงแบบปกติ $f \sim N(\mu,\sigma)$ แต่ไม่รู้ค่าพารามิเตอร์ (μ,σ) ทำการประมาณค่าพารามิเตอร์ด้วยวิธี Maximum likelihood estimate จะได้ค่าประมาณพารามิเตอร์ $(\bar{x},\hat{\sigma})$ และสุ่มตัวอย่าง (X^*) จากการแจกแจงแบบปกติ $f \sim N(\bar{x},\hat{\sigma})$

(2) Nonparametric bootstrap เป็นการจำลองข้อมูลโดยที่ไม่รู้การแจกแจงของข้อมูลโดยที่ข้อมูลตัวอย่าง $X_1, X_2, ..., X_n$ แต่ละตัวเป็นอิสระกันและมีการแจกแจงเหมือนกัน (independent and identical distribution) วิธีการจำลอง nonparametric bootstrap จะทำการสุ่มข้อมูลจากกลุ่มตัวอย่างออกมาทีละตัวด้วยวิธีการสุ่มแบบคืนกลับด้วยความน่าจะเป็น $\frac{1}{n}$ หรือที่เรียกว่า resampling โดยข้อมูลสุ่ม $X_1^*, X_2^*, ..., X_n^*$ แต่ละตัวเป็นอิสระกันและมีการแจกแจงเหมือนกัน เช่น มีข้อมูล $X_1, X_2, ..., X_7$ ทำการสุ่มข้อมูลโดยสุ่มแบบคืนกลับด้วยความน่าจะเป็น $\frac{1}{7}$ ได้ข้อมูลจากการสุ่มคือ $X_5^*, X_1^*, X_4^*, X_1^*, X_7^*, X_6^*, X_2^*$ เป็นต้น

ในงานวิจัยนี้จะใช้การสุ่มแบบ Nonparametric bootstrap ในการจำลองข้อมูลเพื่อนำไปสร้างตัวแบบคะแนนเครดิต เนื่องจากข้อมูลที่นำมาสร้างตัวแบบเป็นข้อมูลที่ไม่รู้การแจกแจงของข้อมูล

2.2 คะแนนเครดิต (Credit Scoring)

คะแนนเครดิตเป็นเครื่องมือที่ช่วยประเมินโอกาสและความสามารถในการชำระหนี้ของผู้กู้ยืม ซึ่งมีขั้นตอนต่างๆดังต่อไปนี้

(1) วิเคราะห์และกำหนดตัวแปรที่น่าจะมีผลต่อการประเมินคะแนนเครดิต

(2) กำหนดตัวแบบเพื่อใช้ในการคิดคะแนนเครดิต โดยตัวแบบที่นิยมใช้ คือ ตัวแบบโลจิสติก ซึ่งเป็นตัวแบบที่ใช้ทำนายการผิดนัดชำระหนี้ผ่านปัจจัยต่างๆ เช่น อายุ เงินเดือน เป็นต้น

(3) การตรวจสอบความถูกต้องของตัวแบบซึ่งสามารถตรวจสอบได้จากพื้นที่ใต้เส้นโค้ง ROC

(4) เลือกตัวแบบที่ดีที่สุดในการนำมาพยากรณ์การผิดนัดชำระหนี้

ในปัจจุบันสถาบันการเงินของไทยได้มีการใช้คะแนนเครดิตกันมากขึ้น ไม่ว่าจะเป็น การให้บัตรเครดิต หรือสินเชื่อด้านต่างๆ นอกจากนั้นในต่างประเทศยังมีการกำหนดดอกเบี้ยเงินกู้จากคะแนนเครดิตกล่าวคือ ถ้าผู้ที่มีคะแนนเครดิตสูงจะมีโอกาสในการชำระหนี้ตรงตามเวลามากกว่าผู้ที่มีคะแนนเครดิตต่ำ ดังนั้นดอกเบี้ยเงินกู้ของผู้ที่มีคะแนนเครดิตสูงก็จะน้อยกว่าผู้ที่มีคะแนนเครดิตต่ำ โดยตัวแบบที่นิยมใช้สำหรับการคิดคะแนนเครดิตคือ ตัวแบบการถดถอยโลจิสติก (logistic regression model)

ตัวแปรต้นที่ใช้ในการพยากรณ์โอกาสในการชำระหนี้ตรงตามเวลานั้นมีได้หลายตัวแปร งานวิจัยนี้จะศึกษาการผิดนัดชำระหนี้ โดยมีตัวแปรต้น 5 ตัวแปร ได้แก่ 1) อายุ โดยช่วงวัยกลางคนมีอัตราการผิดนัดชำระหนี้ที่สูงกว่าช่วงวัยรุ่นและวัยสูงอายุ (Debbaut, Ghent, & Kudlyak, 2014) 2) เพศ โดย 60.3% ของผู้หญิงมีแนวโน้มที่จะเสียค่าธรรมเนียมจากการชำระหนี้ช้าเมื่อเปรียบเทียบกับ 57.4% ของผู้ชาย (Tamara, 2016) 3) เงินฝากธนาคาร พบว่า ผู้ที่มีเงินฝากสูงจะมีโอกาสในการชำระหนี้เต็มมากกว่าผู้ที่มีเงินฝากต่ำ (White, 2007) นอกจากนั้นยังมีตัวแปร 4) ประเภทที่อยู่อาศัยเป็นตัวช่วยบ่งบอกถึงทรัพย์สินที่ถือครองซึ่งคาดว่าจะมีผลกับการคิดคะแนนเครดิต (Lopes, 2008) และ 5) ระยะเวลาในการทำงาน เป็นตัวช่วยบ่งบอกถึงการไม่มีงานทำ การมีงานทำซึ่งส่งผลต่อคะแนนเครดิต (Latoya, 2018)

2.3 ตัวแบบการถดถอยโลจิสติก (Logistic regression model)

ตัวแบบการถดถอยโลจิสติก (logistic regression model) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรต้นกับตัวแปรตาม โดยที่ตัวแปรตามเป็นข้อมูลเชิงคุณภาพ ส่วนตัวแปรต้นสามารถเป็นได้ทั้งข้อมูลเชิงคุณภาพและข้อมูลเชิงปริมาณ อย่างที่กล่าวไว้ข้างต้นตัวแบบการถดถอยโลจิสติก (logistic regression model) ได้ถูกนำมาใช้อย่างแพร่หลายในการพยากรณ์โอกาสในการชำระหนี้ โดยตัวแบบที่ใช้ส่วนมากคือ ตัวแบบการถดถอยโลจิสติกทวิ

2.3.1 ตัวแบบการถดถอยโลจิสติกทวิ (Binary logistic regression model)

ตัวแบบการถดถอยโลจิสติกทวิเป็นการวิเคราะห์ความถดถอยที่ตัวแปรตามเป็นข้อมูลเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า คือ 0 และ 1 ส่วนตัวแปรต้นสามารถเป็นได้ทั้งข้อมูลเชิงคุณภาพและข้อมูลเชิงปริมาณ โดยการศึกษาครั้งนี้กำหนดให้ ตัวแปรตาม (y) มีค่า 2 ค่า คือ 0 แทน กลุ่มเหตุการณ์ที่ไม่ได้สนใจนั่นคือกลุ่มคนที่มีการชำระหนี้ตรงตามเวลาและ 1 แทนกลุ่มเหตุการณ์ที่สนใจนั่นคือ กลุ่มคนที่มีการผิดนัดชำระหนี้ในตัวแบบคะแนนเครดิต ซึ่งมีความสัมพันธ์กับตัวแปรต้น (x)

การเขียนตัวแบบโลจิสติกจะอยู่ในรูปของ log ของ odds[1] เรียกว่า logit หรือ logistic response function โดยสมการโลจิสติกจะอยู่ในรูป log(odds) ซึ่งเรียกว่า logit มีรูปแบบดังนี้

$$\operatorname{logit}(P) = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k$$

โดยที่

$\beta_0, \beta_1, \ldots, \beta_k$	แทนเป็นค่าพารามิเตอร์ของตัวแบบ
$P = \dfrac{e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k}}$	แทนความน่าจะเป็นของการเกิดเหตุการณ์ (y=1)
$X_1, X_2, \ldots, X_k$	แทนตัวแปรต้นตัวที่ 1,2,...,k

[1] $odds = \dfrac{\text{ความน่าจะเป็นของการเกิดเหตุการณ์ (y=1)}}{\text{ความน่าจะเป็นของการไม่เกิดเหตุการณ์ (y=0)}}$

2.3.2 การทดสอบความเหมาะสม (Goodness of fit test) ของตัวแบบ

เพื่อทดสอบว่าตัวแบบที่ได้มามีความเหมาะสมหรือไม่ ซึ่งสามารถตรวจสอบได้โดยวิธีดังต่อไปนี้

2.3.2.1 Hosmer-Lemeshow เป็นการทดสอบความเหมาะสมของแบบจำลอง โดยใช้การทดสอบไคสแควร์ ซึ่งมีสมติฐานดังนี้

$$H_0 : \operatorname{logit}(P) = \beta_0 + \beta_1 X_1 + ... + \beta_k X_k$$
$$H_a : not\ H_0$$

สถิติทดสอบ

$$HL = \sum_{i=1}^{g} \frac{\left(\sum_{j=1}^{n_j} Y_{ij} - \sum_{j=1}^{n_j} \hat{\pi}_{ij} \right)^2}{\sum_{j=1}^{n_j} \hat{\pi}_{ij} (1 - \sum_{j=1}^{n_j} \hat{\pi}_{ij} / n_j)} \sim \chi^2_{g-2}$$

โดยที่ $j = 1, 2, ..., n_j \ ; i = 1, 2, ..., g$

เมื่อ $\hat{\pi}_{ij}$ ความน่าจะเป็นที่คาดหวังตัวที่ j ในกลุ่มที่ i

Y_{ij} ความน่าจะเป็นที่ได้จากการสังเกตตัวที่ j ในกลุ่มที่ i

n_j จำนวนค่าที่สังเกตได้ในแต่ละกลุ่ม

g จำนวนกลุ่ม

2.3.2.2 การใช้ R^2 (Proportion of Correct Prediction) พิจารณาจากค่า R^2 ของ Cox & Snell และ Nagelkerke หรือ Pesudo R^2 เป็นค่าที่บอกสัดส่วนหรือเปอร์เซ็นต์ที่สามารถอธิบายความแปรผันใน Logistic Regression model

2.3.2.3 Walds Statistic เป็นการทดสอบค่านัยสำคัญของ β_i ในตัวแบบว่ามีค่าเท่ากับ 0 หรือไม่ โดยมีสมมติฐานในการทดสอบดังนี้

$$H_0 : \beta_i = 0 \quad ; i = 1, 2, ..., k$$
$$H_a : \beta_i \neq 0$$

สถิติทดสอบ

$$Z = \frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)} \sim N(0,1)$$

เมื่อค่า p-value < 0.05 จะปฏิเสธสมมติฐานว่างนั่นคือ ตัวแบบเต็มรูปมีความเหมาะสมมากกว่าตัวแบบลดรูป ซึ่งสอดคล้องกันกับการเปรียบเทียบพื้นที่ใต้โค้ง ROC นั่นคือตัวแบบเต็มรูปจะมีพื้นที่ใต้โค้ง ROC มากกว่าตัวแบบลดรูป

2.3.2.4 Likelihood ratio test (LRT) เป็นวิธีที่ใช้ในการเปรียบเทียบตัวแบบระหว่างตัวแบบเต็มรูป (full model) กับตัวแบบลดรูป (reduced model) โดยมีสมมติฐานในการทดสอบดังนี้

H_0: reduced model is true

H_a: full model is true

สถิติทดสอบ

LRT = -2log (Likelihood for reduced model) - (-2log (Likelihood for full model))

โดย LRT ประมาณด้วยการแจกแจงแบบไคสแควร์ ที่มีองศาอิสระเท่ากับผลต่างของจำนวนพารามิเตอร์ในตัวแบบเต็มรูปกับตัวแบบลดรูป เมื่อค่า p-value < 0.05 จะปฏิเสธสมมติฐานว่างนั่นคือ ตัวแบบเต็มรูปมีความเหมาะสมมากกว่าตัวแบบลดรูป ซึ่งสอดคล้องกันกับการเปรียบเทียบพื้นที่ใต้โค้ง ROC นั่นคือตัวแบบเต็มรูปจะมีพื้นที่ใต้โค้ง ROC มากกว่าตัวแบบลดรูป

ดังนั้น LRT จึงมีความเหมาะสมที่จะใช้ในการเปรียบเทียบผลจากวิธี Delong test กับวิธีการแปลงข้อมูลของผลต่างพื้นที่ใต้โค้ง ROC แล้วนำไปทดสอบด้วย Z-test ว่าผลที่ได้จากทั้ง 2 วิธีสอดคล้องกับผลจากวิธี LRT หรือไม่

2.4 Receiver Operating Characteristic curve (ROC Curve)

กราฟ ROC เป็นกราฟที่แสดงให้เห็นถึงประสิทธิภาพของ binary classification ซึ่งตัวแปรตาม (y) เป็นตัวแปรเชิงคุณภาพแบ่งออกเป็น 2 กรณีคือ y = 1 เมื่อเกิดเหตุการณ์ที่สนใจหรือผลการทดสอบเป็นบวก และ y = 0 เมื่อเกิดเหตุการณ์ที่ไม่ได้สนใจหรือผลการทดสอบเป็นลบ

จุดตัด (Cut-off point) หมายถึง จุดที่ใช้จำแนกเหตุการณ์ออกเป็นเหตุการณ์ที่สนใจกับเหตุการณ์ที่ไม่ได้สนใจพบว่า สามารถแบ่งกรณีการเปรียบเทียบระหว่างค่าพยากรณ์และค่าสังเกตออกเป็น 4 กรณี ดังนี้

		ค่าสังเกต	
		Positive	Negative
ค่าพยากรณ์	Positive	True positive (TP)	False positive (FP)
	Negative	False negative (FN)	True negative (TN)

True positive (TP) คือ จำนวนเหตุการณ์ที่สนใจมีผลเป็นบวกและมีผลจากการพยากรณ์เป็นบวก

False positive (FP) คือ จำนวนเหตุการณ์ที่สนใจมีผลเป็นลบแต่มีผลจากการพยากรณ์เป็นบวก

False negative (FN) คือ จำนวนเหตุการณ์ที่สนใจมีผลเป็นบวกและมีผลจากการพยากรณ์เป็นลบ

True negative (TN) คือ จำนวนเหตุการณ์ที่สนใจมีผลเป็นลบและมีผลจากการพยากรณ์เป็นลบ

ตัวสถิติที่ใช้วัดความถูกต้องของการพยากรณ์ซึ่งผลที่เกิดจากการพยากรณ์มีเพียง 2 ค่า คือ

