การเปรียบเทียบพลังการทดสอบของการวัดภาวะสารูปดีเมื่อกำหนดตัวแบบไม่ถูกต้อง สำหรับตัวแปรตอบสนองเชิงกลุ่มแบบมีลำดับ

โดย
นายสุจินต์ สุขกุมภาพันธ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2548

ISBN 974-11-6143-3

POWER COMPARISONS OF GOODNESS-OF-FIT MEASURES OF MODEL MISSPECIFICATIONS FOR ORDINAL RESPONSE CATEGORIES

By

Sujin Sukgumpapan

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

Department of Statistics

Graduate School

SILPAKORN UNIVERSITY

2005

ISBN 974-11-6143-3

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง “การเปรียบเทียบพลังการทดสอบของการวัดภาวะสารูปดีเมื่อกำหนดตัวแบบไม่ถูกต้องสำหรับตัวแปรตอบสนองเชิงกลุ่มแบบมีลำดับ” เสนอโดย นายสุจินต์ สุขกุมภาพันธ์ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์

....
(รองศาสตราจารย์ ดร. วิสาข์ จัติวัตร์)
รองอธิการบดีฝ่ายวิชาการ รักษาการแทน
คณบดีบัณฑิตวิทยาลัย
วันที่......เดือน..................พ.ศ............

ผู้ควบคุมวิทยานิพนธ์

รองศาสตราจารย์ วีรานันท์ พงศาภักดี

คณะกรรมการตรวจสอบวิทยานิพนธ์

.... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ปราณี นิลกรณ์)
........./............/............

.... กรรมการ
(รองศาสตราจารย์ วีรานันท์ พงศาภักดี)
........./............/............

.... กรรมการ
(รองศาสตราจารย์ วัฒนา เกาศัลย์)
........./............/............

.... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. กมลชนก พานิชการ)
........./............/............

.... กรรมการ
(รองศาสตราจารย์ ดร. กมล บุษบา)
........./............/............

K 46304202 : สาขาวิชาสถิติประยุกต์

คำสำคัญ : การทดสอบภาวะสารูปดี ตัวแบบโลจิทมัลติโนเมียล ตัวแบบโลจิทที่วางนัยทั่วไป พลังการทดสอบ

สุจินต์ สุขกุมภาพันธ์ : การเปรียบเทียบพลังการทดสอบของการวัดภาวะสารูปดีเมื่อกำหนดตัวแบบไม่ถูกต้อง สำหรับตัวแปรตอบสนองเชิงกลุ่มแบบมีลำดับ (POWER COMPARISONS OF GOODNESS-OF-FIT MEASURES OF MODEL MISSPECIFICATIONS FOR ORDINAL RESPONSE CATEGORIES) อาจารย์ผู้ควบคุมวิทยานิพนธ์ : รศ.วีรานันท์ พงศาภักดี. 82 หน้า ISBN 974-11-6143-3

การทดสอบภาวะสารูปดีของตัวแบบเชิงเส้นที่วางนัยทั่วไป (GLMs) ในปัจจุบันมีตัวสถิติให้เลือกใช้หลากหลาย แต่ตัวแบบ GLMs ที่มีตัวแปรตอบสนองแบบมีลำดับยังมีตัวสถิติทดสอบให้เลือกอย่างจำกัด งานวิจัยนี้มีจุดประสงค์ 2 ประการคือ 1. ศึกษาพลังการทดสอบของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของตัวแบบ GLMs ต่าง ๆ และ 2. เปรียบเทียบพลังการทดสอบของตัวสถิติในข้อ 1. ระหว่างตัวแบบที่ใช้ตัวแปรตอบสนองแบบมีลำดับและไม่มีลำดับ การวิจัยนี้เน้นตัวแบบตามสมมติฐานทางเลือก **partition** ซึ่งเป็นตัวแบบที่สอดคล้องกับการแบ่งส่วนของค่าพยากรณ์ของตัวแปรตอบสนองซึ่งแบ่งเป็น 10 กลุ่มตามลำดับของค่าพยากรณ์ตัวแปรตอบสนอง และในขณะเดียวกันก็ได้ศึกษาพลังการทดสอบตัวแบบตามสมมติฐานทางเลือก **pairwise** ซึ่งเป็นตัวแบบที่มีเทอมผลคูณระหว่างตัวแปรอธิบาย ทั้งสองตัวแบบใช้ในการตรวจสอบภาวะสารูปดีเทียบกับตัวแบบภายใต้สมมติฐานว่าง ซึ่งเป็นตัวแบบง่าย (minimal) ข้อมูลที่ใช้ในการศึกษาได้มาจากการสร้างแบบจำลองข้อมูลที่มีขนาด 600, 1000 และ1500 กำหนดเซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ เป็น { -0.6, 0.2, 0.707, -0.707} และใช้ตัวแปรอธิบาย $X_1 \sim$Bernoulli(0.5), $X_2 \sim$Normal(0,1) และค่าของตัวแปรตอบสนอง Y จำลองแบบจากตัวแบบที่ถูกต้องหรือตัวแบบจริงที่กำหนด

ผลการวิจัยพบว่าเมื่อกำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ **pairwise** และให้ β_{12} ซึ่งเป็นพารามิเตอร์ที่สอดคล้องกับเทอมผลคูณของตัวแปรอธิบาย x_1 และ x_2 มีค่าตั้งแต่ 0-1.4 จากแบบจำลอง 500 ชุด ให้ค่าพลังการทดสอบจากตัวแบบ **pairwise** หรือตัวแบบที่ถูกต้อง สูงกว่าและลู่เข้าสู่ 100% เร็วกว่าพลังการทดสอบจากตัวแบบ **partition** โดยเฉพาะเมื่อกำหนดขนาดตัวอย่างมากขึ้น แต่เมื่อกำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ **quadratic** และให้ β_{22} มีค่าตั้งแต่ 0-1.4 พบว่าพลังการทดสอบจากตัวแบบ **partition** สูงกว่าพลังการทดสอบของตัวแบบ **pairwise** และมีค่าลู่เข้าสู่ 100% เมื่อ ค่า β_{22} ซึ่งเป็นพารามิเตอร์ที่สอดคล้องกับเทอมกำลังสองของตัวแปรอธิบาย x_2 มีค่ามากขึ้น ในขณะที่พลังการทดสอบจากตัวแบบ **pairwise** ต่ำและไม่ลู่เข้าสู่ 100% จึงชี้ให้เห็นว่าในการทดสอบภาวะสารูปดีของตัวแบบตามสมมติฐานว่าง **minimal** เนื่องจากในทางปฏิบัตินั้นเราไม่ทราบว่าตัวแบบจริงเป็นตัวแบบใดจึงควรที่จะใช้ตัวแบบตามสมมติฐานทางเลือกเป็นตัวแบบ **partition** โดยใช้ร่วมกับตัวสถิติอัตราส่วนภาวะน่าจะเป็น

การเปรียบเทียบพลังการทดสอบระหว่างตัวแบบที่ใช้ตัวแปรตอบสนองแบบมีลำดับ และไม่มีลำดับ จากผลการศึกษาในทุกขนาดตัวอย่างพบว่าเมื่อกำหนดให้ตัวแปรตอบสนองเป็นแบบมีลำดับแล้วการทดสอบสอบภาวะสารูปดีที่ตัวแปรตามเป็นแบบไม่มีลำดับภายใต้ตัวแบบ **baseline-category logit** และตัวแบบ **baseline-partition** ทำให้สูญเสียพลังการทดสอบไปบางส่วน โดยให้ผลลัพธ์ทำนองเดียวกันเมื่อกำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ **pairwise** และตัวแบบ **quadratic** ด้วย

ภาควิชาสถิติ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ปีการศึกษา 2548

ลายมือชื่อนักศึกษา ....

ลายมือชื่ออาจารย์ผู้ควบคุมวิทยานิพนธ์ ....

K46304202: MAJOR : APPLIED STATISTICS

KEYWORDS: GOODNESS-OF-FIT TESTS / MULTINOMIAL LOGIT MODEL/ GENERALIZED LOGIT MODEL/ POWER

SUJIN SUKGUMPHAPHAN : POWER COMPARISONS OF GOODNESS-OF-FIT MEASURES OF MODEL MISSPECIFICATIONS FOR ORDINAL RESPONSE CATEGORIES. THESIS ADVISOR: ASSOC.PROF. VEERANUN PONGSAPUKDEE. 82 pp. ISBN 974-11-6143-3.