Sensitivity or True positive rate (TPR) คือ อัตราส่วนของจำนวนค่าพยากรณ์ที่มีผลเป็นบวกที่ทำนายถูกต่อจำนวนเหตุการณ์ที่สนใจที่มีผลเป็นบวก จะได้

$$sensitivity = \frac{TP}{TP + FN}$$

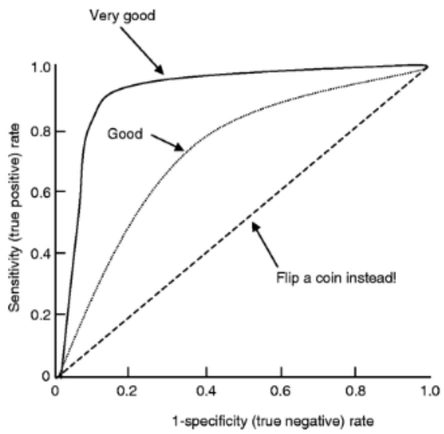
Specificity or True negative rate (TNR) คือ อัตราส่วนของจำนวนค่าพยากรณ์ที่มีผลเป็นลบที่ทำนายถูกต่อจำนวนเหตุการณ์ที่สนใจที่มีผลเป็นลบ จะได้

$$specificity = \frac{TN}{FP + TN}$$

Empirical ROC Curve หรือ ROC curve เป็นกราฟที่พล็อตระหว่างค่า Sensitivity และ 1-specificity โดย แกน y แทน Sensitivity และ แกน x แทน 1-specificity ที่จุดตัดต่างๆเพื่อแบ่งผลลัพธ์ของการพยากรณ์ออกเป็น 2 กลุ่ม คือกลุ่มที่เกิดเหตุการณ์ และกลุ่มที่ไม่เกิดเหตุการณ์

พื้นที่ใต้โค้ง ROC เป็นดัชนีที่ใช้ในการบ่งชี้ความถูกต้องหรือความน่าเชื่อถือของตัวแบบ ตัวแบบใดที่มีพื้นที่ใต้โค้ง ROC มากที่สุดจะถือว่าเป็นตัวแบบดีที่สุด โดยการหาพื้นที่ AUC สามารถทำได้

โดยการอินทิเกรตพื้นที่ใต้กราฟด้วยวิธี Trapezoidal rule หรือวิธี Nonparametric คือ Mann-Whitney U-statistic (Bamber, 1975) และถ้า $A\hat{UC} < 0.5$ จะถือว่าตัวแบบไม่มีความน่าเชื่อถือ กล่าวคือไม่สามารถแยกจำแนกเหตุการณ์ที่สนใจออกจากกลุ่มเหตุการณ์ที่ไม่สนใจได้ โดยกราฟ ROC จะอยู่ใต้เส้นประใน รูปที่ 2.1 และจากรูปที่ 2.1 พบว่าเส้นโค้ง Good จะมีพื้นที่ใต้กราฟที่น้อยกว่าเส้นโค้ง Very good ดังนั้นตัวแบบของเส้นโค้ง Very good จะมีความน่าเชื่อถือมากกว่า (Tape, n.d.) ได้แบ่งการวัดประสิทธิภาพของตัวแบบจากพื้นที่ใต้กราฟดังตารางที่ 2.1



รูปที่ 2.1 แสดงเส้นโค้ง ROC

(จาก:https://med.mahidol.ac.th/fammed/sites/default/files/public/pdf/EBM_Diagnostic_study.pdf)

ตารางที่ 2.1 แสดงประสิทธิภาพของตัวแบบจากพื้นที่ใต้โค้ง ROC (Tape, n.d.)

AUC	ประสิทธิภาพ
0.9-1.0	ยอดเยี่ยม
0.8-0.9	ดี
0.7-0.8	ปานกลาง
0.6-0.7	ไม่ดี
0.5-0.6	แย่

2.5 การเปรียบเทียบพื้นที่ใต้กราฟ ROC

การเปรียบเทียบพื้นที่ใต้โค้ง ROC เป็นการทดสอบพื้นที่ใต้กราฟ ROC ของ 2 ตัวแบบมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ Bamber (1975) ได้ทำการเปรียบเทียบพื้นที่ใต้โค้ง ROC โดยการคำนวณด้วยวิธี Trapezoidal กับวิธี Mann-Whitney U-statistic พบว่าทั้งสองวิธีได้พื้นที่ใต้โค้ง ROC ที่เท่ากัน McClish (1989) ได้ทำการศึกษาเกี่ยวกับการหาพื้นที่ใต้โค้ง ROC และการเปรียบเทียบพื้นที่ใต้โค้ง ROC ของข้อมูลชุดเดียวกันที่มีการแจกแจงแบบปกติ (DeLong, 1988) ได้ทำการศึกษาเกี่ยวกับการหาพื้นที่ใต้โค้ง ROC และการเปรียบเทียบพื้นที่ใต้โค้ง ROC ของข้อมูลชุดเดียวกันที่ไม่มีการแจกแจงแบบปกติด้วยสถิติทดสอบ Z (Z-test) ซึ่งเป็นวิธีการที่ได้รับความนิยมมาก

2.5.1 Delong method

การทดสอบความแตกต่างของพื้นที่ใต้โค้ง ROC ด้วยวิธี Delong test เป็นการทดสอบโดยใช้ตัวสถิติ Mann-Whitney ในการคำนวณหาพื้นที่ใต้โค้ง ROC ซึ่งกำหนดตัวแปรดังต่อไปนี้

ให้ m แทนจำนวนของเหตุการณ์ที่สนใจ

n แทนจำนวนของเหตุการณ์ที่ไม่ได้สนใจ

$X_i \,;\, i = 1, 2, ..., m$ แทนค่าของเหตุการณ์ในกลุ่มที่สนใจ

$Y_j \,;\, j = 1, 2, ..., n$ แทนค่าของเหตุการณ์ในกลุ่มที่ไม่ได้สนใจ

$\hat{\theta}$ แทนความน่าจะเป็นของจำนวนค่าของเหตุการณ์ที่สนใจที่มากกว่าหรือเท่ากับค่าของเหตุการณ์ที่ไม่ได้สนใจต่อเหตุการณ์ทั้งหมด จะได้

$$\hat{\theta} = \frac{1}{mn} \sum_{j=1}^{n} \sum_{i=1}^{m} \psi(X_i, Y_j)$$

เมื่อ

$$\psi(X, Y) = \begin{cases} 1 & ; Y < X \\ 0.5 & ; Y = X \\ 0 & ; Y > X \end{cases}$$

ให้ S_{10} และ S_{01} เป็นเมทริกซ์ขนาด $k \times k$ เมื่อ k แทนจำนวนตัวแบบที่จะทำการทดสอบ

สมาชิกในแถวที่ r หลักที่ s ของเมทริกซ์ S_{10} คือ $s_{10}^{r,s}$

$$s_{10}^{r,s} = \frac{1}{m-1}\sum_{i=1}^{m}[V_{10}^{r}(X_i)-\hat{\theta}^{r}][V_{10}^{s}(X_i)-\hat{\theta}^{s}]$$

สมาชิกในแถวที่ r หลักที่ s ของเมทริกซ์ S_{01} คือ $s_{01}^{r,s}$

$$s_{01}^{r,s} = \frac{1}{n-1}\sum_{j=1}^{m}[V_{01}^{r}(Y_j)-\hat{\theta}^{r}][V_{01}^{s}(Y_j)-\hat{\theta}^{s}]$$

เมื่อ $V_{10}^{r}(X_i) = \frac{1}{n}\sum_{i=1}^{n}\psi(X_i^r, Y_j^r) \quad ; \ i = 1,2,\ldots,m$

$V_{01}^{r}(Y_j) = \frac{1}{m}\sum_{i=1}^{n}\psi(X_i^r, Y_j^r) \quad ; \ j = 1,2,\ldots,n$

$\hat{\theta}^{r}, \hat{\theta}^{s}$ แทนความน่าจะเป็นของจำนวนค่าของเหตุการณ์ที่สนใจที่มากกว่าหรือเท่ากับค่าของเหตุการณ์ที่ไม่ได้สนใจต่อเหตุการณ์ทั้งหมดของตัวแบบที่ r และตัวแบบที่ s ตามลำดับ

จะได้ ค่าประมาณของ covariance matrix ของเวกเตอร์พารามิเตอร์ $\hat{\theta} = (\hat{\theta}^1, \hat{\theta}^2, \ldots, \hat{\theta}^k)$ คือ

$$S = \frac{1}{m}S_{10} + \frac{1}{n}S_{01}$$

โดยมีสมมติฐานการทดสอบดังนี้

$H_0 : AUC_p = AUC_{p-k}$

$H_a : AUC_p \neq AUC_{p-k}$

เมื่อ AUC_p แทน ตัวแบบเต็มรูป

AUC_{p-k} แทน ตัวแบบลดรูป

ตัวสถิติที่ใช้ทดสอบ

$$z = \frac{(eAUC_p - eAUC_{p-k}) - (AUC_p - AUC_{p-k})}{(LSL^T)^{1/2}}$$

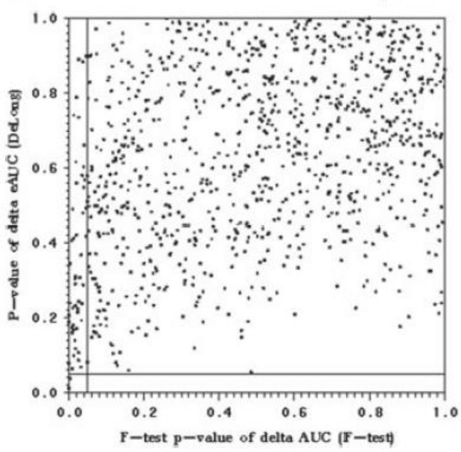
เมื่อ L แทนเวกเตอร์แถวของสัมประสิทธิ์

$eAUC_p$ แทนพื้นที่ใต้กราฟ ROC ของตัวแบบเต็มรูป

$eAUC_{p-k}$ แทนพื้นที่ใต้กราฟ ROC ของตัวแบบลดรูป

2.5.2 การนำวิธี Delong test ไปใช้ในทางที่ผิด

Demler et al. (2012) ได้ทำการศึกษาการทดสอบของ Delong พบว่าเมื่อลองทำการสุ่มพื้นที่ใต้กราฟมาแล้วพบว่าผลต่างของพื้นที่ใต้กราฟนั้นไม่เป็นการแจกแจงแบบปกติ การใช้ Delong test สำหรับข้อมูลของผลต่างของพื้นที่ใต้โค้ง ROC ที่ไม่มีการแจกแจงแบบปกตินั้นให้ผลการวิเคราะห์ที่เป็นไปในทางตรงกันข้ามกับ (F-test) นั่นคือ F-test พบความแตกต่างของตัวแบบเต็มรูปและตัวแบบลดรูปอย่างมีนัยสำคัญทางสถิติแต่ Delong test ไม่พบความแตกต่างของ 2 ตัวแบบในทางกลับกันเมื่อ Delong test พบความแตกต่างของตัวแบบเต็มรูปและตัวแบบลดรูปอย่างมีนัยสำคัญทางสถิติแต่ F-test ไม่พบความแตกต่างของ 2 ตัวแบบอย่างมีนัยสำคัญทางสถิติ แสดงดังรูปที่ 2.2 Demler et al. จึงสรุปว่า Delong test เป็นวิธีที่ไม่เหมาะสมในการใช้เปรียบเทียบผลต่างของพื้นที่ใต้โค้ง ROC ที่ไม่เป็นการแจกแจงแบบปกติ



รูปที่ 2.2 กราฟแสดงค่า p-value ระหว่าง F-test กับ Delong test
(ที่มา: Demler et al., 2012)

โดยการศึกษานี้จะทำการตรวจสอบว่าข้อมูลผลต่างพื้นที่ใต้โค้ง ROC มีการแจกแจงแบบปกติหรือไม่ หากไม่มีการแจกแจงแบบปกติจะต้องทำการแปลงข้อมูลก่อน

2.6 การแปลงข้อมูล (Transformation)

การแปลงข้อมูลคือการเปลี่ยนข้อมูลให้มีลักษณะตรงตามที่เราต้องการศึกษาโดยวิธีการแปลงข้อมูลนั้นมีได้หลายวิธี เช่น การแปลงข้อมูลของ Tukey (1957) การแปลงข้อมูลของ Box and Cox

(1964) เป็นต้นซึ่งวิธีที่ได้รับความนิยมสำหรับการแก้ปัญหาความแปรปรวนไม่เท่ากันและการแก้ปัญหาการแจกแจงปกติ คือ วิธีการแปลงข้อมูลแบบ Box-Cox

2.6.1 Box-Cox Transformation

เป็นวิธีการแปลงข้อมูลที่ได้รับความนิยมมาก เนื่องจากเป็นวิธีการแปลงข้อมูลที่ให้ผลในการแก้ไขปัญหาได้ดี และยังมีรูปแบบการแปลงที่ไม่ยุ่งยาก ซึ่งรูปแบบการแปลง คือ

$$y_i^{\lambda} = \begin{cases} \dfrac{y_i^{\lambda} - 1}{\lambda} & ; \lambda \neq 0 \\ \ln(y_i) & ; \lambda = 0 \end{cases}$$

เมื่อ λ คือ ค่าพารามิเตอร์ที่ได้มาจากข้อมูล

แต่การแปลงข้อมูลแบบ Box-Cox ข้อมูลที่นำมาแปลงข้อมูลจะต้องเป็นข้อมูลที่มีค่าเป็นบวก ซึ่ง (Yeo & Johnson, 2000) ได้ทำการขยายการเปลี่ยนแปลงข้อมูลสำหรับข้อมูลที่มีบางค่าติดลบ

2.6.2 Yeo-Johnson Transformation

เป็นวิธีการแปลงข้อมูลสำหรับข้อมูลที่มีบางค่าติดลบหรือมีบางค่าเป็นศูนย์ ซึ่งรูปแบบการแปลงคือ

$$y_i^{(\lambda)} = \begin{cases} \dfrac{(y_i + 1)^{\lambda} - 1}{\lambda} & ; \lambda \neq 0, y \geq 0 \\ \log(y_i + 1) & ; \lambda = 0, y \geq 0 \\ -\dfrac{(y_i + 1)^{(2-\lambda)} - 1}{2 - \lambda} & ; \lambda \neq 2, y < 0 \\ -\log(-y_i + 1) & ; \lambda = 2, y < 0 \end{cases}$$

โดยผู้วิจัยจะใช้การแปลงข้อมูลด้วยวิธี Yeo-Johnson Transformation เนื่องจากผลต่างของพื้นที่ใต้โค้ง ROC มีบางค่าที่ติดลบ

บทที่ 3
วิธีการศึกษา

ในงานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบพื้นที่ใต้เส้นโค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ สำหรับการจำลองข้อมูลและการวิเคราะห์ข้อมูลทั้งหมดจะทำงานด้วยโปรแกรม R

3.1 ลักษณะและที่มาของข้อมูล

ข้อมูลที่นำมาสร้างตัวแบบโลจิสติกเพื่อทำนายการผิดนัดชำระหนี้ ซึ่งข้อมูลประกอบไปด้วยตัวแปรต้นคือ อายุ เพศ เงินฝากธนาคาร ประเภทที่อยู่อาศัย และระยะเวลาในการทำงาน ตัวแปรตามคือ การผิดนัดชำระหนี้ ซึ่งทำการจำลองจากข้อมูล German credit ใน package caret ในโปรแกรม R ข้อมูลมีจำนวนทั้งหมด 1000 คน ซึ่งประกอบไปด้วย

(1) ข้อมูลการผิดนัดชำระหนี้ เป็นข้อมูลเชิงคุณภาพโดยแบ่งเป็นผู้ที่ไม่ผิดนัดชำระหนี้ (70%) และผู้ที่ผิดนัดชำระหนี้ (30%)

(2) อายุ เป็นข้อมูลเชิงปริมาณที่มีอายุเฉลี่ย 35.55 ปี และส่วนเบี่ยงเบนมาตรฐาน 11.38 ปี

(3) เพศ เป็นข้อมูลเชิงคุณภาพแบ่งเป็นเพศหญิง (30.9%) และเพศชาย (69.1%)

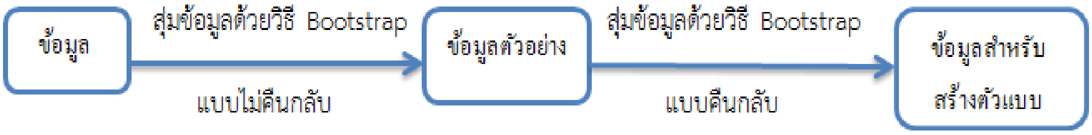
(4) เงินฝากธนาคาร เป็นข้อมูลเชิงคุณภาพแบ่งเป็น เงินฝากน้อยกว่า 100 Deutsche Mark (DM) (60.3%) เงินฝากระหว่าง 100DM ถึง 500 DM (10.3%) เงินฝากระหว่าง 500DM ถึง 1000 DM (6.3%) เงินฝากมากกว่า 1000DM (4.8%) และไม่มีเงินฝากธนาคาร (18.3%)

(5) ประเภทที่อยู่อาศัย เป็นข้อมูลเชิงคุณภาพแบ่งเป็น ที่พักอาศัยของตนเอง (71.3%) อาศัยผู้อื่นอยู่เช่น บ้านญาติ (17.9%) และเช่าอาศัย (10.8%)

(6) ระยะเวลาในการทำงาน เป็นข้อมูลเชิงคุณภาพแบ่งเป็น ว่างงาน (6.2%) ระยะเวลาทำงานน้อยกว่า 1 ปี (17.2%) ระยะเวลาทำงานระหว่าง 1 ปี ถึง 4 ปี (33.9%) ระยะเวลาทำงานระหว่าง 4 ปี ถึง 7 ปี (17.4%) และระยะเวลาทำงานมากกว่า 7 ปี (25.3%)

3.2 การจำลองข้อมูล

ทำการจำลองเพื่อนำข้อมูลมาสร้างตัวแบบโลจิสติกสำหรับใช้ในการทำนายการผิดนัดชำระหนี้ โดยใช้โปรแกรม R สำหรับการจำลองข้อมูลโดยใช้วิธี Bootstrap ซึ่งทำการสุ่มทั้งหมด 2 ครั้ง ดังรูปที่ 3.1 โดยวิธีการสุ่มดังนี้



รูปที่ 3.1 แสดงการสุ่มตัวอย่างข้อมูลสำหรับการสร้างตัวแบบ

(1) สุ่มข้อมูลจากข้อมูล German credit ด้วยวิธี Bootstrap แบบไม่คืนกลับจำนวน 300 500 และ 1,000 เพื่อนำมาเป็นข้อมูลตัวอย่างในการนำไปสร้างตัวแบบ

(2) สุ่มข้อมูลจากข้อมูลในข้อที่ (1) ด้วยวิธี Bootstrap แบบคืนกลับนำมาสร้างตัวแบบจำนวน 50 ตัวแบบที่ขนาดตัวอย่าง 300 500 และ 1,000 เพื่อนำไปสร้างตัวแบบและวัดประสิทธิภาพของตัวแบบ

ข้อมูลที่ได้จากการสุ่มในครั้งที่ 2 จะนำข้อมูลที่ทำการสุ่มได้แบ่งเป็น 2 ส่วนคือ ส่วนแรกสำหรับการสร้างตัวแบบ (70% ของข้อมูล) ส่วนที่สองสำหรับการตรวจสอบตัวแบบ (30% ของข้อมูล) โดยใช้วิธีการสุ่มแบบไม่คืนกลับ แล้วนำข้อมูลทั้ง 2 ส่วนไปสร้างตัวแบบและการตรวจสอบตัวแบบ

3.3 การสร้างตัวแบบ

ตัวแบบที่นำมาใช้ในการวิเคราะห์ข้อมูลคือตัวแบบโลจิสติก โดยให้ y=1 แทนกลุ่มคนที่ผิดนัดชำระหนี้ และ y=0 แทนกลุ่มคนที่ไม่ผิดนัดชำระหนี้ โดยนำข้อมูลที่ทำการสุ่มในส่วนแรกมาสร้างตัวแบบเต็มรูปและตัวแบบลดรูปที่ขนาดตัวอย่าง 300 500 และ 1,000 ซึ่งตัวแบบที่เกิดขึ้นทั้งหมดคือ

ตัวแบบเต็มรูป

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{saving} + \beta_5 X_{present}$$

ตัวแบบลดรูป

- 4 ตัวแปร

$$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving} + \beta_4 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{saving}$$

- 3 ตัวแปร

$$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving} + \beta_3 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving} + \beta_3 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving} + \beta_3 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing}$$

- 2 ตัวแปร

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving}$$

$$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing}$$

$$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving}$$

$$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving}$$

$$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{present}$$

$$logit(P) = \beta_0 + \beta_1 X_{saving} + \beta_2 X_{present}$$

- 1 ตัวแปร

$$logit(P) = \beta_0 + \beta_1 X_{gender}$$

$$logit(P) = \beta_0 + \beta_1 X_{age}$$

$$logit(P) = \beta_0 + \beta_1 X_{housing}$$

$$logit(P) = \beta_0 + \beta_1 X_{saving}$$

$$logit(P) = \beta_0 + \beta_1 X_{present}$$

เมื่อ $\beta_0, \beta_1, ..., \beta_5$ แทน ค่าพารามิเตอร์ในตัวแบบ

X_{gender} แทน ตัวแปรเพศ

X_{age} แทน ตัวแปรอายุ

$X_{housing}$ แทน ตัวแปรประเภทที่อยู่อาศัย

X_{saving} แทน ตัวแปรเงินฝากในธนาคาร

$X_{present}$ แทน ตัวแปรระยะเวลาในการทำงาน

เมื่อได้ตัวแบบมาแล้ว ขั้นตอนต่อไปคือการนำตัวแบบที่ได้มาตรวจสอบความถูกต้องของตัวแบบ

3.4 การตรวจสอบความถูกต้องของตัวแบบ

การตรวจสอบความถูกต้องของตัวแบบสามารถทำได้หลายวิธี ในงานวิจัยนี้ผู้วิจัยจะทำการตรวจสอบความถูกต้องของตัวแบบด้วยวิธีโค้ง ROC โดยนำข้อมูลส่วนที่สอง มาตรวจสอบความถูกต้องของตัวแบบ ทั้งตัวแบบเต็มรูปและตัวแบบลดรูป ซึ่งเส้นโค้ง ROC เป็นกราฟพล็อตระหว่างค่า Sensitivity และ 1-specificity โดย แกน y แทน Sensitivity และ แกน x แทน 1-specificity ที่จุดตัดต่างๆ แล้วนำเส้นโค้ง ROC ที่ได้จากตัวแบบเต็มรูปและตัวแบบลดรูปไปคำนวณหาค่าพื้นที่ใต้โค้งของแต่ละตัวแบบเพื่อนำไปเปรียบเทียบต่อไป

3.5 เปรียบเทียบพื้นที่ใต้โค้ง ROC

การเปรียบเทียบพื้นที่ใต้โค้ง ROC เพื่อทดสอบว่าพื้นที่ของตัวแบบ 2 ตัวแบบมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ โดยงานวิจัยนี้มีวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC 2 วิธี คือวิธี Delong testและวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง (Transform) แล้วนำไปใช้ Z-test

- วิธี Delong test มีขั้นตอนการเปรียบเทียบดังนี้
  (1) คำนวณพื้นที่ใต้โค้งของตัวแบบเต็มแบบและตัวแบบลดรูป แล้วนำข้อมูลไปหาผลต่างพื้นที่ใต้โค้ง ROC
  (2) คำนวณค่า covariance matrix ของเวกเตอร์พารามิเตอร์
  (3) ทำการเปรียบเทียบความแตกต่างของพื้นที่ใต้โค้ง ROC ด้วยสถิติทดสอบ

$$z = \frac{(eAUC_p - eAUC_{p-k}) - (AUC_p - AUC_{p-k})}{(LSL^T)^{1/2}}$$

ซึ่งวิธีการทดสอบความแตกต่างของพื้นที่ใต้โค้งด้วยวิธี Delong test ผู้วิจัยจะใช้ package pROC ในโปรแกรม R ในการคำนวณหาค่า p-value ของความแตกต่างพื้นที่ใต้กราฟระหว่างตัวแบบเต็มรูปและตัวแบบลดรูป

- วิธี Transform มีการเปรียบเทียบดังนี้
  (1) คำนวณพื้นที่ใต้โค้งของตัวแบบเต็มแบบและตัวแบบลดรูป แล้วนำข้อมูลไปหาผลต่างพื้นที่ใต้โค้ง ROC
  (2) นำข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ไปแปลงข้อมูลด้วยวิธี Yeo-Johnson Transformation
  (3) นำข้อมูลที่แปลงไปทดสอบโดยใช้ Z-test ในการเปรียบเทียบความแตกต่างของพื้นที่ใต้โค้ง ROC

3.6 การเปรียบเทียบวิธีการทดสอบ

การเปรียบเทียบวิธีการทดสอบทั้งสองวิธี ผู้วิจัยทำการเปรียบเทียบจากความสอดคล้องของทั้งสองวิธีกับวิธี Likelihood ratio test ว่าทั้งสองวิธีมีความสอดคล้องกับวิธี Likelihood ratio test มากน้อยเพียงใด โดยพิจารณากราฟ p-value จาก Likelihood ratio test เทียบกับ Delong test และ วิธีการแปลงข้อมูล (Transform) ผลต่างพื้นที่ใต้โค้ง ROC โดยให้แกน x แทน ค่า p-value ที่ได้จากวิธี Likelihood ratio test และแกน y แทนค่า p-value ที่ได้จากวิธี Delong test และวิธีการแปลงข้อมูล ที่ขนาดตัวอย่าง 300 500 และ 1,000

นอกจากนั้นผู้วิจัยยังพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ (correlation) ของค่า p-value ระหว่างวิธี Delong test กับ วิธี Likelihood ratio test และวิธี Transform กับ วิธี Likelihood ratio test ที่ขนาดตัวอย่าง 300 500 และ 1,000 เมื่อได้ค่าสัมประสิทธิ์สหสัมพันธ์จากวิธี Delong test และวิธี Transform ผู้วิจัยจะทำการตรวจสอบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของทั้งสองวิธีมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ ถ้ามีความแตกต่างกันแล้วค่าสัมประสิทธิ์สหสัมพันธ์จากวิธีใดจะมีค่ามากกว่าอย่างมีนัยสำคัญทางสถิติ ที่ขนาดตัวอย่าง 300 500 และ 1000 ตัวสถิติทดสอบ t-test และทดสอบว่าที่ขนาดตัวอย่างที่แตกต่างกันนั้นค่าสัมประสิทธิ์สหสัมพันธ์ของค่า p-values ระหว่างวิธี Transform กับ วิธี Likelihood ratio test มีค่าแตกต่างกันอย่างมีนัยสำคัญหรือไม่

3.7 การนำวิธี Transform มาใช้ในการเปรียบเทียบพื้นที่ใต้โค้ง ROC

การเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปและตัวแบบลดรูป เพื่อทดสอบว่าตัวแบบเต็มรูปและตัวแบบลดรูปว่าตัวแบบใดมีความเหมาะสมมากกว่ากัน โดยทดสอบจากการเปรียบเทียบพื้นที่ใต้โค้ง ROC ซึ่งมีขั้นตอนการทดสอบดังนี้

(1) ทำการสุ่มข้อมูลจากข้อมูลตัวอย่างด้วยวิธี Bootstrap แบบคืนกลับเพื่อนำข้อมูลมาสร้างตัวแบบเต็มรูปและตัวแบบลดรูปจำนวน 50 ตัวแบบ

(2) คำนวณพื้นที่ใต้โค้ง ROC ของตัวแบบเต็มแบบและตัวแบบลดรูป แล้วนำค่าพื้นที่ใต้โค้ง ROC ไปหาผลต่างพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปและตัวแบบลดรูปทั้ง 50 ตัวแบบ