Currently there are many statistics for testing goodness-of-fit for generalized linear models or GLMs but there still be limited global measures for the goodness-of-fit for ordinal response GLMs. This research is conducted for two purposes, firstly, to evaluate the power of the tests of the likelihood ratio and Wald statistics for several GLMs, secondly, to compare the power of the tests between the models with ordinal and nominal response.

This present study emphasizes the power of the test for the alternative **partition** models that could correspond to the method for partitioning the covariate space into 10 regions under the ordered predicted probabilities of the response. Meanwhile, another alternative **pairwise** model which possesses the interaction between two explanatory variables is also studied. Both alternatives models are compared with the null hypothesis **minimal** model. Data were simulated under the sample sizes of 600, 1000 and 1500, given the set of parameters $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ of { -0.6, 0.2 0, 0.707, -0.707} and the explanatory variables are X_1~Bernoulli(0.5) and X_2~Normal(0,1). The outcome, y's were generated through the correctly specified models.

The research results show that given the correct **pairwise** model, for varied β_{12} from 0-1.4, with 500 simulation runs, the power of the tests from the **pairwise** model are more rapidly approach to 100% than those of the **partition** model, especially, when the sample sizes are increased. But these results are directly opposite when the correct model is the **quadratic** model for varied β_{22} from 0-1.4. Thus, the **partition** model together with the likelihood ratio statistic is more stable and appropriate for testing goodness-of-fit for the **minimal** model, due to in practice, we don't know the true model.

The comparative studies between fitting the ordinal and the nominal response categories to the ordinal response data, show that we do lost some power of the tests when testing goodness-of-fit for the **baseline-category logit** model instead of the **baseline-partition** model. These results are revealed similarly for both the correct **pairwise** and **quadratic** models.

Department of Statistics Graduate School, Silpakorn University Academic Year 2005

Student's signature ………………………………

Thesis Advisor's signature …………………….……..…

กิตติกรรมประกาศ

การศึกษาและเรียบเรียงวิทยานิพนธ์ฉบับนี้ สามารถเสร็จสมบูรณ์เป็นรูปเล่ม และเสร็จลุล่วงไปได้ด้วยดี ผู้เขียนได้รับความอนุเคราะห์และกรุณาจากรองศาสตราจารย์วีรานันท์ พงศาภักดี อาจารย์ผู้ควบคุมวิทยานิพนธ์ ที่ให้คำแนะนำและช่วยเหลือในการเรียบเรียงวิทยานิพนธ์ และแก้ไขข้อบกพร่องต่าง ๆ ด้วยความเมตตา เอาใจใส่อย่างยิ่ง

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.ปราณี นิลกรณ์ รองศาสตราจารย์ วัฒนา เกาศัลย์ ผู้ช่วยศาสตราจารย์ ดร. กมลชนก พานิชการ รองศาสตราจารย์ ดร. กมล บุษบา กรรมการตรวจสอบวิทยานิพนธ์ ที่กรุณาให้คำแนะนำแก้ไข เพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบคุณบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากรที่กรุณาให้ความอนุเคราะห์ สนับสนุนและอนุมัติให้ดำเนินงานวิจัยในครั้งนี้

ขอขอบคุณเพื่อน ๆ และรุ่นพี่คณะที่ให้คำแนะนำต่าง ๆ ซึ่งเป็นส่วนสำคัญที่ทำให้การเรียบเรียงงานวิจัยฉบับนี้เป็นไปด้วยดี

สุดท้ายขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และญาติพี่น้องในความรัก ความอบอุ่นที่มีให้ตลอดเวลา กำลังใจในยามท้อถอยและการสนับสนุนอย่างดีมาโดยตลอด และขอขอบคุณบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร สำหรับทุนอุดหนุนในการทำวิจัยครั้งนี้

ผู้วิจัยขออำนาจคุณพระศรีรัตนตรัย ดลบันดาลให้ทุกท่านจงมีแต่ความสุข ความเจริญประสบแต่ความสำเร็จ ไร้โรคภัยทั้งในโลกนี้และโลกหน้า

สารบัญ

บทที่ หน้า

สารบัญตาราง

ตารางที่ หน้า

สารบัญรูป

รูปที่ หน้า

รูปที่		หน้า

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวิเคราะห์ข้อมูลเชิงกลุ่มเป็นวิธีการวิเคราะห์ข้อมูลเชิงสถิติที่สนใจศึกษาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ โดยเฉพาะเมื่อข้อมูลจัดอยู่ในรูปตารางการณ์จรที่สอดคล้องกับวิธีการวางแผนการสุ่มตัวอย่าง วิธีวิเคราะห์เน้นศึกษาตัวแปรตอบสนองเชิงกลุ่ม (categorical response variable) เป็นสำคัญ (Agresti 2002) เพื่อมุ่งตรวจสอบความสัมพันธ์แบบต่าง ๆ เช่น แบบมีเงื่อนไข (conditional association) แบบมาจินัล (marginal association) ตลอดจนตรวจสอบอิทธิพลของตัวแปรอธิบายหรือตัวแปรร่วม (explanatory or covariate) ที่ส่งผลกระทบต่อตัวแปรตอบสนอง ซึ่งอาจเป็นผลกระทบโดยตรง (direct effect) หรือผลกระทบทางอ้อม (indirect effect) และวิธีการที่สำคัญ ซึ่งเป็นมาตรฐานของการวิเคราะห์ข้อมูลเชิงกลุ่มคือ เทคนิคการสร้างตัวแบบเชิงเส้นที่วางนัยทั่วไป (generalized linear modeling techniques) มาช่วยอธิบายลักษณะความสัมพันธ์แบบต่าง ๆ ระหว่างตัวแปร และสามารถนำตัวแบบซึ่งได้ตรวจสอบภาวะสารูปดีแล้วไปวิเคราะห์เพื่อประเมินลักษณะของอิทธิพลที่มีต่อตัวแปรตอบสนองและการจำแนกกลุ่มต่อไป

สมการเชิงเส้นที่ใช้แสดงความสัมพันธ์ระหว่างตัวแปรตอบสนอง และตัวแปรอธิบายเรียกว่า ตัวแบบเชิงเส้น (Linear Models) ซึ่งตัวแบบเชิงเส้นต่าง ๆ จัดให้อยู่ในเซตของตัวแบบเชิงเส้นที่วางนัยทั่วไป (Generalized linear models) เช่นตัวแบบการวิเคราะห์ความแปรปรวน (ANOVA Models) ตัวแบบโลจิสติก (Logistic models) ฯลฯ ตัวแบบเชิงเส้นที่วางนัยทั่วไป (GLMs) ประกอบด้วยส่วนประกอบ 3 ส่วนคือ **ส่วนประกอบเชิงสุ่ม (random component)** ซึ่งแสดงการแจกแจงความน่าจะเป็นของตัวแปรตอบสนอง **ส่วนประกอบแบบมีระบบ (systematic component)** ซึ่งแสดงฟังก์ชันผลรวมเชิงเส้นของตัวแปรอธิบายที่ใช้เป็นตัวพยากรณ์ และส่วนประกอบ **link function** ซึ่งใช้อธิบายฟังก์ชันความสัมพันธ์ระหว่างส่วนประกอบแบบมีระบบกับส่วนประกอบเชิงสุ่มข้างต้น

ส่วนประกอบที่ 1 ของ GLM คือส่วนประกอบที่เกี่ยวข้องกับสมมติฐานของการแจกแจงของตัวแปรเชิงสุ่ม (Y) ที่เป็นตัวแปรตอบสนอง สมมติว่าค่าสังเกต Y มีขนาด N หน่วยที่เป็นอิสระกัน นั่นคือ $\mathbf{y} = (y_1,\ldots,y_N)$ แต่ละส่วนประกอบของ $\mathbf{y}$ คือ y_i , $i = 1,\ldots,N$ มีการแจกแจงในกลุ่มเอกซ์โปเนนเชียล(exponential dispersion family) ซึ่งอยู่ในรูปของ

$$f(y_i;\theta_i,\phi) = \exp\{[y_i\theta_i-b(\theta_i)]/a(\phi) + c(y_i,\phi)\}.$$

โดยที่ a(.) , b(.) , c(.) แทนฟังก์ชันต่าง ๆ สำหรับพารามิเตอร์

θ เรียกว่า natural parameter

ϕ เรียกว่า dispersion parameter

ส่วนประกอบที่ 2 ของ GLM คือ ส่วนประกอบแบบมีระบบ ทำหน้าที่เชื่อมเวกเตอร์ η โดยที่ $\boldsymbol{\eta} = (\eta_1,\ldots\eta_N)'$ กับเซตของตัวแปรอธิบาย ให้มีรูปแบบเชิงเส้นดังนี้

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} \qquad \text{หรือ} \qquad \eta_i = \sum_j \beta_j x_{ij} \quad i = 1,\ldots, N \;\; j = 1,\ldots, p$$

เมื่อ **X** แทนเมทริกซ์ของตัวแปรอธิบายที่ประกอบด้วยค่าสังเกตขนาด Nxp

$\boldsymbol{\beta}$ แทนเวกเตอร์ของพารามิเตอร์ $(\beta_1,\ldots, \beta_p)'$ ขนาด p

$\boldsymbol{\eta}$ แทนเวกเตอร์ของตัวพยากรณ์เชิงเส้น (linear predictor) ขนาด N

p แทนจำนวนตัวแปรอธิบาย

โดยที่ $\eta_i = g(\mu_i)$

เมื่อ g แทนฟังก์ชันแบบ monotonic differentiable function ของค่าเฉลี่ยค่าสังเกตของตัวแปรตอบสนอง (Y)

μ_i แทนค่าเฉลี่ยของตัวแปรตอบสนองจากค่าสังเกตที่ i

ดังนั้นตัวแบบที่ต้องการจะเชื่อมระหว่างค่าเฉลี่ยค่าสังเกตของ **y** กับตัวแปรอธิบาย คือ

$$g(\mu_i) = \sum_j \beta_j x_{ij} \qquad i = 1,\ldots, N \quad j = 1,\ldots, p$$

ส่วนประกอบที่ 3 ของ GLM คือ link function ต่าง ๆ สำหรับเชื่อมส่วนประกอบเชิงสุ่มและส่วนประกอบแบบมีระบบเข้าด้วยกัน ตัวอย่างของ link function เช่น $\log(m_i)$, $\log[\frac{p(x)}{1-p(x)}]$

เมื่อ m_i แทนค่าคาดหวังของจำนวนนับในแต่ละเซลล์ของตารางการณ์จร

p(x) แทนความน่าจะเป็นของ y เมื่อกำหนด x

ในการวิเคราะห์ข้อมูลโดยใช้ตัวแบบ GLMs ที่มีตัวแปรตอบสนองเชิงกลุ่ม (categorical response) ตัวแปรอธิบาย (explanatory variables) อาจจะเป็นแบบแบ่งกลุ่ม (nominal) หรือแบบมีลำดับ (ordinal) หรือ แบบต่อเนื่อง (continuous) ก็ได้ ตัวแบบ GLMs ต่าง ๆ ถูกนำเสนอสำหรับวิเคราะห์ข้อมูลเชิงกลุ่มแบบมีลำดับ (Agresti 1984) แต่ตัวแบบโลจิทที่วางนัยทั่วไป (generalized logit models) ถูกใช้บ่อยในทางปฏิบัติ เช่น ตัวแบบโลจิททั่วไปสำหรับตัวแปรตามที่มีลำดับรวมทั้งตัวแบบลอกลิเนียร์ (loglinear models) (Haberman 1974:589-600) ตัวแบบคอนทินิวเอชันออดส์ (continuation odds models) (Fienberg and Mason 1979:1-67) และ ตัวแบบ proportional odds models (McCullagh 1980:109-142) สำหรับกรณีพิเศษของตัวแปรตามแบบไบนารี (binary) มีหลายวิธีการที่ใช้ในการประเมินภาวะสารูปดี ได้มีการนำเสนอ

วิธีต่าง ๆ บนพื้นฐานของการแบ่งส่วนข้อมูลของตัวแปรตามออกเป็นกลุ่มย่อย Tsiatis (1980:250-251) และ Hosmer and Lemeshow (1980:1043-1069) ได้เสนอตัวสถิติทดสอบภาวะสารูปดีที่คำนวณในรูปควอดราติกฟอร์มของค่าสังเกตในกลุ่มย่อยต่าง ๆ ส่วน Tsiatis (1980:250-251) ได้เสนอตัวสถิติสำหรับข้อมูลที่มีการแบ่งกลุ่มย่อยตามตัวแปรร่วม แต่ไม่ได้เสนอวิธีการแบ่งข้อมูลภายใต้ตัวแปรร่วมให้อยู่ในกลุ่มย่อยที่เหมาะสม อย่างไรก็ตามนอกจากวิธีสำหรับกรณีพิเศษต่าง ๆ การประเมินภาวะสารูปดีของตัวแบบ GLMs โดยทั่วไปนิยมใช้ตัวสถิติอัตราส่วนภาวะน่าจะเป็น (likelihood ratio deviance statistic) ซึ่งเป็นตัวสถิติที่มีความคล่องตัวและเป็นวิธีวัดที่ใช้ได้ในหลายสถานการณ์ (global measures of goodness-of-fit statistic) สำหรับทดสอบตัวแบบ GLMs แบบต่าง ๆ ดังกล่าวข้างต้น

Hosmer and Lemeshow (1980:1043-1069) ได้เสนอการแบ่งข้อมูลเป็นกลุ่มย่อยหรือขอบเขตโดยแบ่งตามเปอร์เซ็นต์ไทล์ของตัวประมาณความน่าจะเป็นที่ได้จากการสร้างตัวแบบ Logistic Regression models แต่เทคนิคนี้ไม่ได้ขยายไปสู่ตัวแปรไม่ต่อเนื่องที่มีมากกว่า 2 ระดับ

Lipsitz, Fitzmaurice and Molenberghs (1996:175-190)ได้นำเสนอการขยายขอบเขตของวิธีดังกล่าวสำหรับตัวแปรตอบสนองแบบมีลำดับ 3 ระดับ ซึ่งตัวแบบที่ใช้คือตัวแบบ **proportional odds model** (McFadden 1974:105-142) คือ

$$L_{ik} = L_{ik}(p_{ij}) = \alpha_k + \mathbf{x}_i'\boldsymbol{\beta} \quad \text{......} \quad (1)$$

เมื่อ

$$L_{ik} = \log\left(\frac{p_{i1} + \ldots + p_{ik}}{p_{i,k+1} + \ldots + p_{iK}}\right) \qquad i = 1,\ldots,N \quad k = 1,\ldots,K\text{-}1$$

$\mathbf{x}_i$ แทนเวกเตอร์ขนาด p ของค่าสังเกตของตัวแปรอธิบาย

$\boldsymbol{\beta}$ แทนเวกเตอร์ของพารามิเตอร์ $(\beta_1,\ldots,\ \beta_p)'$ ขนาด p

$p_{ik} = p(y_i = k|\mathbf{x}_i)$

K แทนจำนวนกลุ่มหรือระดับของค่าสังเกต y_i

ตัวแบบนี้ (ในกรณีที่ ตัวแปรตอบสนองเป็นแบบมีลำดับ) น่าจะให้ผลสารสนเทศเหมาะสมกว่าการใช้ตัวแบบ **baseline-category logit models** (Fadden 1974) คือ

$$\log\frac{p_{ik}(\mathbf{x})}{p_{iK}(\mathbf{x})} = \alpha_k + \mathbf{x}_i'\boldsymbol{\beta}_k \quad k = 1\,,\,.\,.\,.\,,K\text{-}1 \ \text{......} \quad (2)$$

เมื่อ $\mathbf{x}_i$ แทนเวกเตอร์ขนาด p ของค่าสังเกตที่ i

$\boldsymbol{\beta}_k$ แทนเวกเตอร์ขนาด p ของพารามิเตอร์ที่สอดคล้องกับโลจิทที่ k ของตัวแปรตอบสนอง Y

โดยทั่วไปใช้ในกรณีที่ตัวแปรตอบสนองเป็นแบบไม่มีลำดับ (nominal categorical variable)