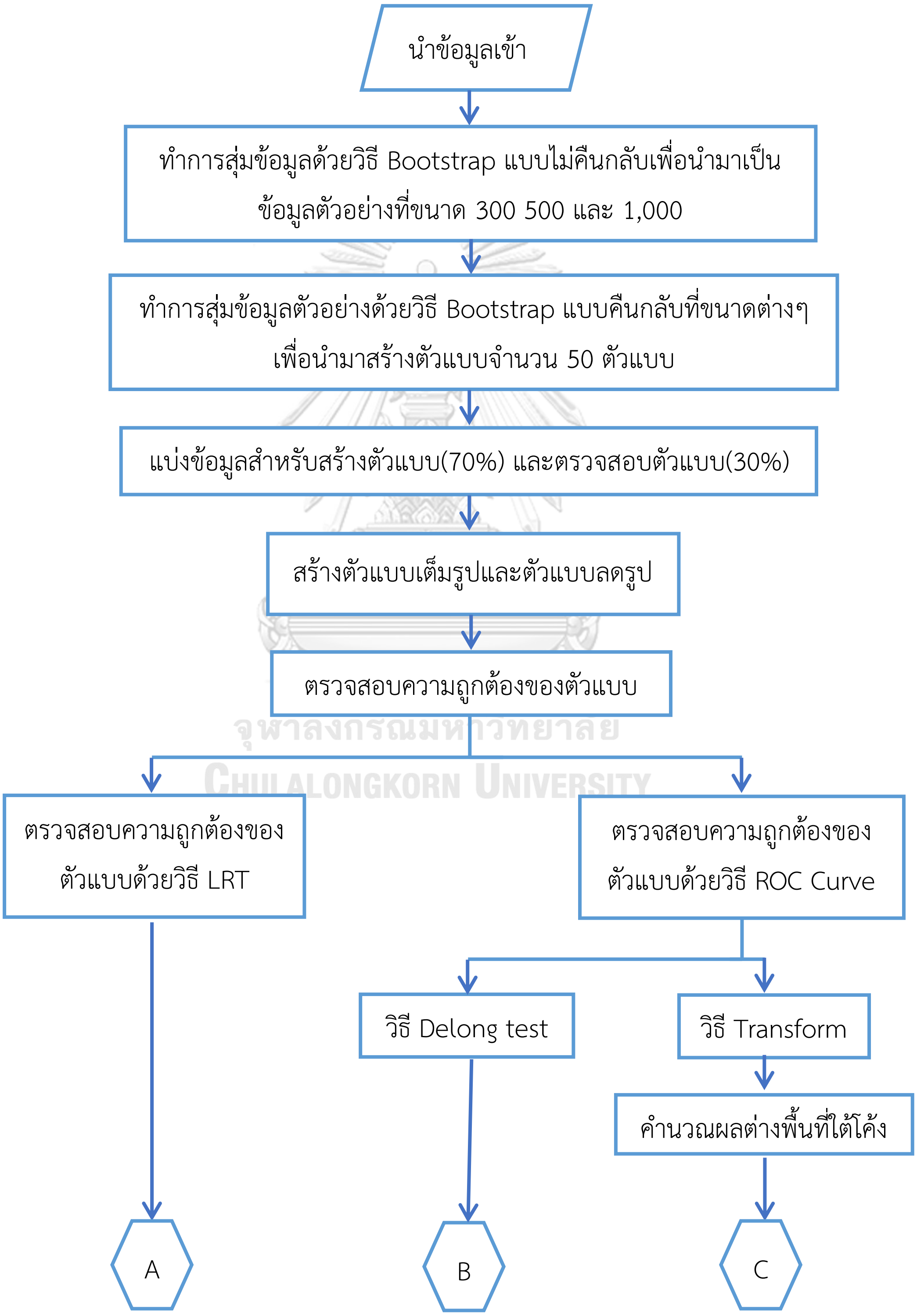
(3) นำข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ในขั้นที่ (2) ไปคำนวณหาค่าพารามิเตอร์สำหรับการแปลงข้อมูลด้วยวิธี Yeo-Johnson Transformation

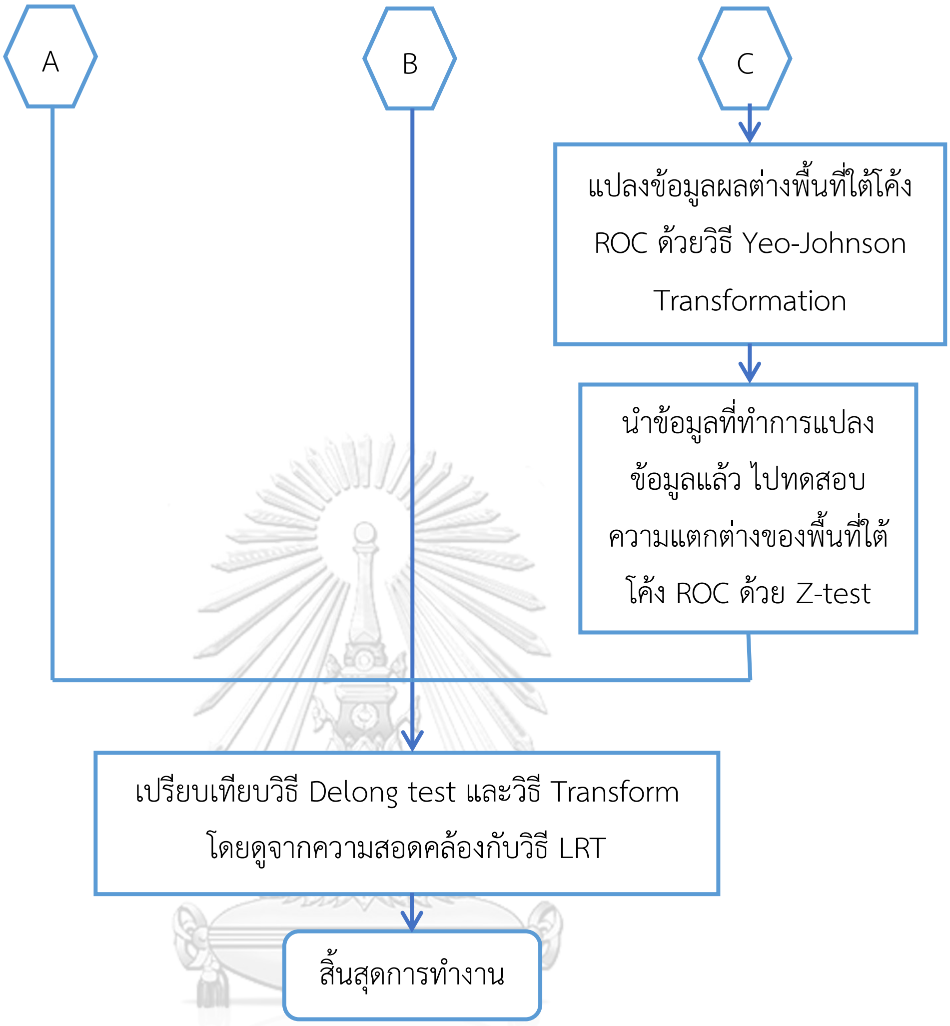
(4) ทำการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC จากข้อมูลตัวอย่างด้วยวิธี Yeo-Johnson Transformation โดยใช้ค่าพารามิเตอร์จากขั้นที่ (3) และนำข้อมูลผลต่างพื้นที่ใต้โค้ง ROC จากข้อมูลตัวอย่างที่ทำการแปลงแล้วไปใช้ Z-test ในการเปรียบเทียบความแตกต่างของพื้นที่ใต้โค้ง ROC

(5) ทำการสรุปผลความแตกต่างของพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปและตัวแบบลดรูป

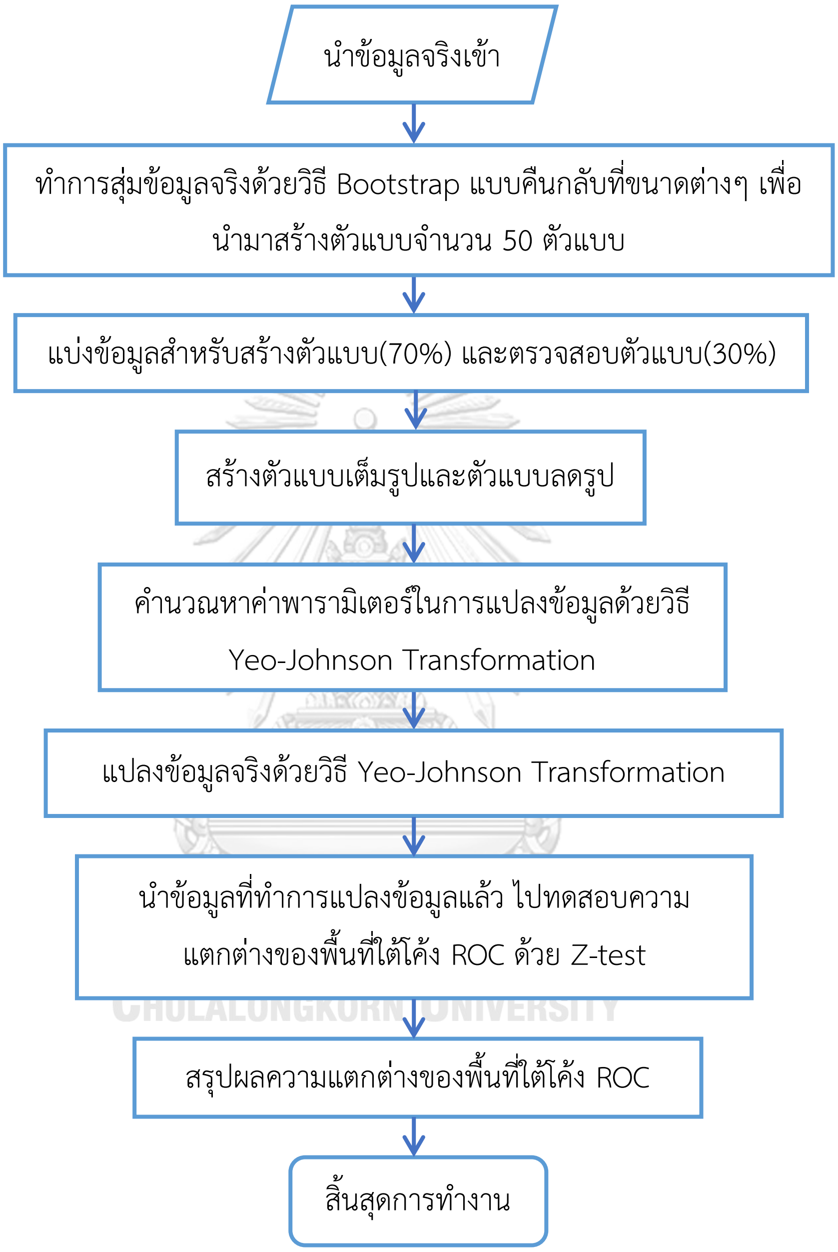
3.8 ขั้นตอนการทำงานของโปรแกรม R

จากที่กล่าวมาข้างต้น สามารถเขียนแผนภาพการทำงานของโปรแกรม R สำหรับเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างวิธี Transform และวิธี Delong test ได้ดังนี้




A
B
C
แปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ด้วยวิธี Yeo-Johnson Transformation
นำข้อมูลที่ทำการแปลงข้อมูลแล้ว ไปทดสอบความแตกต่างของพื้นที่ใต้โค้ง ROC ด้วย Z-test
เปรียบเทียบวิธี Delong test และวิธี Transform โดยดูจากความสอดคล้องกับวิธี LRT
สิ้นสุดการทำงาน

และสามารถเขียนแผนภาพการทำงานของโปรแกรม R สำหรับการนำวิธี Transform มาใช้ในการเปรียบเทียบพื้นที่ใต้โค้ง ROC ได้ดังนี้



บทที่ 4
ผลการศึกษา

งานวิจัยนี้มีวัตถุประสงค์เพื่อหาวิธีเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ และนำวิธีใหม่นี้มาเปรียบเทียบวิธี Delong test ว่าทั้ง 2 วิธีมีประสิทธิภาพแตกต่างกันอย่างไร โดยดูจากความสอดคล้องของค่า p-value ของวิธี Delong test และวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC เทียบกับวิธี Likelihood ratio test เนื่องจากวิธี Likelihood ratio test เป็นการทดสอบความเหมาะสมของตัวแบบระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูป จึงเป็นวิธีที่มีความเหมาะสมสำหรับดูความสอดคล้องจาก 2 วิธี นอกจากนี้งานวิจัยนี้ทำการตรวจสอบว่าขนาดตัวอย่างมีผลต่อประสิทธิภาพของทั้งสองวิธีหรือไม่ โดยข้อมูลที่ใช้ในการวิจัยนำมาจากข้อมูล German credit ใน package caret ในโปรแกรม R ด้วยวิธี Bootstrap ข้อมูลที่จำลองได้จะถูกแบ่งเป็น 2 ส่วนคือ ส่วนแรกสำหรับการสร้างตัวแบบ (70% ของข้อมูล) ส่วนที่สองสำหรับการตรวจสอบตัวแบบ (30% ของข้อมูล) ซึ่งข้อมูลที่ใช้ในการสร้างตัวแบบมาจากข้อมูลชุดเดียวกัน ซึ่งผลการวิจัยแบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 ผลการเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างวิธี Transform และวิธี Delong test

ส่วนที่ 2 การนำวิธี Transform มาใช้ในการเปรียบเทียบพื้นที่ใต้โค้ง ROC

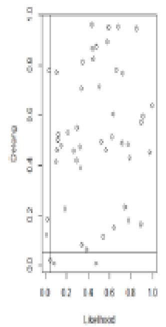
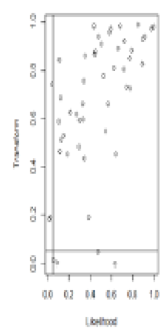
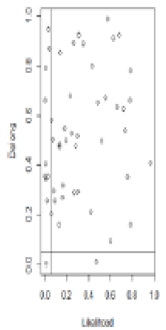
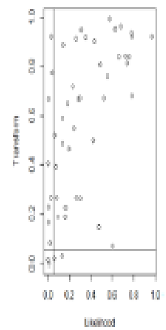
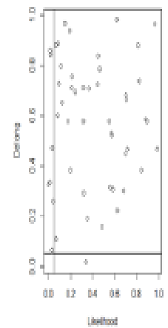
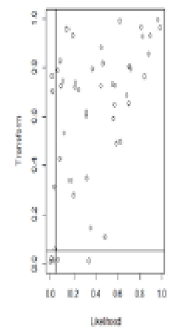
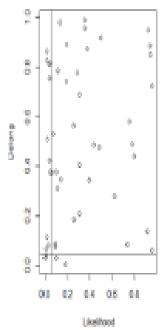
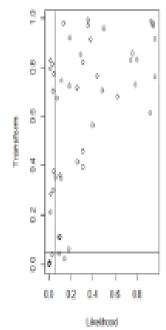
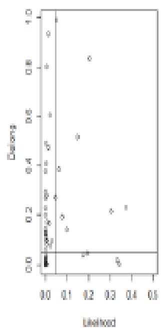
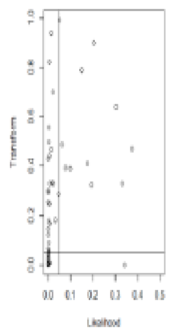
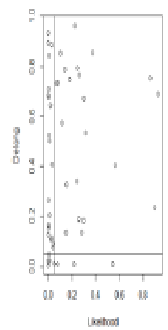
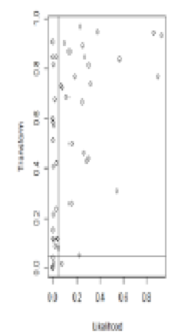
4.1 ส่วนที่ 1: ผลการเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างวิธี Transform และวิธี Delong test