การเลือกตัวแบบ GLMs ตัวแบบหนึ่งตัวแบบใด นอกจากศึกษาความสัมพันธ์ระหว่างตัวแปรแล้วเรายังใช้ตัวแบบมาพยากรณ์ตัวแปรตอบสนองเชิงกลุ่มด้วย โดยที่ตัวแบบที่ใช้วิเคราะห์ความสัมพันธ์และพยากรณ์นั้นควรมีภาวะสารูปดี การทดสอบภาวะสารูปดีสำหรับตัวแบบเชิงเส้นของตัวแปรแบบมีลำดับที่ใช้พยากรณ์ตัวแปรตอบสนองเชิงกลุ่มนั้นสามารถใช้ likelihood ratio deviance และ ตัวสถิติเพียร์สันไคสแควร์ โดยจะวัดภาวะสารูปดีของตัวแบบเทียบกับตัวแบบเต็ม การแจกแจงแบบไคสแควร์จะเป็นการแจกแจงโดยประมาณที่ดีของการแจกแจงของตัวสถิติภาวะสารูปดีเมื่อข้อมูลส่วนใหญ่มีความถี่ในแต่ละเซลล์มากกว่า 5 ตัวแบบ **proportional odds models** ตั้งชื่อโดย McCullagh (1980:109-142) ซึ่งได้ใช้ Fisher scoring algorithms ในการคำนวณค่าประมาณแบบภาวะน่าจะเป็นสูงสุด (MLEs) ของพารามิเตอร์ของตัวแบบ อย่างไรก็ตาม ตัวสถิติที่ใช้ในการทดสอบความเหมาะสมของตัวแบบดังกล่าวอาจไม่ลู่เข้าสู่การแจกแจงแบบไคสแควร์เมื่อตัวแปรอธิบายเป็นตัวแปรต่อเนื่อง Lipsitz et al. (1996:175-190) ได้นำเสนอตัวแบบที่ใช้ในการทดสอบภาวะสารูปดีของตัวแบบ **proportional odds models** ที่มีตัวแปรตอบสนองแบบมีลำดับตั้งแต่ 3 ระดับขึ้นไป โดยพัฒนามาจากตัวสถิติของ Hosmer and Lemeshow (1980:1043-1069) ซึ่งเป็นตัวสถิติที่ใช้ทดสอบภาวะสารูปดีของตัวแบบโลจิสติกที่มีตัวแปรอธิบายแบบต่อเนื่องและตัวแปรตอบสนองเป็นตัวแปรแบบแบ่งกลุ่มที่มี 2 ระดับ โดยอาศัยการแบ่งกลุ่มของข้อมูลออกเป็นกลุ่มย่อยตามค่าประมาณค่าเฉลี่ยของตัวแปรตอบสนอง แต่ Lipsitz et al. (1996:175-190) ยังไม่ศึกษาการทดสอบภาวะสารูปดีโดยอาศัยการแบ่งกลุ่มข้อมูลดังกล่าวกับตัวแบบ **baseline categories logit** ที่มีตัวแปรตอบสนองเชิงกลุ่มที่มีระดับตั้งแต่ 3 ระดับขึ้นไป

งานวิจัยนี้สนใจศึกษาตัวแบบ **proportional odds models** ซึ่งเป็นตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับที่มีส่วนประกอบเชิงสุ่มแบบ ordinal multinomial มากกว่า 2 กลุ่มหรือระดับ ส่วนประกอบแบบมีระบบใช้ตัวแปรอธิบายแบบผสม และส่วนประกอบ Link function ใช้ Cumulative probabilities Link function โดยจะศึกษาเปรียบเทียบกับตัวแบบ **baseline-category logit mode**l ซึ่งใช้ตัวแปรตอบสนองเป็นตัวแปรที่ไม่มีลำดับและแตกต่างจากตัวแบบข้างต้นที่ส่วนประกอบเชิงสุ่มเป็นแบบ nominal multinomial และ Link function เป็นแบบ Beseline-logit Link function และต้องการเปรียบเทียบพลังการทดสอบ(%)ระหว่างการใช้ตัวแปรตอบสนองแบบมีลำดับและไม่มีลำดับ ในเวลาเดียวกันยังสนใจการวิเคราะห์พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด และตัวสถิติวาลด์ ซึ่งเป็นที่นิยมสำหรับการตรวจสอบภาวะสารูปดีของตัวแบบ GLMs(Clogg and Shihadeh 1994) โดยเน้นการใช้ตัวแบบต่าง ๆ เช่น

$$L_{ik} = \alpha_k + \mathbf{x}_i'\boldsymbol{\beta} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2,\ldots, K-1 \qquad g = 1,\ldots, G-1 \;\ldots. \qquad (3)$$

ของ และ Lipsitz et al. (1996:175-190) เป็นตัวแบบทางเลือก

เมื่อ $\mathbf{x}_i$ แทนเวกเตอร์ขนาด p ของค่าสังเกตที่ i

$\boldsymbol{\beta}$ แทนเวกเตอร์ขนาด p ของพารามิเตอร์ที่สอดคล้องกับโลจิทของตัวแปรตอบสนอง Y

I_{ig} แทนอินดิเคเตอร์ (ในที่นี้ให้เป็นตัวแปรดัมมี) ของค่าสังเกตที่ i ซึ่งได้จากการแบ่งค่าสังเกตออกเป็นกลุ่มย่อย G กลุ่ม ตามลำดับของค่าพยากรณ์ค่าเฉลี่ยของตัวแปรตาม

γ_g แทนพารามิเตอร์ที่สอดคล้องกับตัวแปรดัมมีของกลุ่มย่อย g ของค่าสังเกต

วัตถุประสงค์ของการวิจัย

1. ศึกษาพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ภายใต้ตัวแบบต่าง ๆ ที่ใช้ในการทดสอบภาวะสารูปดี
2. เพื่อเปรียบเทียบพลังการทดสอบ(%)ภายใต้ตัวแบบที่ใช้สำหรับตัวแปรตอบสนองแบบมีลำดับ และไม่มีลำดับเมื่อตัวแปรตอบสนองเป็นแบบมีลำดับโดยใช้ตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด และตัวสถิติวาลด์ ภายใต้ตัวแบบเชิงเส้นที่วางนัยทั่วไปแบบต่าง ๆ และขนาดตัวอย่างต่าง ๆ กัน

ขอบเขตของการวิจัย

ขอบเขตของงานวิจัยนี้คือ

1. สร้างแบบจำลองเพื่อศึกษาพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์สำหรับตัวแบบตามสมมติฐานว่างคือตัวแบบ(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad k = 1,2 \qquad i = 1,2\ldots, n \;\ldots\ldots\ldots\ldots \qquad (4)$$

และเปรียบเทียบตัวแบบทางเลือกระหว่างตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2 \qquad i = 1,2,\ldots, n \;\ldots\ldots \qquad (5)$$

และตัวแบบ(6) **pairwise**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad k = 1,2 \qquad i = 1,2,\ldots n \;\ldots\ldots \qquad (6)$$

เมื่อ X_1 มีการแจกแจงแบบ Bernoulli(0.5)

X_2 มีการแจกแจงแบบ Normal(0,1)

ตัวแปรตอบสนองเป็นตัวแปรแบบมีลำดับ 3 ระดับ (K = 3)

กำหนด G = 10

$$L_{ik} = \left(\frac{p_{i1} + \ldots + p_{ik}}{p_{i,k+1} + \ldots + p_{i3}} \right) \qquad k = 1,2$$

$$I_{ig} = \begin{cases} 1 & \text{if } \hat{\mu}_i \text{ is in region } g, \\ 0 & \text{otherwise}, \end{cases} \quad i = 1,.., n \quad g = 1,.., 9$$

เมื่อ $\hat{\mu}_i = \hat{p}_{i1} + 2\hat{p}_{i2} + 3\hat{p}_{i3}$ แทนค่าเฉลี่ยของตัวแปรตอบสนองของค่าสังเกตที่ i กำหนดระดับนัยสำคัญในการทดสอบ 0.05

1.1 กำหนดตัวแบบที่ถูกต้องคือ(6)

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad , k = 1,2 \qquad i = 1,2,\ldots n$$

กำหนดเซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ดังนี้ { -0.6, 0.2, 0.707, -0.707}
ในการศึกษาพลังการทดสอบ(%)ของตัวสถิติกำหนดค่า β_{12} ให้มีค่าเปลี่ยนแปลงจาก 0-1.4 โดยเพิ่มขึ้นครั้งละ 0.1

1.2 กำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ(7) **quadratic**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2 \qquad k = 1,2 \quad \ldots\ldots\ldots\ldots \quad (7)$$

โดยใช้เซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ เช่นเดียวกับข้อ 1.1
ในการศึกษาพลังการทดสอบ(%)ของตัวสถิติกำหนดค่า β_{22} ให้มีค่าเปลี่ยนแปลงจาก 0-1.4 โดยเพิ่มขึ้นครั้งละ 0.1

2 สร้างแบบจำลองเพื่อเปรียบเทียบพลังการทดสอบ(%)เมื่อใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับ และตัวแบบสำหรับตัวแปรตอบสนองแบบไม่มีลำดับ โดยใช้ตัวสถิติอัตราส่วนภาวะน่าจะเป็นตัวสถิติวาลด์

กำหนดตัวแบบตามสมติฐานว่าง เป็น(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad k = 1,2 \qquad i = 1,2\ldots, n$$

และตัวแบบตามสมมติฐานทางเลือกเป็นตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + x_1\beta_1 + x_2\beta_2 + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2 \qquad i = 1,\ldots, n$$