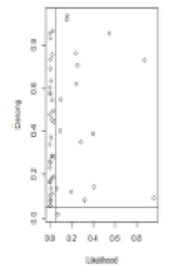
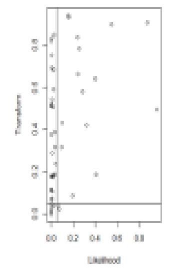
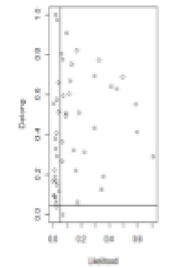
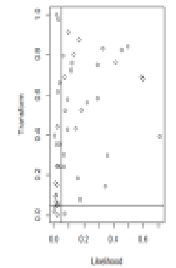
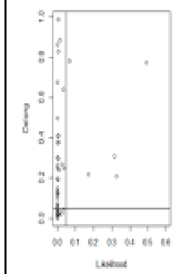
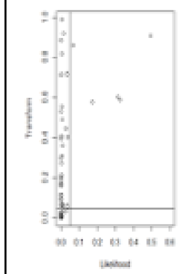
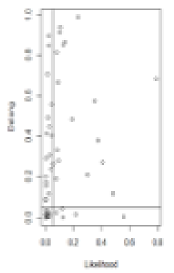
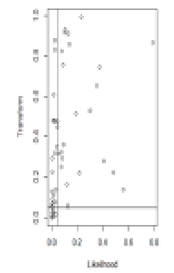
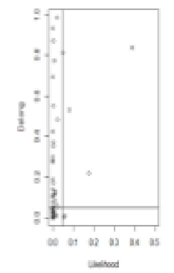
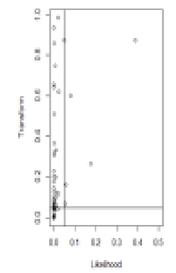
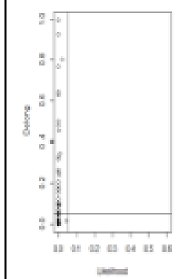
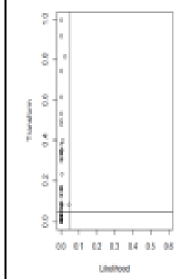
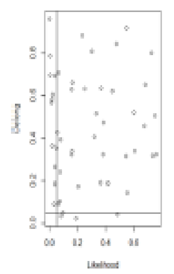
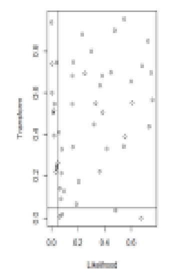
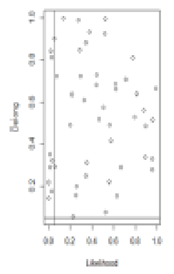
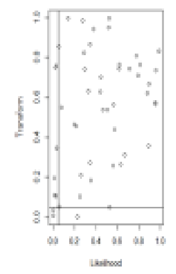
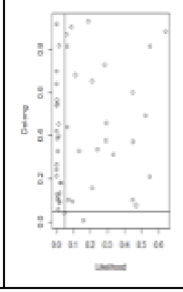
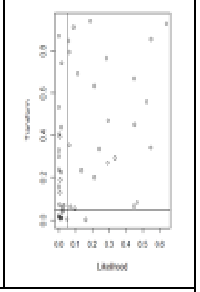
ในส่วนนี้ผู้วิจัยต้องเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างวิธี Transform และวิธี Delong test โดยดูจากความสอดคล้องของค่า p-value ของวิธี Delong test และวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC เทียบกับวิธี Likelihood ratio test ซึ่งพิจารณาจากกราฟของค่า p-value ระหว่างวิธี Delong test และวิธี Likelihood ratio test เมื่อแกน x แทน p-value ของวิธี Likelihood ratio test และแกน y แทน p-value ของวิธี Delong test และกราฟของค่า p-value ระหว่างวิธี Transform และวิธี Likelihood ratio test เมื่อ แกน x

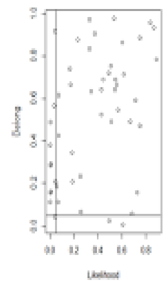
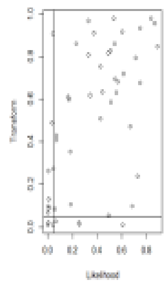
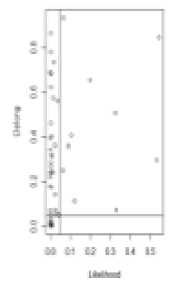
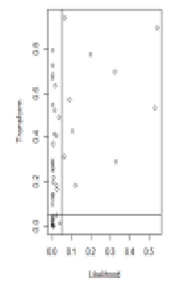
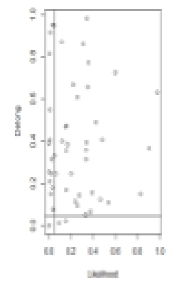
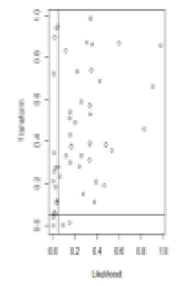
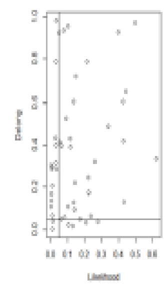
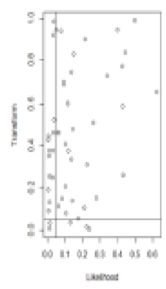
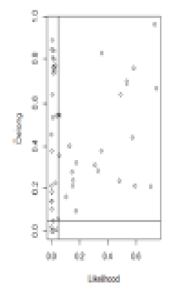
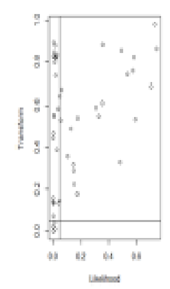
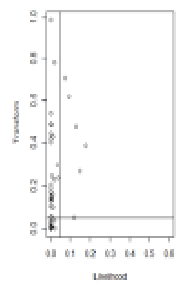
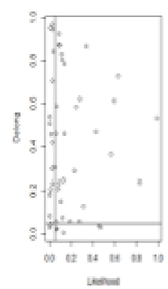
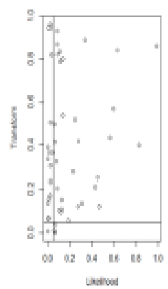
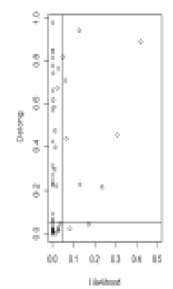
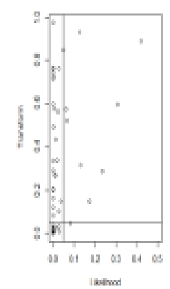
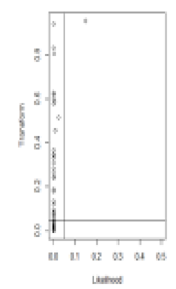
แทน p-value ของวิธี Likelihood ratio test และแกน y แทน p-value ของวิธี Transform ที่ขนาดตัวอย่าง 300 500 และ 1,000 ดังตารางที่ 4.1

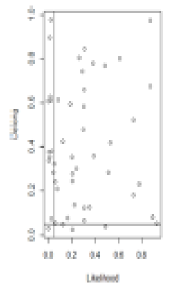
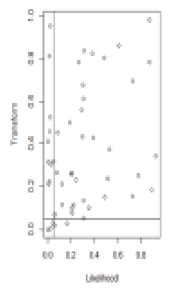
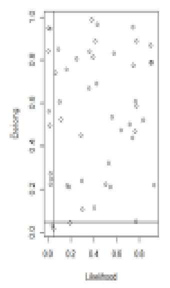
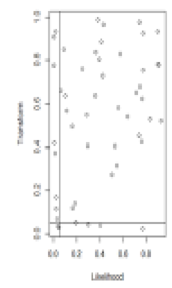
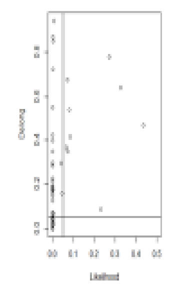
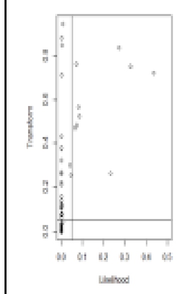
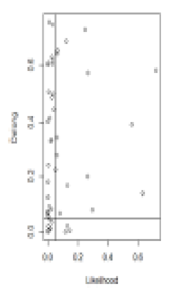
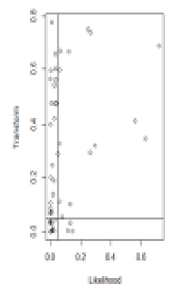
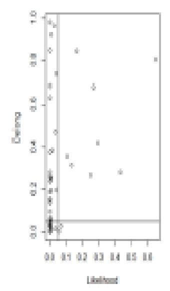
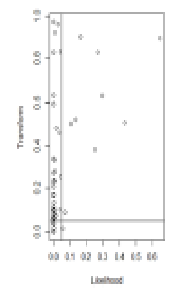
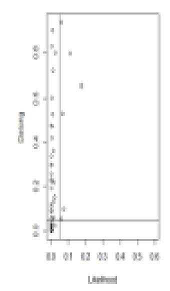
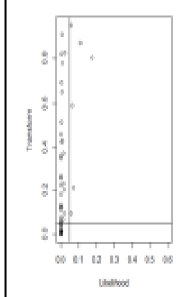
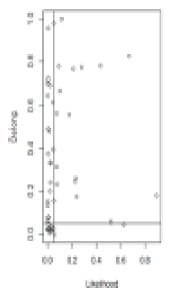
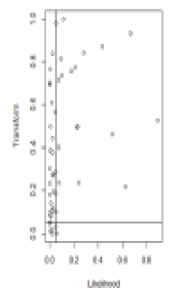
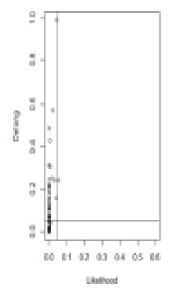
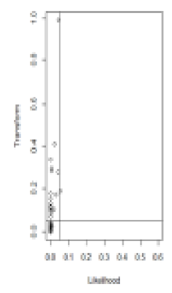
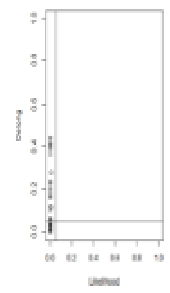
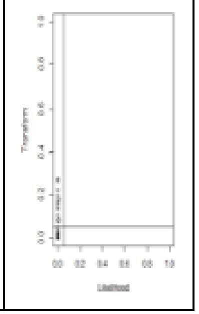
นอกจากนั้นยังผู้วิจัยพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ (correlation) ของค่า p-value ระหว่างวิธี Delong test กับ วิธี Likelihood ratio test และวิธีแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC แล้วนำไปทดสอบด้วย Z-test กับ วิธี Likelihood ratio test เพื่อดูความสอดคล้อง ซึ่งได้ผลการทดสอบดังตารางที่ 4.2 ที่ขนาดตัวอย่าง 300 500 และ 1,000

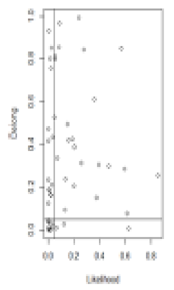
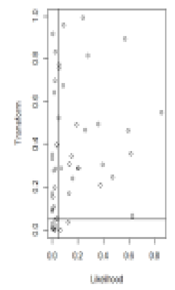
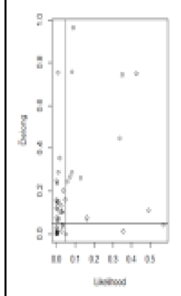
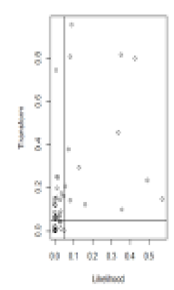
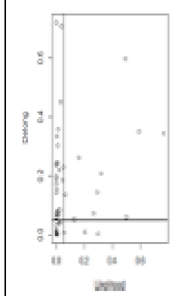
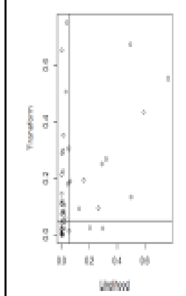
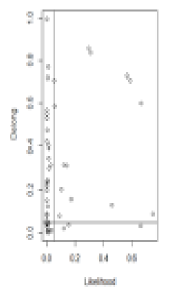
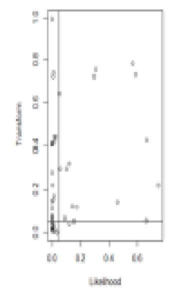
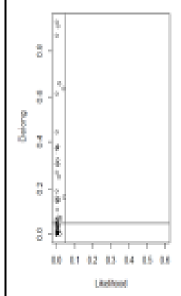
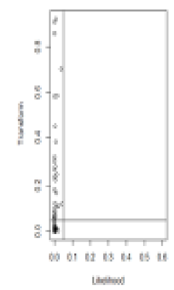
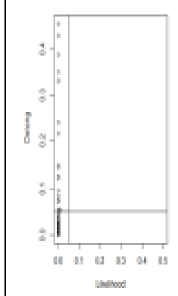
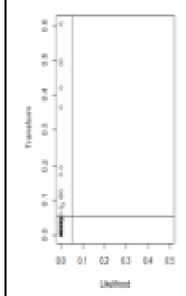
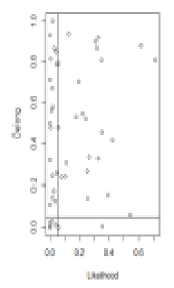
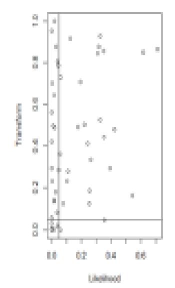
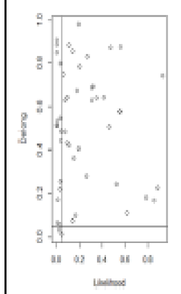
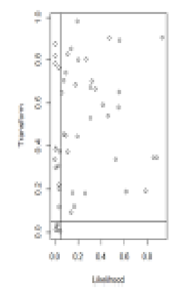
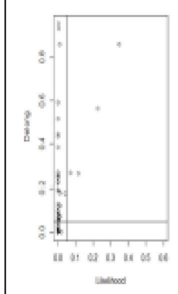
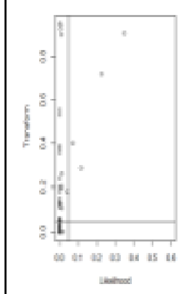
ตารางที่ 4.1 กราฟแสดงค่า p-value จากวิธี Likelihood ratio test เทียบกับวิธี Delong test และ วิธีการแปลงข้อมูล(Transform) ผลต่างพื้นที่ใต้โค้ง ROC

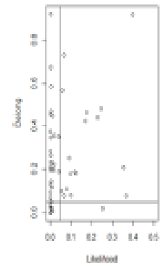
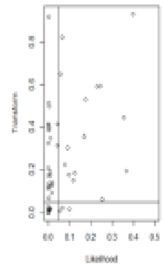
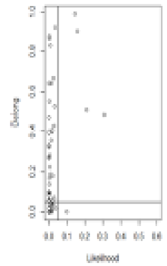
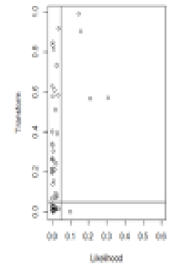
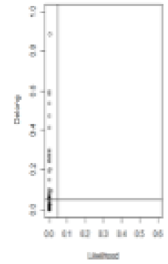
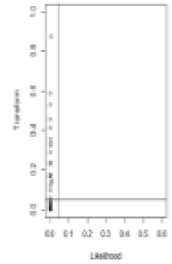
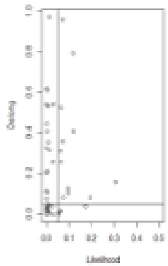
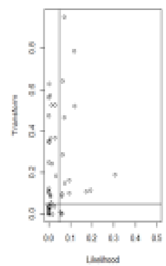
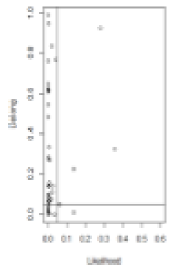
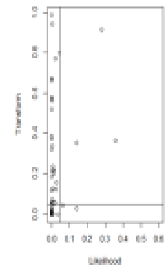
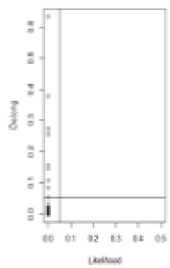
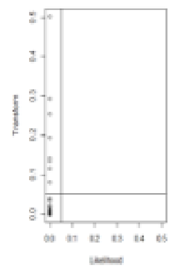
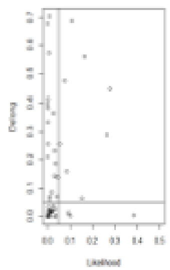
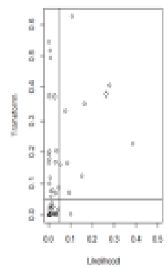
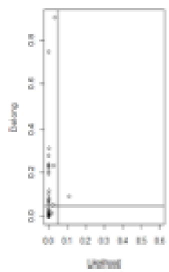
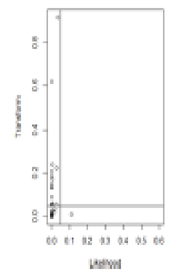
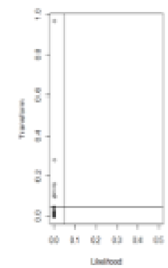
ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
4 ตัวแปร						
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$						

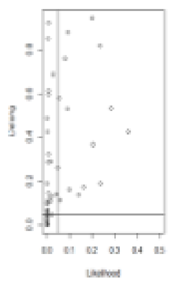
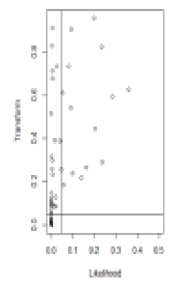
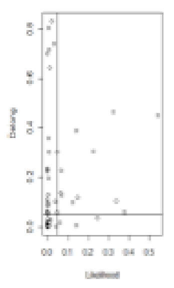
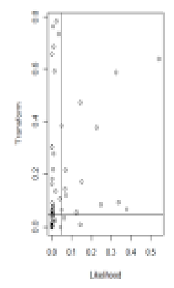
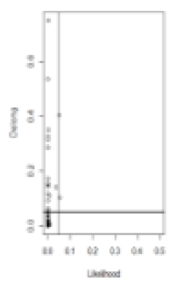
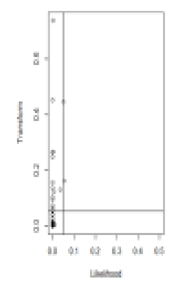
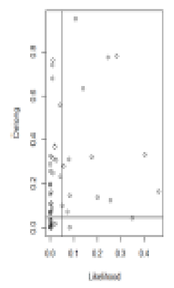
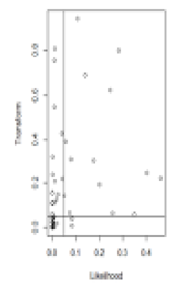
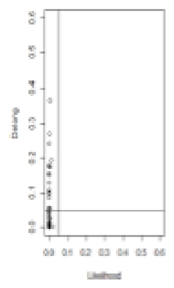
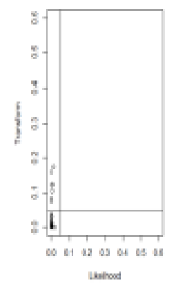
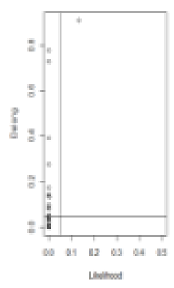
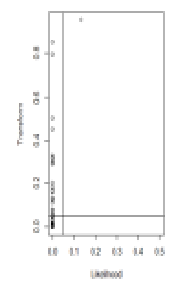
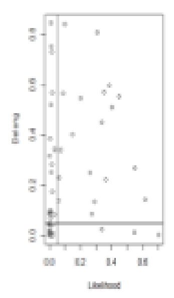
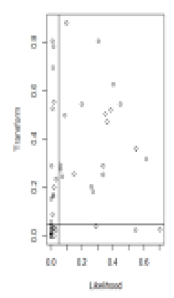
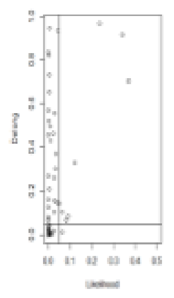
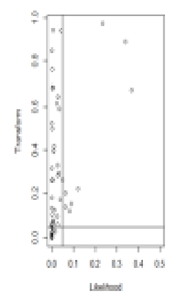
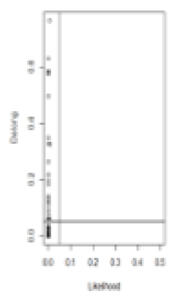
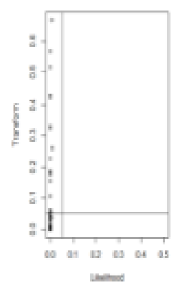
ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving} + \beta_4 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{saving}$						
3 ตัวแปร						

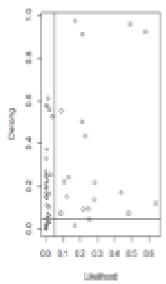
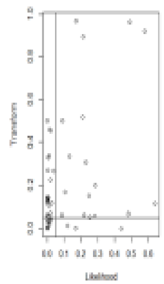
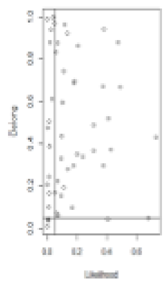
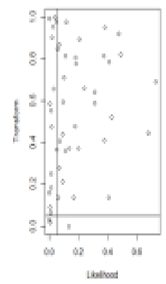
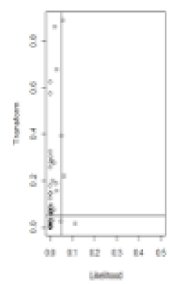
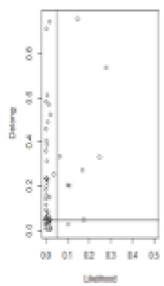
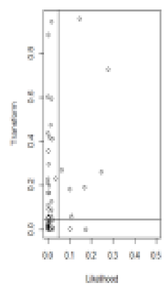
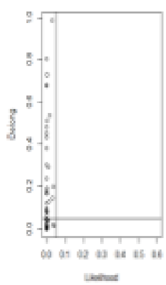
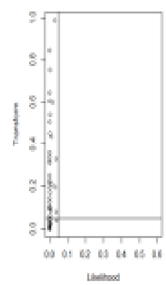
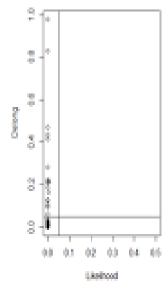
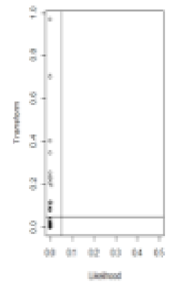
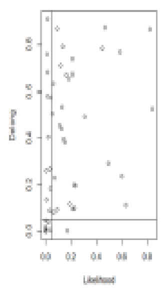
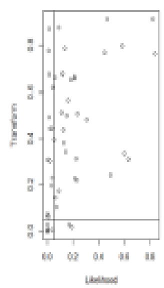
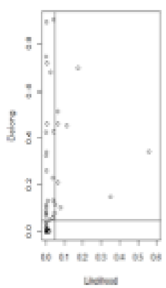
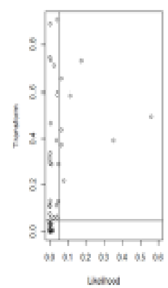
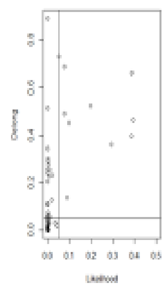
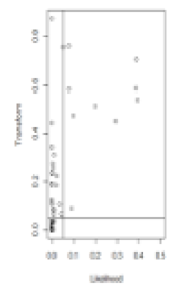
ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving} + \beta_3 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving} + \beta_3 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{present}$						

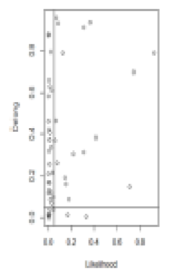
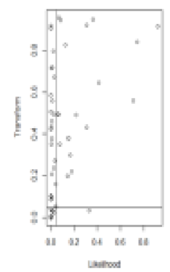
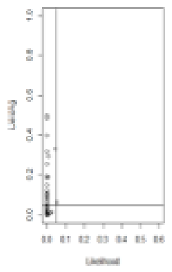
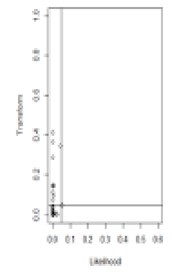
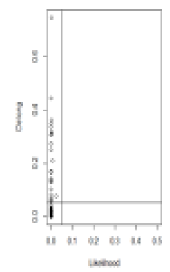
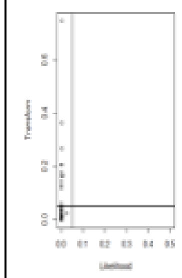
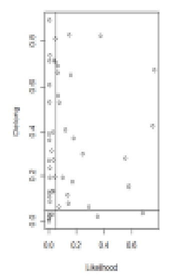
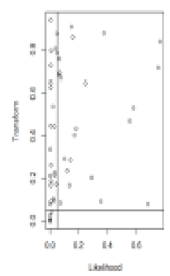
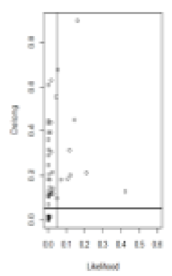
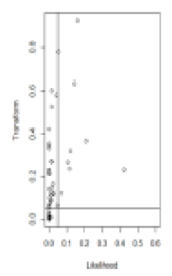
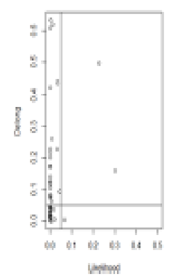
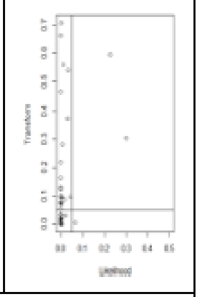
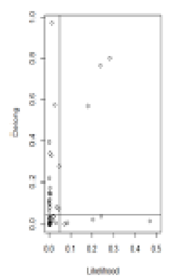
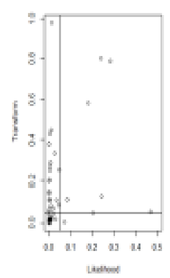
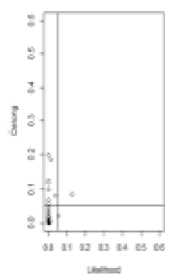
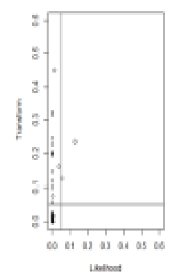
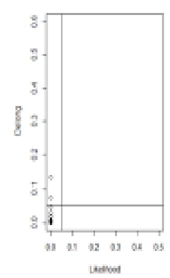
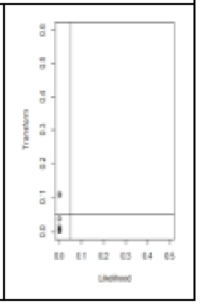
ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving} + \beta_3 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{present}$						

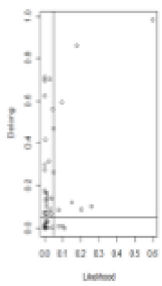
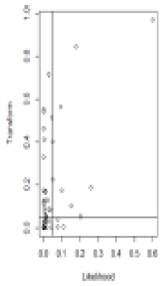
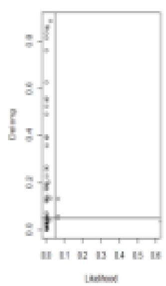
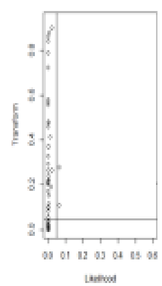
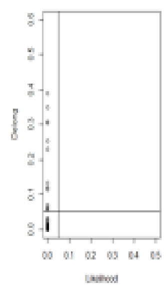
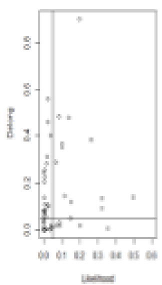
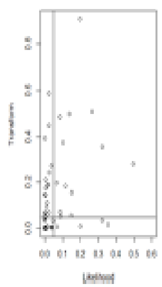
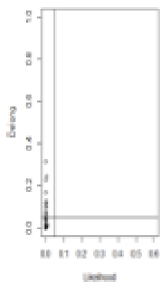
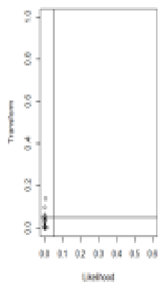
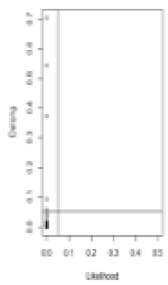
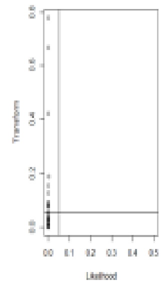
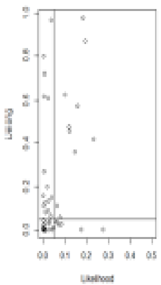
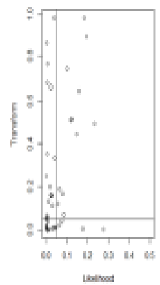
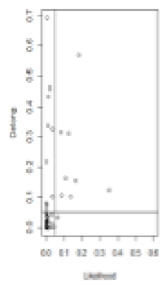
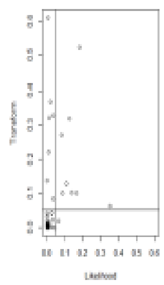
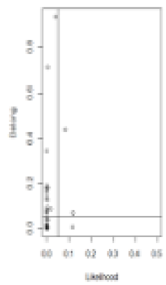
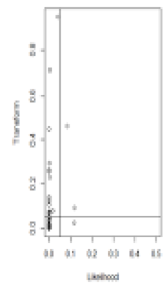
ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving}$						