เปรียบเทียบกับกรณีที่ใช้ตัวแบบตามสมมติฐานว่างเป็นตัวแบบ(8) **baseline-category logit**

$$\log \frac{p_{ik}}{p_{i3}} = \alpha_k + \beta_{1k} x_{1i} + \beta_{2k} x_{2i} \qquad k = 1,2 \qquad i = 1,\ldots, n \quad \ldots\ldots\ldots \quad (8)$$

เมื่อ $p_{ik} = p(y_i = k | \mathbf{x}_i)$

$\mathbf{x}_i$ แทนเวกเตอร์ค่าสังเกตของตัวแปรอธิบาย ขนาด p

และตัวแบบตามสมมติฐานทางเลือกเป็นตัวแบบ(9) **baseline-partition**

$$\log \frac{p_k(\mathbf{x})}{p_3(\mathbf{x})} = \alpha_k + \beta_{1k} x_{1i} + \beta_{2k} x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_{gk} \qquad k = 1,2 \qquad i = 1,\ldots n \quad \ldots\ldots\ldots \quad (9)$$

ตามลำดับ

2.1 กำหนดตัวแบบที่ถูกต้องคือ(6) **pairwise**

$$L_{ik} = \alpha_k + x_{1i}\beta_1 + x_{2i}\beta_2 + \beta_{12}x_{1i}x_{2i} \qquad k = 1,2 \qquad i = 1,\dots, n$$

เมื่อ X_1 มีการแจกแจงแบบ Bernoulli(0.5)

X_2 มีการแจกแจงแบบ Normal(0,1)

กำหนดระดับนัยสำคัญในการทดสอบ 0.05

กำหนดเซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ดังนี้ {- 0.6, 0.2, 0.707, -0.707}

ในการศึกษาพลังการทดสอบ(%)ของตัวสถิติกำหนดค่า β_{12} ให้มีค่าเปลี่ยนแปลงจาก 0-1.4 โดยเพิ่มขึ้นครั้งละ 0.1

ในแต่ละเซตเงื่อนไขของการจำลองแบบจะทำการศึกษาที่ขนาดตัวอย่าง 600, 1000 และ 1500

2.2 กำหนดตัวแบบที่ถูกต้องคือ(7) **quadratic**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2 \qquad k = 1,2 \qquad i = 1,\dots, n$$

เมื่อ X_1 มีการแจกแจงแบบ Bernoulli(0.5)

X_2 มีการแจกแจงแบบ Normal(0,1)

กำหนดระดับนัยสำคัญในการทดสอบ 0.05

กำหนดเซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ดังนี้ {- 0.6, 0.2, 0.707, -0.707}

ในการศึกษาพลังการทดสอบ(%)ของตัวสถิติกำหนดค่า β_{22} ให้มีค่าเปลี่ยนแปลงจาก 0-1.4 โดยเพิ่มขึ้นครั้งละ 0.1

ในแต่ละแต่ละเซตเงื่อนไขของการจำลองแบบจะทำการศึกษาที่ขนาดตัวอย่าง 600, 1000 และ 1500 ทำการจำลองแบบซ้ำ 500 ครั้ง โดยใช้ข้อมูลของตัวแปรอธิบายต่าง ๆ ชุดเดิม(fixed) และใช้ข้อมูลของตัวแปรตอบสนองเชิงสุ่ม (random) ภายใต้ตัวแบบที่ถูกต้องหรือตัวแบบที่ถูกต้องของการวิจัย การประมวลผลลัพธ์ของการวิจัยใช้ macro program ที่เขียนขึ้นมาประมวลผลร่วมกับ Minitab version 11.

นิยามศัพท์

ตัวสถิติวาลด์ (Wald statistic)

ในการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ของตัวแบบต่าง ๆ (ตัวแปรตอบสนองเป็นตัวแปรไม่ต่อเนื่อง)

$$H_0; \boldsymbol{\beta} = \overline{\mathbf{o}}$$
$$H_1; \boldsymbol{\beta} \neq \overline{\mathbf{o}}$$

สามารถใช้ตัวสถิติทดสอบตัวใดตัวหนึ่ง ดังต่อไปนี้

Wald Statistic $W = \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right)' \left[\mathrm{cov}(\hat{\boldsymbol{\beta}})\right]^{-1} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right)$

เมื่อ $\hat{\boldsymbol{\beta}}$ แทนเวกเตอร์ค่าประมาณของเวกเตอร์พารามิเตอร์ $\boldsymbol{\beta}$

ตัวสถิติอัตราส่วนภาวะน่าจะเป็น (Likelihood ratio statistic)

Likelihood Ratio Statistic $D = -2\log\Lambda = -2\log(\lambda_0 / \lambda_1) = -2(D_0 - D_1)$

เมื่อ λ_0 และ λ_1 แทนภาวะน่าจะเป็นภายใต้สมมติฐานว่างและภาวะน่าจะเป็นภายใต้สมมติฐานทางเลือก ตามลำดับ

D_0 และ D_1 แทน ค่าสูงสุดของ $\log(\lambda_0)$ และ $\log(\lambda_1)$ ตามลำดับ

ตัวสถิติที่กล่าวมานี้มีการแจกแจงเมื่อใกล้อนันต์เป็นแบบไคสแควร์ (χ^2)

พลังการทดสอบ (Power of the test) หมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง (H_0) เมื่อสมมติฐานว่าง(H_0) ไม่จริง ซึ่งในงานวิจัยนี้เรียกสั้น ๆ ว่า power

ตัวแบบ (Model)

ตัวแบบ หมายถึง สมการแสดงความสัมพันธ์ระหว่างตัวแปรตอบสนองกับตัวแปรอธิบายซึ่งอยู่ในรูปของโครงสร้างตัวแบบเชิงสถิติ (statistical model) หรือของตัวแบบเชิงเส้นที่วางนัยทั่วไป (generalized linear models)

ชื่อของตัวแบบในงานวิจัยนี้

ตัวแบบ(1) **proportional odds models**

$$L_{ik} = L_{ik}(p_{ij}) = \alpha_k + \mathbf{x}_i'\boldsymbol{\beta} \quad k = 1, \ldots, K\text{-}1 \quad i = 1, \ldots, n$$

เป็นตัวแบบที่อยู่ในรูปทั่วไปของตัวแบบในกลุ่ม **proportional odds models**

ตัวแบบ(2) **baseline-category logit models**

$$\log\frac{p_{ik}(\mathbf{x})}{p_{iK}(\mathbf{x})} = \alpha_k + \mathbf{x}_i'\boldsymbol{\beta}_k \quad k = 1, \ldots, K\text{-}1 \quad i = 1, \ldots, n$$

เป็นตัวแบบที่อยู่ในรูปทั่วไปของตัวแบบในกลุ่ม **baseline-category logit models**

ตัวแบบ(3) **proportional odds-partition**

$$L_{ik} = \alpha_k + \mathbf{x}_i'\boldsymbol{\beta} + \sum_{g=1}^{G-1} I_{ig}\gamma_g, \; k = 1,2,\ldots,\; K-1,\; g = 1,\ldots,\; G-1$$

เป็นตัวแบบที่อยู่ในรูปทั่วไปของตัวแบบ **proportional odds models** ที่เพิ่มตัวแปรดัมมี (I_{ig}) ที่ได้จากการแบ่งกลุ่มข้อมูลตามค่าประมาณค่าเฉลี่ยของตัวแปรตอบสนอง

ตัวแบบ(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad , k = 1,2 \qquad i = 1,2\ldots,\; n$$

เป็นตัวแบบที่อยู่ในกลุ่ม **proportional odds models** ซึ่งเป็นตัวแบบที่ไม่ซับซ้อน

ตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2 \qquad i = 1,2,...,\ n$$

เป็นตัวแบบในกลุ่ม **proportional odds models** ที่เพิ่มตัวแปรดัมมีที่ได้จากการแบ่งกลุ่มข้อมูลตามค่าประมาณค่าเฉลี่ยของตัวแปรตอบสนอง

ตัวแบบ(6) **pairwise**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad k = 1,2 \qquad i = 1,2,...n$$

เป็นตัวแบบที่อยู่ในกลุ่ม **proportional odds models** ที่เพิ่มเทอมผลคูณของตัวแปรอธิบาย X_1 และ X_2

ตัวแบบ(7) **quadratic**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2 \qquad k = 1,2$$

เป็นตัวแบบที่อยู่ในกลุ่ม **proportional odds models** ที่เพิ่มเทอมกำลังสองของตัวแปร X_2

ตัวแบบ(8) **baseline-category logit**

$$\log \frac{p_{ik}}{p_{i3}} = \alpha_k + \beta_{1k} x_{1i} + \beta_{2k} x_{2i} \qquad k = 1,2 \qquad i = 1,...,\ n$$

เป็นตัวแบบที่อยู่ในกลุ่ม **baseline-category logit models** ซึ่งเป็นตัวแบบที่ไม่ซับซ้อน

ตัวแบบ(9) **baseline-partition**

$$\log \frac{p_k(\mathbf{x})}{p_3(\mathbf{x})} = \alpha_k + \beta_{1k} x_{1i} + \beta_{2k} x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_{gk} \qquad k = 1,2 \qquad i = 1,...n$$

เป็นตัวแบบที่อยู่ในกลุ่ม **baseline-category logit mode**ls ที่เพิ่มตัวแปรดัมมีที่เกิดจากแบ่งกลุ่มข้อมูลตามค่าประมาณค่าเฉลี่ยของตัวแปรตอบสนอง

ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถเปรียบเทียบพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นซึ่งเป็นตัวสถิติที่นิยมในอดีตถึงปัจจุบันเทียบกับตัวสถิติวาลด์ ภายใต้ตัวแบบเชิงเส้นที่วางนัยทั่วไป หลากหลายตัวแบบ
2. เห็นความแตกต่างของพลังการทดสอบ(%)ของการใช้ตัวสถิติภายใต้ตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับ กับการใช้ตัวแบบสำหรับตัวแปรตอบสนองที่ไม่มีลำดับ โดยที่จริง ๆ แล้วตัวแปรตอบสนองที่ต้องการศึกษาเป็นตัวแปรแบบมีลำดับ
3. เห็นพฤติกรรมของตัวแบบทางเลือกที่น่าสนใจในการทดสอบภาวะสารูปดีของตัวแบบภายใต้สมมติฐานว่าง

บทที่ 2

ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ตัวแบบโลจิทสำหรับตัวแปรตอบสนองแบบมัลติโนเมียล (Logit Model for Multinomial Responses)

ในกรณีที่ตัวแปรตอบสนองมีผลลัพธ์ที่เป็นไปได้มากกว่า 2 กลุ่มหรือระดับ (polychromous) ตัวแบบที่ใช้ในการวิเคราะห์ความสัมพันธ์ระหว่าง ตัวแปรตอบสนองกับตัวแปรอธิบายได้รับการพัฒนามาจากตัวแบบการถดถอยลอจิสติกและอาจเรียกรวม ๆ ว่า **logit model for multinomial responses** (Agresti 2002) สำหรับกรณีที่ตัวแปรตอบสนองที่มีผลลัพธ์ที่เป็นไปได้มากกว่า 2 กลุ่มหรือระดับแบบแบ่งกลุ่ม (nominal) เป็นที่รู้จักดี คือตัวแบบ baseline-category logit model (McFadden 1974:105-142)

ตัวแบบโลจิทสำหรับตัวแปรตอบสนองแบบไม่มีลำดับ (Nominal Response: Baseline-Category Logit Model)

สำหรับกรณีที่ตัวแปรตอบสนองที่มี K กลุ่มหรือระดับอย่างไม่มีลำดับ (normimal categories or levels) ตัวแบบที่ใช้จะสอดคล้องกับออส์ดของระดับที่สนใจเทียบกับมาตรฐาน (baseline-category) จำนวน K-1 คู่

ให้ $p_k(\mathbf{x})$ = P(Y=k|**x**) เมื่อ **x** แทนเมทริกซ์ของตัวแปรอธิบายขนาด nxp และ $\sum_k p_k(\mathbf{x}) = 1$ สำหรับตัวแปรตอบสนองที่มี K ระดับจะมีการแจกแจงแบบมัลติโนเมียลด้วยความน่าจะเป็น $\{p_1(\mathbf{x}), \ldots, p_K(\mathbf{x})\}$ ตัวแบบโลจิทสำหรับแต่ละคู่ระหว่างตัวแปรตอบสนองระดับที่ k กับระดับฐาน(ตัวแปรตอบสนองระดับที่ K) คือ

$$\log \frac{p_k(\mathbf{x})}{p_K(\mathbf{x})} = \alpha_k + \beta_k' \mathbf{x} \qquad , k = 1,2,\ldots, K-1$$

ค่าประมาณของความน่าจะเป็นของตัวแปรตอบสนองระดับที่ k คือ

$$p_k(\mathbf{x}) = \frac{\exp(\alpha_k + \beta_k' \mathbf{x})}{1 + \sum_{k=1}^{K-1} \exp(\alpha_k + \beta_k' \mathbf{x})}$$

สำหรับกรณี ที่ตัวแปรตอบสนองที่มีผลลัพธ์ที่เป็นไปได้มากกว่า 2 กลุ่มหรือระดับ แบบมีลำดับ (ordinal categories or levels) อาจเรียกรวม ๆ ว่า **multinomial logit-models for ordinal responses** (Agresti 1984)

ตัวแบบมัลติโนเมียลโลจิทสำหรับตัวแปรตอบสนองแบบมีลำดับ(Ordinal Responses: Multinomial Logit Models)

เนื่องจากตัวแปรตอบสนองเป็นตัวแปรเชิงกลุ่ม ให้ Y_{ik} แทนค่าสังเกตที่ i (i = 1 ,2, . . ., n) เป็นตัวแปรตอบสนองกลุ่มหรือระดับที่ k (k = 1 , 2, . . ., K) $Y_{ik} = 1$ เมื่อค่าสังเกตนั้นเป็นตัวแปรตามกลุ่มหรือระดับที่ k และเท่ากับ 0 ที่อื่น ๆ ดังนั้น $\sum_{k=1}^{K} Y_{ik} = 1$ เราสามารถจัด Y_{ik} ให้อยู่ในรูปเวกเตอร์ $\mathbf{Y_i} = (Y_{i1}, \ldots, Y_{ik})'$ สัดส่วนของตัวแปรตอบสนองที่ k คือ $p_{ik} = E(Y_{ik})$, $\sum_{k=1}^{K} p_{ik} = 1$ เวกเตอร์สุ่ม $\mathbf{Y_i}$ มีการแจกแจงแบบมัลติโนเมียลด้วยเวกเตอร์สัดส่วนความน่าจะเป็น $\mathbf{p_i} = (p_{i1}, \ldots, p_{iK})'$ ค่าสังเกตแต่ละค่ามีเวกเตอร์ของตัวแปรอธิบายเป็น $\mathbf{x_i} = (x_{i1}, \ldots, x_{ip})'$

มีลิงค์ฟังก์ชันต่าง ๆ ที่เป็นทางเลือกหลายแบบซึ่งอยู่ในรูปตัวแบบทั่วไปดังนี้

$$L_{ik} = L_{ik}(p_i) = \alpha_k + x_i'\beta$$

สำหรับ k = 1, . . .,K-1 ตัวแบบ **cumulative logit models** คือ

$$L_{ik} = \log\left(\frac{p_{i1} + \ldots + p_{ik}}{p_{i,k+1} + \ldots + p_{iK}}\right)$$

ตัวแบบ **continuation-ratio logit models** คือ

$$L_{ik} = \log\left(\frac{p_{i,k+1}}{p_{i1} + \ldots + p_{ik}}\right)$$

และ ตัวแบบ **adjacent-categories logit models** คือ

$$L_{ik} = \log(p_{i,k+1} / p_{ik})$$

นอกจากนี้ก็ยังมีลิงค์ฟังก์ชันง่าย ๆ ที่เราสามารถเลือกใช้ได้ เช่น cumulative probit

$$L_{ik} = \Phi^{-1}(p_{i,k+1} + \ldots + p_{iK})$$

เมื่อ $\Phi(\,)$ คือ ฟังก์ชันการแจกแจงสะสมแบบปกติ และตัวแบบ **cumulative complementary-log-log logit models**

$$L_{ik} = \log\{-\log(p_{i,k+1} + \ldots + p_{iK})\}$$

ซึ่งเป็นตัวแบบที่มีความนิยมตัวแบบหนึ่ง สำหรับทุกตัวแบบดังกล่าว $p_{\mathbf{i}} = p_{\mathbf{i}}(\beta)$ เมื่อ $\beta = (\alpha', \beta')$

ในส่วนถัดไปของงานวิจัยนี้ผู้วิจัยได้นำเสนอการทดสอบภาวะสารูปดีที่เหมาะสำหรับตัวแบบที่กล่าวมาข้างต้น ตามลำดับดังนี้