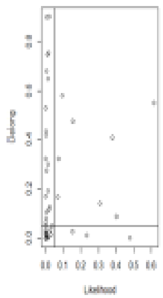
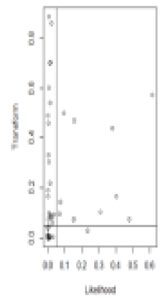
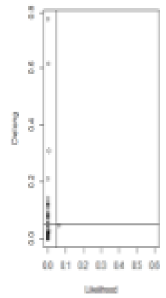
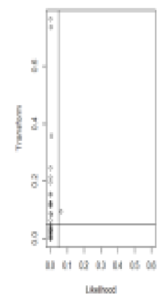
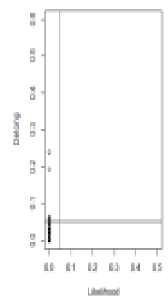
ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing}$						
2 ตัวแปร						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing}$						

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving}$						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing}$						

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving}$						
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving}$						

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{present}$						
$logit(P) = \beta_0 + \beta_1 X_{saving} + \beta_2 X_{present}$						
1 ตัวแปร						
$logit(P) = \beta_0 + \beta_1 X_{gender}$						

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{age}$						
$logit(P) = \beta_0 + \beta_1 X_{housing}$						
$logit(P) = \beta_0 + \beta_1 X_{saving}$						

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{present}$						

ผู้วิจัยทำการเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC ด้วยวิธี Transform และวิธี Delong test โดยพิจารณาจากความสอดคล้องของทั้งสองวิธีกับวิธี Likelihood ratio test จากตารางที่ 4.1 พบว่าเมื่อเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปจำนวน 4 ตัวแปร พบว่ากราฟค่า p-value จากวิธี Transform เทียบกับวิธี Likelihood ratio test มีแนวโน้มเป็นเส้นตรง 45 องศาแต่กราฟค่า p-value จากวิธี Delong test เทียบกับวิธี Likelihood ratio test ค่อนข้างกระจาย นั่นคือ การเปรียบเทียบพื้นที่ใต้โค้ง ROC ด้วยวิธี Transform มีความสอดคล้องกับวิธี Likelihood ration test มากกว่าวิธี Delong test นอกจากนั้นเมื่อพิจารณากราฟจากวิธี Transform บริเวณที่ ค่า p-value ของตัวแบบที่พบนัยสำคัญทางสถิติ (ค่า p-value น้อยกว่า 0.05) จากวิธี Transform และ Likelihood ratio test และพิจารณากราฟจากวิธี Delong test บริเวณที่ ค่า p-value ของตัวแบบที่พบนัยสำคัญทางสถิติจากวิธี Delong test และ Likelihood ratio test พบว่าวิธี Transform มีจำนวนตัวแบบที่พบนัยสำคัญทางสถิติมากกว่าวิธี Delong test นั่นคือวิธี Transform มีจำนวนตัวแบบที่พบนัยสำคัญทางสถิติที่สอดคล้องกับวิธี Likelihood ratio test มากกว่าวิธี Delong test

ตารางที่ 4.2 แสดงค่าสัมประสิทธิ์สหสัมพันธ์ของค่า p-value ระหว่างวิธี Delong test กับ วิธี Likelihood และวิธี Transform กับ วิธี Likelihood ที่ขนาดตัวอย่าง 300 500 และ 1,000

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
4 ตัวแปร						
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$	0.2107	0.5672	0.1524	0.5807	0.0159	0.5156
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$	0.1530	0.5605	0.0192	0.2858	0.1046	0.4423
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving} + \beta_4 X_{present}$	0.0823	0.3724	0.1492	0.3578	0.2277	0.4204
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{present}$	0.1466	0.3043	0.3121	0.3460	0.0512	0.1007
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{saving}$	0.0699	0.2446	0.0174	0.3235	0.1379	0.3427
3 ตัวแปร						
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving} + \beta_3 X_{present}$	0.3872	0.5544	0.2359	0.4401	0.0071	0.3337
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving} + \beta_3 X_{present}$	0.2067	0.3126	0.1531	0.3849	0.1860	0.2776
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{present}$	-0.0176	0.1578	0.1887	0.3004	0.4551	0.4201
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving}$	0.0393	0.2781	0.1404	0.2905	0.3459	0.5097
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving} + \beta_3 X_{present}$	0.1129	0.2991	0.3175	0.4478	0.4134	0.4818
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{present}$	0.0930	0.3179	0.5414	0.6340	0.1255	0.2248
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving}$	0.0186	0.1968	0.2940	0.4174	0.2271	0.4320

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{present}$	0.1878	0.2535	0.1561	0.1851	0.0472	0.1034
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving}$	0.1523	0.2602	-0.0372	0.1749	0.3732	0.4465
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing}$	0.1907	0.3676	0.3151	0.3625	0.1493	0.2509
2 ตัวแปร						
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age}$	0.2397	0.3237	0.1583	0.2242	0.2239	0.2609
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing}$	0.1315	0.3375	0.2020	0.1891	0.4061	0.4040
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving}$	0.1753	0.3017	0.0804	0.2171	0.2026	0.3172
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{present}$	0.2299	0.2781	0.1962	0.3780	0.5986	0.5279
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing}$	0.0885	0.2118	0.4100	0.4116	0.3283	0.3568
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving}$	0.2693	0.3369	0.0126	0.2080	0.1723	0.2651
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{present}$	0.2371	0.2456	0.07974	0.1677	0.0585	0.0686
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving}$	0.2717	0.3828	0.1730	0.3060	0.4879	0.5801
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{present}$	0.2292	0.3724	0.1108	0.2329	0.0556	0.0763
$logit(P) = \beta_0 + \beta_1 X_{saving} + \beta_2 X_{present}$	0.0127	0.2010	0.1809	0.3613	0.0007	0.2503
1 ตัวแปร						
$logit(P) = \beta_0 + \beta_1 X_{gender}$	0.2885	0.3105	0.0692	0.1798	0.0852	0.1259

ตัวแบบลดรูป	N = 300		N = 500		N = 1000	
	Delong test	Transform	Delong test	Transform	Delong test	Transform
$logit(P) = \beta_0 + \beta_1 X_{age}$	0.4152	0.4945	0.1566	0.2325	0.0522	0.1165
$logit(P) = \beta_0 + \beta_1 X_{housing}$	0.1413	0.2854	0.1211	0.2233	0.0016	0.1146
$logit(P) = \beta_0 + \beta_1 X_{saving}$	0.3402	0.3440	0.3012	0.3036	0.2377	0.2406
$logit(P) = \beta_0 + \beta_1 X_{present}$	0.0591	0.1352	0.0187	0.0554	0.1723	0.2336

จากตารางที่ 4.2 จะเห็นได้ว่าวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC มีค่าสัมประสิทธิ์สหสัมพันธ์ของค่า p-value ที่สอดคล้องกับวิธี Likelihood ratio test มากกว่าวิธี Delong test ดังนั้นผู้วิจัยจึงทำการทดสอบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของวิธี Transform มีค่ามากกว่าค่าสัมประสิทธิ์สหสัมพันธ์ของวิธี Delong test อย่างมีนัยสำคัญทางสถิติหรือไม่ โดยใช้สถิติทดสอบ t-test ซึ่งมีสมมติฐานการทดสอบดังนี้

$$H_0: \mu_{delong} \geq \mu_{transform}$$

$$H_a: \mu_{delong} < \mu_{transform}$$

โดยที่

μ_{delong} แทนค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT กับวิธี Delong test

$\mu_{transform}$ แทนค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT กับวิธี Transform

เนื่องจากการใช้สถิติทดสอบ t–test จะต้องมีการทดสอบค่าความแปรปรวนของทั้งสองกลุ่มว่าเท่ากันหรือไม่เท่ากันโดยมีสมมติฐานการทดสอบดังนี้

$$H_0: \sigma^2_{delong} = \sigma^2_{transform}$$

$$H_a: \sigma^2_{delong} \neq \sigma^2_{transform}$$

โดยที่

σ^2_{delong} แทนค่าความแปรปรวนของค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT กับวิธี Delong test

$\sigma^2_{transform}$ แทนค่าความแปรปรวนของค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT กับวิธี Transform

โดยใช้ตัวสถิติทดสอบ F–testซึ่งได้ผลการทดสอบดังตารางที่ 4.3

ตารางที่ 4.3 แสดงค่า p-value ที่เกิดจากการทดสอบความเท่ากันของค่าความแปรปรวนของค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value จากวิธี Delong test และวิธี Transform ที่ขนาด 300 500 และ 1,000

N	p-value	ผลสรุป
300	0.9753	ไม่สามารถปฏิเสธ H_0
500	0.9274	ไม่สามารถปฏิเสธ H_0
1,000	0.7036	ไม่สามารถปฏิเสธ H_0

จากตารางที่ 4.3 พบว่าค่า p-value จากทั้ง 3 ขนาดตัวอย่างมีค่ามากกว่า 0.05 นั่นคือมีหลักฐานไม่เพียงพอที่จะสรุปได้ว่า แทนค่าความแปรปรวนของค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ที่ได้จากวิธี Transform และวิธี Delong test มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 จากผลการทดสอบดังกล่าวทำให้การทดสอบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของ Transform มีค่ามากกว่าค่าสัมประสิทธิ์สหสัมพันธ์ของวิธี Delong test อย่างมีนัยสำคัญทางสถิติหรือไม่ โดยใช้สถิติทดสอบ t-test แบบ Pooled variance ซึ่งได้ผลการทดสอบดังตารางที่ 4.4

ตารางที่ 4.4 แสดงค่า p-value ที่เกิดจากการทดสอบความเท่ากันของค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากวิธี Delong test และวิธี Transform ที่ขนาด 300 500 และ 1,000

N	p-value	ผลสรุป
300	7.699×10^{-10}	ปฏิเสธ H_0
500	1.583×10^{-8}	ปฏิเสธ H_0
1,000	1.423×10^{-5}	ปฏิเสธ H_0

จากตารางที่ 4.4 พบว่าค่า p-value จากทั้ง 3 ขนาดตัวอย่างมีค่าน้อยกว่า 0.05 นั่นคือมีหลักฐานเพียงพอที่จะสรุปได้ว่า ค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากวิธี Transform มีค่ามากกว่าค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากวิธี Delong test มีนัยสำคัญทางสถิติที่ขนาด 300 500 และ 1,000 ที่ระดับนัยสำคัญ 0.05

นอกจากนี้งานวิจัยยังทดสอบว่าที่ขนาดตัวอย่างที่แตกต่างกันนั้นค่าสัมประสิทธิ์สหสัมพันธ์ของค่า p-value ระหว่างวิธี Transform กับ วิธี Likelihood ratio test มีค่าแตกต่างกันอย่างมีนัยสำคัญหรือไม่ โดยใช้ตัวสถิติทดสอบ F-test ซึ่งมีสมมติฐานการทดสอบดังนี้

H_0: $\mu_{300} = \mu_{500} = \mu_{1000}$

H_a: มี $\mu_i \neq \mu_j$ อย่างน้อย 1 คู่ โดยที่ $i \neq j$

โดยที่

μ_{300} แทนค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT กับ วิธี Transform ที่ขนาดตัวอย่าง 300

μ_{500} แทนค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT กับ วิธี Transform ที่ขนาดตัวอย่าง 500

μ_{1000} แทนค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT กับ วิธี Transform ที่ขนาดตัวอย่าง 1000

โดยได้ค่า p-value เท่ากับ 0.7027 ซึ่งมีค่ามากกว่า 0.05 นั่นคือมีหลักฐานไม่เพียงพอที่จะสรุปได้ว่า ค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ของค่า p-value ระหว่างวิธี Transform กับ วิธี Likelihood ratio test ที่ขนาด 300 500 และ 1,000 มีค่าแตกต่างกันที่ระดับนัยสำคัญ 0.05 แสดงให้เห็นว่า ขนาดตัวอย่างไม่มีผลกับประสิทธิภาพของการเปรียบเทียบพื้นที่ใต้โค้ง ROC ด้วยวิธี Delong test

4.2 ส่วนที่ 2: การนำวิธี Transform มาใช้ในการเปรียบเทียบพื้นที่ใต้โค้ง ROC

ในส่วนนี้ผู้วิจัยต้องการเปรียบเทียบพื้นที่ใต้โค้ง ROC ด้วยวิธี Transform เพื่อทดสอบความแตกต่างของพื้นที่ใต้โค้ง ROC ระหว่าง ตัวแบบเต็มรูปและตัวแบบลดรูป ที่ขนาดตัวอย่าง 300 500 และ 1,000 โดยนำข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูปไปแปลงข้อมูลด้วยวิธี Yeo-Johnson Transformation แล้วนำไปใช้ Z-test ในการเปรียบเทียบพื้นที่ใต้โค้ง ซึ่งได้ค่า p-value ดังตารางที่ 4.5

ตารางที่ 4.5 แสดงค่า p-value ของการเปรียบเทียบพื้นที่ใต้โค้ง ROC ด้วยวิธี Transform ที่ขนาด 300 500 และ 1,000

ตัวแบบลดรูป	N=300	N=500	N=1000
4 ตัวแปร			
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$	0.9818	0.5667	0.3832
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving} + \beta_4 X_{present}$	0.5903	0.2346	0.5666
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving} + \beta_4 X_{present}$	0.0352	0.2942	0.0237
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{present}$	0.2958	0.0024	0.0591
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing} + \beta_4 X_{saving}$	0.3992	0.7939	0.1407
3 ตัวแปร			
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving} + \beta_3 X_{present}$	0.8863	0.0073	0.5520
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving} + \beta_3 X_{present}$	0.2151	0.2250	0.0126
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{present}$	0.3927	0.0180	0.1985
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing} + \beta_3 X_{saving}$	0.7919	0.4776	0.1359
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving} + \beta_3 X_{present}$	0.0072	0.0488	0.0131
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{present}$	0.3060	0.0020	0.0014
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing} + \beta_3 X_{saving}$	0.1476	0.1802	0.1979
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{present}$	0.0041	0.0772	0.0000
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{saving}$	0.0643	0.1002	0.1177
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age} + \beta_3 X_{housing}$	0.1197	0.0538	0.0106