การประเมินตัวแบบ (Assessing the Fit of the Model)

ให้ $\mathbf{y}' = (y_1, \ldots y_n)$ แทนเวกเตอร์ผลลัพธ์ของตัวแปรตอบสนอง

$\hat{\mathbf{y}}' = (\hat{y}_1, \ldots \hat{y}_n)$ แทนเวกเตอร์ค่าพยากรณ์ของตัวแปรตอบสนอง

เราจะสรุปว่าตัวแบบเหมาะสมกับข้อมูลเมื่อ 1.) ผลรวมของระยะห่างระหว่าง $\mathbf{y}$ กับ $\hat{\mathbf{y}}$ มีค่าน้อย 2.) คู่ลำดับ $(y_i, \hat{y}_i)$, i = 1, . . ., n มีรูปแบบไม่เป็นระบบ (unsystematic) ของความสัมพันธ์

กับส่วนเหลือ หรือมีรูปแบบบ้างก็น้อย (Hosmer and Lemeshow 1980:1043-1069) ดังนั้นการประเมินความเหมาะสมของตัวแบบจึงเกี่ยวข้องกับการหาระยะห่างระหว่าง y_i กับ $\hat{y}_i$ และทำการตรวจสอบองค์ประกอบของระยะห่างดังกล่าวอย่างถี่ถ้วน

ค่าวัดความเหมาะสมของตัวแบบถูกคำนวณหลังจากที่เราสร้างตัวแบบ (model fitting) แต่เนื่องจากค่าวัดเหล่านี้ไม่ได้แสดงเฉพาะแต่ละค่าสังเกต (ค่าวัดของตัวสถิติต่าง ๆ จะวัดออกมาเป็นผลรวมของความคลาดเคลื่อนของทุกค่าสังเกต) ดังนั้นถ้าค่าวัดความเหมาะสมของตัวแบบดังกล่าวมีค่าน้อย เราก็ไม่สามารถมองเห็นถึงข้อมูลส่วนน้อย(ถ้ามี) ที่เบี่ยงเบนไปจากค่าพยากรณ์ และค่าวัดความเหมาะสมของตัวแบบถ้ามีค่ามากนั้นจะบ่งบอกถึงความไม่เหมาะสมของตัวแบบ ก่อนที่จะหาตัวสถิติสำหรับประเมินความเหมาะสมของตัวแบบต้องพิจารณาสิ่งที่มีอิทธิพลต่อค่าวัดดังกล่าวก่อน นั่นคือองศาอิสระ ซึ่งขึ้นอยู่กับจำนวน covariate pattern (จำนวนกลุ่มของตัวแปรร่วมหรือตัวแปรอธิบายที่แตกต่างกัน)เขียนแทนด้วย J ถ้าในข้อมูล n ค่าสังเกตไม่มีเซตของตัวแปรอธิบายแต่ละค่าสังเกตซ้ำกันจะทำให้ J = n ตัวอย่างเช่นถ้าข้อมูลมีตัวแปรอธิบายเป็น อายุ เชื้อชาติ เพศ และน้ำหนัก จะเห็นได้ว่าเซตของตัวแปรร่วมจะมีความหลากหลายมากดังนั้นแต่ละเซตของตัวแปรอธิบายจะมีผลลัพธ์ของตัวแปรตอบสนองค่าเดียว แต่ถ้าใช้เพียงตัวแปรอธิบาย เชื้อชาติ และเพศ โดยกำหนดให้ตัวแปรแต่ละกลุ่มหรือระดับมี 2 ระดับ ก็จะได้เซตของตัวแปรร่วมที่แตกต่างกัน 4 แบบ เช่น(00 01 10 11) นั่นคือ จำนวน covariate pattern เท่ากับ 4 ดังนั้นถ้าเก็บข้อมูลขนาด n > 4 ในแต่ละเซตของตัวแปรอธิบายจะมีผลลัพธ์ของตัวแปรตอบสนองมากกว่าหนึ่งค่า

ในกระบวนการเลือกตัวแบบที่เหมาะสมไม่จำเป็นต้องหาจำนวน covariate pattern เนื่องจากองศาอิสระ ที่จำเป็นต่อการหา p-value สามารถหาได้จากผลต่างของจำนวนตัวแปรในแต่ละตัวแบบ แต่ในการประเมินความเหมาะสมของตัวแบบใดตัวแบบหนึ่ง จำเป็นจะต้องทราบจำนวน covariate pattern เนื่องจากจำนวนดังกล่าวมีผลโดยตรงกับองศาอิสระ

สมมติในตัวแบบที่ต้องการประเมินความเหมาะสมมีตัวแปรร่วม p ตัวแปร $\mathbf{x}' = (x_1, \ldots x_p)$ ถ้าในข้อมูลมีเวกเตอร์ของตัวแปรร่วมเหมือนกันอย่างน้อยหนึ่งคู่จะทำให้ J < n จำนวนซ้ำของเวกเตอร์ในแต่ละแบบที่แตกต่างกันเขียนแทนด้วย $m_j, j = 1,2,3,\ldots n$ และ $\sum_j m_j = n$ ให้ y_j แทนจำนวนผลลัพธ์ที่สนใจ (y = 1) จะได้ $\sum y_j = n_1$ แทนผลรวมของจำนวนผลลัพธ์ที่สนใจทั้งหมด การแจกแจงของตัวสถิติที่ใช้ทดสอบความเหมาะสมของตัวแบบหาได้เมื่อ n มีขนาดใหญ่มากพอ ซึ่งเรียกว่า n-asymptotics แต่กรณีที่ J < n เมื่อกำหนดให้ J คงที่ และ n มีขนาดใหญ่จะทำให้ m_j มีแนวโน้มทีจะมีค่าใหญ่ด้วย กรณีนี้การแจกแจงของตัวสถิติจะหาได้เมื่อ m_j มีขนาดใหญ่ เรียกว่า m-asymptotics

ตัวสถิติไคสแควร์พียร์สันและดีเวียนส์ (Pearson Chi-Square and Deviance)

ในตัวแบบถดถอยเชิงเส้น ผลรวมระยะห่างระหว่างค่าสังเกตกับค่าพยากรณ์และการวิเคราะห์อิทธิพลที่มีผลต่อค่าพยากรณ์ ค่าวัดที่ได้จะเป็นฟังก์ชันของส่วนเหลือ หรือผลต่างระหว่างค่าสังเกตและค่าพยากรณ์ของตัวแปรตาม $(\mathbf{y} - \hat{\mathbf{y}})$ ในการถดถอยโลจิสติกสามารถวัดผลต่างระหว่างค่าสังเกตกับค่าพยากรณ์ของตัวแปรตอบสนองได้ในทำนองเดียวกันและยังทำได้หลายวิธี ค่าพยากรณ์ของตัวแปรตอบสนองขึ้นกับค่าประมาณความน่าจะเป็นของผลลัพธ์ระดับต่าง ๆ ของตัวแปรตอบสนองที่ได้จากแต่ละ covariate pattern ดังนั้นค่าพยากรณ์ $\hat{y}_i$ คำนวนได้จาก

$$m_j\hat{p}_j = m_j \exp\left[\sum_j \hat{\beta}_j x_{ij}\right] / \left\{1 + \exp\left[\sum_j \hat{\beta}_j x_{ij}\right]\right\}$$

ส่วนเหลือเพียร์สันนิยามโดย

$$r(y_j, \hat{p}_j) = \frac{(y_j - m_j\hat{p}_j)}{\sqrt{m_j\hat{p}_j(1-\hat{p}_j)}}$$

ตัวสถิติไคสแคร์เพียร์สันคือ

$$X^2 = \sum_{j=1}^{J} r(y_j, \hat{p}_j)^2$$

Deviance Residual นิยามโดย

$$d(y_j, \hat{p}_j) = \pm\left\{2\left[y_j \ln\left(\frac{y_j}{m_j\hat{p}_j}\right) + (m_j - y_j)\ln\left(\frac{m_j - y_j}{m_j(1-\hat{p}_j)}\right)\right]\right\}^{1/2}$$

ตัวสถิติ Deviance คือ

$$D = \sum_{j=1}^{J} d(y_j, \hat{p}_j)^2$$

การแจกแจงของตัวสถิติ X^2 และ D ภายใต้สมมติฐานตัวแบบถูกต้อง คือการแจกแจงแบบไคสแควร์ องศาอิสระเท่ากับ J-(p+1) สำหรับตัวสถิติ D สอดคล้องกับตัวสถิติอัตราส่วนภาวะน่าจะเป็น (LR) ของตัวแบบที่มี J พารามิเตอร์ (LR_1) เทียบกับตัวแบบที่มี p+1 พารามิเตอร์ (LR_2)