ตัวแบบลดรูป	N=300	N=500	N=1000
2 ตัวแปร			
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{age}$	0.2496	0.0464	0.0052
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{housing}$	0.3461	0.0564	0.0038
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{saving}$	0.5731	0.0965	0.0006
$logit(P) = \beta_0 + \beta_1 X_{gender} + \beta_2 X_{present}$	0.1678	0.0081	0.0114
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{housing}$	0.4095	0.2635	0.0044
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{saving}$	0.3155	0.2133	0.0632
$logit(P) = \beta_0 + \beta_1 X_{age} + \beta_2 X_{present}$	0.4247	0.0251	0.0019
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{saving}$	0.1928	0.4716	0.1186
$logit(P) = \beta_0 + \beta_1 X_{housing} + \beta_2 X_{present}$	0.8136	0.0021	0.0110
$logit(P) = \beta_0 + \beta_1 X_{saving} + \beta_2 X_{present}$	0.1563	0.0172	0.0185
1 ตัวแปร			
$logit(P) = \beta_0 + \beta_1 X_{gender}$	0.1118	0.0153	0.0091
$logit(P) = \beta_0 + \beta_1 X_{age}$	0.2916	0.0894	0.0002
$logit(P) = \beta_0 + \beta_1 X_{housing}$	0.0630	0.0001	0.0034
$logit(P) = \beta_0 + \beta_1 X_{saving}$	0.3510	0.0132	0.0069
$logit(P) = \beta_0 + \beta_1 X_{present}$	0.1274	0.0087	0.0318

บทที่ 5
บทสรุปและข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC จากตัวแบบการถดถอยโลจิสติก 2 ตัวแบบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ โดย 2 ตัวแบบคือ ตัวแบบเต็มรูป (full model) และ ตัวแบบลดรูป (reduced model) ซึ่งเป็นแบบจำลองที่ตัวแบบซ้อนกัน (nested model) กัน โดยข้อมูลในการสร้างตัวแบบมาจาก German credit ใน package caret ในโปรแกรม R ที่ขนาดตัวอย่าง 300 500 และ 1,000 ตัวแปรต้นที่ใช้ในการพยากรณ์โอกาสในการชำระหนี้ คือ อายุ เพศ เงินฝาก ประเภทที่อยู่อาศัย และระยะเวลาในการทำงาน โดยมีสรุปผลวิจัยดังนี้

5.1 สรุปผลการวิจัย

จากการศึกษา การเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพื้นที่ใต้โค้ง ROC ด้วยวิธี Transform และวิธี Delong test โดยดูจากความสอดคล้องของค่า p-value ของวิธี Delong test และวิธี Transform เทียบกับวิธี Likelihood ratio test ซึ่งพิจารณาจากกราฟของค่า p-value ระหว่างวิธี Delong test และวิธี Likelihood ratio test พบว่าเมื่อเปรียบเทียบตัวแบบเต็มรูปกับตัวแบบลดรูปจำนวนตัวแปร 4 ตัวแปรกราฟค่า p-value จากวิธี Transform เทียบกับวิธี Likelihood ratio test มีแนวโน้มเป็นเส้นตรง 45 องศาแต่กราฟค่า p-value จากวิธี Delong test เทียบกับวิธี Likelihood ratio test ค่อนข้างกระจาย และเมื่อพิจารณากราฟจากวิธี Transform บริเวณที่ค่า p-value ของตัวแบบที่พบนัยสำคัญทางสถิติจากวิธี Transform และ Likelihood ratio test และพิจารณากราฟจากวิธี Delong test บริเวณที่ ค่า p-value ของตัวแบบที่พบนัยสำคัญทางสถิติจากวิธี Delong test และ Likelihood ratio test พบว่าวิธี Transform มีจำนวนตัวแบบที่พบนัยสำคัญทางสถิติมากกว่าวิธี Delong test นั่นคือวิธี Transform มีความสอดคล้องกับวิธี Likelihood ratio test มากกว่าวิธี Delong test นอกจากการพิจารณาความสอดคล้องของทั้ง 2 วิธี กับวิธี Likelihood ratio test จากกราฟผู้วิจัยยังพิจารณาความสอดคล้องจากค่าสัมประสิทธิ์สหสัมพันธ์ (correlation) ของค่า p-value ระหว่างวิธี Delong test กับ วิธี Likelihood ratio test และวิธี Transform กับ

วิธี Likelihood ratio test เพื่อดูความสอดคล้องของทั้ง 2 วิธีกับวิธี ratio test พบว่าที่ขนาดตัวอย่าง 300 500 และ 1,000 ค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากวิธี Transform มีค่ามากกว่าค่าเฉลี่ยค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากวิธี Delong test อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 และพบว่าที่ขนาดตัวอย่าง 300 500 และ 1,000 มีค่าเฉลี่ยสัมประสิทธิ์สหสัมพันธ์ของค่า p-value ระหว่างวิธี Transform กับ วิธี Likelihood ratio test ไม่แตกต่างกันที่ระดับนัยสำคัญ 0.05 นั่นคือขนาดตัวอย่างไม่มีผลกับประสิทธิภาพของการเปรียบเทียบพื้นที่ใต้โค้ง ROC ด้วยวิธี Delong test

5.2 อภิปรายผล

การศึกษานี้เป็นการศึกษาตัวแบบโลจิสติกซึ่งสามารถใช้วิธี LRT ในการเปรียบเทียบ 2 ตัวแบบได้ แต่ยังมีตัวแบบทางสถิติจำนวนมากที่ไม่สามารถเปรียบเทียบได้ด้วยวิธี LRT เช่น Support Vector Machine เป็นต้น ดังนั้นจึงต้องมีวิธีการที่ใช้ในการเปรียบเทียบพื้นที่ใต้โค้ง ROC แต่เนื่องจากวิธี Delong test เป็นวิธีการเปรียบเทียบพื้นที่ ROC ที่ไม่เหมาะกับข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ที่ไม่มีการแจกแจงปกติ ดังนั้นงานวิจัยนี้จึงทำการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ให้มีการแจกแจงแบบปกติแล้วนำไปทดสอบด้วย Z-test แล้วนำผลมาเปรียบเทียบกับวิธี Delong test โดยวิธีการแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC ให้มีการแจกแจงแบบปกติแล้วนำไปทดสอบด้วย Z-test ใหม่ที่ผู้วิจัย เรียกว่า วิธี Transform

การศึกษาความแตกต่างของพื้นที่ใต้โค้ง ROC โดยการเปรียบเทียบระหว่างวิธี Delong test และวิธี Transform พบว่าวิธี Transform มีค่าค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ที่มีความสอดคล้องกับวิธี LRT มากกว่าวิธี Delong test เนื่องจากการแปลงข้อมูลผลต่างพื้นที่ใต้โค้งด้วยวิธี Yeo-Johnson Transformation จะทำให้ข้อมูลผลต่างพื้นที่ใต้โค้งมีการแจกแจงแบบปกติจึงมีความเหมาะสมที่จะใช้ Z-test ในการเปรียบเทียบความแตกต่างของพื้นที่ใต้โค้งมากกว่าข้อมูลผลต่างพื้นที่ใต้โค้งที่ไม่มีการแจกแจงแบบปกติ ซึ่งสอดคล้องกับ Demler et al. (2012) ที่กล่าวไว้ถึงการนำ Delong test ไปใช้ในทางที่ผิดเนื่องจากข้อมูลผลต่างพื้นที่ใต้โค้งไม่ได้มีการแจกแจงแบบปกติ นอกจากนั้นพบว่าขนาดตัวอย่างที่แตกต่างกันในการสร้างตัวให้ค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ไม่แตกต่างกัน

การนำวิธี Transform ไปใช้ในการเปรียบเทียบพื้นที่ใต้โค้ง ROC จะทำการจำลองข้อมูลตัวอย่างด้วยวิธีการสุ่มแบบ Bootstrap แบบคืนกลับจำนวน 50 ชุดข้อมูลซึ่งข้อมูลที่สุ่มได้นำมาเป็นตัวแทนของข้อมูลตัวอย่างในการหาค่าพารามิเตอร์เพื่อนำไปแปลงข้อมูลตัวอย่าง ด้วยวิธี Yeo-Johnson Transformation การสุ่มแบบ Bootstrap ทำให้ได้ค่าพารามิเตอร์ที่เหมาะสมกับข้อมูลตัวอย่าง เนื่องจากเป็นการสุ่มข้อมูลจากข้อมูลตัวอย่าง ทำให้พบว่าการทดสอบความแตกต่างของพื้นที่ใต้โค้ง ROC ด้วยวิธี Transform มีความสอดคล้องกับวิธี Likelihood ratio test มากกว่าวิธี Delong test

5.3 ข้อเสนอแนะ

(1) ในงานวิจัยนี้ผู้วิจัยสร้างตัวแบบจำนวน 50 ตัวแบบเพื่อนำไปแปลงข้อมูลผลต่างพื้นที่ใต้โค้ง ROC การเปลี่ยนจำนวนตัวแบบอาจจะส่งผลกระทบต่อค่าสัมประสิทธิ์สหสัมพันธ์ของ p-value ระหว่างวิธี LRT และวิธี Transform

(2) ในงานวิจัยนี้ทำการเปรียบเทียบพื้นที่ใต้โค้ง ROC จากตัวแบบโลติสจิก ผู้วิจัยอื่นอาจทำการเปรียบเทียบพื้นที่ใต้โค้ง ROC จากตัวแบบอื่นๆที่น่าสนใจ เช่น Support Vector Machine ตัวแบบโพรบิท เป็นต้น

(3) ในงานวิจัยนี้ทำการเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูป ผู้วิจัยอื่นอาจทำการเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบอื่นๆที่ไม่ใช่ตัวแบบที่ซ้อนทับกัน (nested model) เช่น ในงานวิจัยนี้อาจเปรียบเทียบพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบ default~gender+age+housing กับ ตัวแบบ default~present+housing เป็นต้น

รายการอ้างอิง

Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of mathematical psychology, 12*(4), 387-415.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 211-252.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application* (Vol. 1): Cambridge university press.

Debbaut, P., Ghent, A. C., & Kudlyak, M. (2014). Are young borrowers bad borrowers? Evidence from the Credit CARD Act of 2009.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845.

Demler, O. V., Pencina, M. J., & D'Agostino, R. B. (2012). Misuse of DeLong test to compare AUCs for nested models. *Statistics in Medicine, 31*(23), 2577-2587.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics, 7*, 1-26.

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. In. Wiley: New York.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29-36.

Kuhn, M. (2008). Caret package. *Journal of statistical software, 28*(5), 1-26.

Latoya, I. (2018). How Long Does it Take for Your Credit Score to Improve. Retrieved from https://www.thebalance.com/how-long-does-it-take-for-your-credit-score-to-improve-960555

Lopes, P. (2008). Credit card debt and default over the life cycle. *Journal of Money, Credit and Banking, 40*(4), 769-790.

McClish, D. K. (1989). Analyzing a portion of the ROC curve. *Medical Decision Making, 9*(3), 190-195.

Tamara, E. H. (2016). Credit card race, age, gender statistics. Retrieved from https://www.creditcards.com/credit-card-news/race-age-gender-statistics.php

Tape. (n.d.). The Area Under an ROC Curve. Retrieved from http://gim.unmc.edu/dxtests/ROC3.htm

Tukey, J. W. (1957). On the comparative anatomy of transformations. *The Annals of Mathematical Statistics*, 602-632.

White, M. J. (2007). Bankruptcy reform and credit cards. *Journal of Economic Perspectives, 21*(4), 175-200.

Yeo, I. K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika, 87*(4), 954-959.

ภาคผนวก

คำสั่งการวิเคราะห์ข้อมูลด้วยโปรแกรม R

```
library(pROC)
library(car)
library(lmtest)
library(plyr)

#read data
data<-read.table('C:/Users/Administrator/Desktop/sheet3.csv',header=T,sep=",")
attach(data)

# 1 st simulation for N=1000
data1<-data[sample(nrow(data),1000),]

#Create 50 models
diffauc<-c()
z<-c()
likelihood<-c()
ftest<-c()
LR<-c()
for(i in 1:50){
set.seed(6401+141*i)
data2<-data1[sample(nrow(data1),1000,replace=TRUE),]
train<-data2[sample(nrow(data2),floor(0.7*1000)),]
test<-data2[sample(nrow(data2),nrow(data2)-floor(0.7*1000)),]
```

```
fullmodel<-
glm(default~gender+age+housing+savings+present,data=train,family=binomial(link=
"logit"))
predict1<-predict(fullmodel,newdata=test)
roc1<-roc(test$default,predict1)
reducedmodel<-
glm(default~age+housing+savings+present,data=train,family=binomial(link="logit"))
predict2<-predict(reducedmodel,newdata=test)
roc2<-roc(test$default,predict2)
de<-roc.test(roc1,roc2,method="delong")
Z<-de$statistic
z<-c(z,Z)
diff=auc(roc1)-auc(roc2)
diffauc<-c(diffauc,diff)
LRT<-deviance(creditlog2)-deviance(creditlog1)
like<-1-pchisq(LRT,df.residual(creditlog2)-df.residual(creditlog1))
likelihood<-c(likelihood,like)
}

#p-value from Delong test
delong<-2*pnorm(-abs(z))
## p-value from Transform
po<-powerTransform(diffauc~1, family="yjPower")
da<-yjPower(diffauc,po$start)
s<-sd(da)
za<-da/s
transform<-2*pnorm(-abs(za))
```

```
#Correlation Coefficient
cor(likelihood,delong)
cor(likelihood,transform)

#Comparing AUC by Transform
train1<-data1[sample(nrow(data1),floor(0.7*1000)),]
test1<-data1[sample(nrow(data1),nrow(data1)-floor(0.7*1000)),]
fullmodel<-
glm(default~gender+age+housing+savings+present,data=train1,family=binomial(link="
logit"))
predict1<-predict(fullmodel,newdata=test1)
roc1<-roc(test1$default,predict1)
reducedmodel<-
glm(default~age+housing+savings+present,data=train1,family=binomial(link="logit"))
predict2<-predict(reducedmodel,newdata=test1)
roc2<-roc(test1$default,predict2)
diffauc1<-auc(roc1)-auc(roc2)
da1<-yjPower(diffauc1,po$start)
za1<-da1/s
2*pnorm(-abs(za1))
```

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวเบญจพร เอี่ยมประโคน เกิดวันอาทิตย์ที่ 30 ตุลาคม พ.ศ. 2537 สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาคณิตศาสตร์ เกียรตินิยมอันดับหนึ่ง ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2558 และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาวิชาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2559