โดยที่ $LR = LR_1 - LR_2 = -2(L_1 - L_p) - (-2)(L_2 - L_p) = -2D_1 - (-2D_2) = -2(D_1 - D_2)$

เมื่อ L_p , L_1 และ L_2 แทน maximize log-likelihood ของตัวแบบเต็ม (saturated model) และของตัวแบบลดรูป (reduced models) M_2 และ M_1 ตามลำดับ และ M_2 มีจำนวนพารามิเตอร์มากกว่า M_1

ข้อเสียของตัวสถิติเหล่านี้คือ เมื่อ J ≈ n การแจกแจงภายใต้ n-asymptotics คือ χ^2_{n-p-1} จะเห็นว่าจำนวนพารามิเตอร์จะเพิ่มขึ้นตามขนาดของตัวอย่าง เช่นนี้ตัวสถิติข้างต้นภายใต้การแจกแจง χ^2_{n-p-1} จะให้ค่า p-value ที่ไม่ถูกต้อง (Hosmer and Lemeshow 1989)

McCullagh and Nelder (1983) แสดงให้เห็นว่า X^2 และ D มีค่าคาดหวังน้อยกว่าองศาอิสระ J-(p+1) ทางหนึ่งที่จะเหลีกเลี่ยงปัญหาดังกล่าวคือการจัดกลุ่มข้อมูลเพื่อให้แต่ละกลุ่มสอดคล้องกับ m-asymptotics เพื่อให้เห็นชัดเจนจึงจัดข้อมูลให้เป็นตารางการณ์จรขนาด (2xJ) แนวนอนแบ่งกลุ่มเป็น J กลุ่มที่สอดคล้องกับ covariate pattern แนวตั้งจัดกลุ่มตามผลลัพธ์ที่เป็นไปได้ของตัวแปรตอบสนอง y = 1, 0 ค่าคาดหวังภายใต้สมมติฐานตัวแบบถูกต้องของตัวแปรตอบสนอง y = 1 ในแต่ละเซลล์คือ $m_j\hat{p}_j$ และค่าคาดหวังภายใต้สมมติฐานตัวแบบถูกต้องของตัวแปรตอบสนอง y = 0 คือ $m_j(1-\hat{p}_j)$ ดังนั้นตัวสถิติ X^2 สามารถคำนวณได้จากผลรวมกำลังสองของระยะห่างระหว่างค่าสังเกตและค่าคาดหวังในแต่ละเซลล์ และการแจกแจงของตัวสถิตินี้คือ $\chi^2_{n-(p+1)}$ เมื่อความถี่คาดหวังของแต่ละเซลล์มีค่าใหญ่ m-asymptotics กรณีที่ J ≈ n จะส่งผลให้ความถี่คาดหวังในแต่ละเซลล์มีค่าน้อยซึ่งไม่สอดคล้องกับ m-asymptotics วิธีแก้ปัญหานี้คือการกำหนดจำนวนกลุ่มที่แน่นอน G กลุ่มโดยที่ G < J แต่ละกลุ่มได้จากการรวมกลุ่มที่มีความถี่คาดหวังน้อย ๆ

Hosmer and Lemeshow (1980:1043-1069) และ Lemeshow and Hosmer(1982:92-106) เสนอวิธีการจัดกลุ่มตามค่าประมาณความน่าจะเป็นของตัวแปรตอบสนอง $\hat{p}$ กรณี J = n (ตัวแปรอิสระแบบต่อเนื่อง)

วิธีการคือจัด n คอลัมน์ให้เรียงตามค่า $\hat{p}$ โดยเรียงจากน้อยไปมาก ดังนั้นคอลัมน์แรกจะมี $\hat{p}$ น้อยที่สุด และคอลัมน์สุดท้ายจะเป็นคอลัมน์ที่มี $\hat{p}$ ที่มีค่ามากที่สุด หลังจากนั้นทำการยุบกลุ่ม วิธีการจัดกลุ่มมี 2 วิธีคือ 1.) จัดกลุ่มตามเปอร์เซ็นต์ไทล์ของ $\hat{p}$ 2.) จัดกลุ่มโดยกำหนดคัทพอยท์ (cut point)สำหรับจัดกลุ่ม $\hat{p}$

สำหรับวิธีแรกใช้ G = 10 นั่นคือกลุ่มแรกมีจำนวนค่าสังเกต $n'_1 = n/10$ ค่าสังเกตที่สอดคล้องกับ $\hat{p}$ น้อยที่สุด และกลุ่มสุดท้ายมีค่าสังเกต $n'_{10} = n/10$ ค่าสังเกตที่สอดคล้องกับ $\hat{p}$ ที่มีค่ามากทีสุด

วิธีที่ 2 ใช้ G = 10 เช่นกัน และใช้คัทพอยท์สำหรับจัดกลุ่มเป็น k/10 , k = 1,2, . . ., 9 ตัวอย่างเช่นกลุ่มแรกคือกลุ่มที่มีค่าสังเกตสอดคล้องกับ $\hat{p} < 1/10 = 0.1$และกลุ่มสุดท้ายเป็นกลุ่มที่มีค่าสังเกตที่สอดคล้องกับ $\hat{p} \geq 9/10 = 0.9$ สำหรับการจัดกลุ่มทั้งสองแบบข้างต้นตัวสถิติที่ใช้คือ

$$\hat{C} = \sum_{k=1}^{g} \frac{(o_k - n_k' \bar{p}_k)^2}{n_k' \bar{p}_k (1 - \bar{p}_k)}$$

เมื่อ n_k' แทน จำนวนค่าสังเกตในกลุ่มที่ k

$$o_k = \sum_{j=1}^{n_k'} y_j$$

แทนจำนวนผลลัพธ์ที่สนใจของตัวแปรตอบสนอง (y = 1) ในกลุ่มที่ k

$$\bar{p}_k = \sum_{j=1}^{n_k'} m_j \hat{p}_j \ / \ n_k'$$

แทนค่าเฉลี่ยของ $\hat{p}_j$ ของกลุ่มที่ k

Hosmer and Lemeshow (1980:1043-1069) แสดงให้เห็นว่า $\hat{C}$ มีการแจกแจงเมื่อใกล้อนันต์แบบ χ^2_{g-2} เมื่อ J = n แต่ไม่ได้แสดงการแจกแจงของ $\hat{C}$ กรณีที่ J ≈ n

อีกทางเลือกหนึ่งสำหรับเทอมส่วนในการคำนวณค่า $\hat{C}$ ถ้าเราพิจารณา o_k เป็นผลรวมของตัวแปรตอบสนองมีการแจกแจงที่ไม่เหมือนกัน เราสามารถทำให้กำลังสองของผลต่างระหว่างค่าสังเกตกับความถี่คาดหวังเป็นค่ามาตรฐานโดยการหารด้วย

$$\sum_{j=1}^{n_k'} p_j' (1 - \hat{p}_j)$$

และสามารถแสดงได้โดยง่ายว่า

$$\sum_{j=1}^{n_k'} p_j' (1 - \hat{p}_j) = n_k' \bar{p}_j (1 - \bar{p}_j) - \sum_{j=1}^{n_k'} (\hat{p}_j - \bar{p}_k)^2$$

ซึ่งจะเห็นได้ว่าถ้าใช้ $\sum_{j=1}^{n_k'} p_j' (1 - \hat{p}_j)$ จะทำให้ค่า $\hat{C}$ มีค่าเพิ่มขึ้นเล็กน้อย แต่ในทางปฏิบัติมักคำนวณ $\hat{C}$ โดยใช้เทอมผลหารเป็น $n_k' \bar{p}_k (1 - \bar{p}_k)$

Hosmer , Lemeshow and Klar (1988:911-924) แสดงให้เห็นว่าการจัดกลุ่มโดยแบ่งตามเปอร์เซ็นต์ไทล์ของ $\hat{p}$ ให้การแจกแจงของ $\hat{C}$ เข้าใกล้แบบ χ^2_{G-2} ดีกว่า การแบ่งกลุ่มโดยใช้คัทพอยท์โดยเฉพาะกรณีที่ $\hat{p}$ มีค่าน้อย ๆ($\hat{p}$ < 0.2) และจำนวนกลุ่มที่นิยมใช้คือ G = 10

เนื่องจากการแจกแจงของตัวสถิติ $\hat{C}$ ขึ้นกับ m-asymptotics ดังนั้น p-value ที่ถูกต้องจึงขึ้นกับความถี่คาดหวังในแต่ละเซลล์ของตารางการณ์จรที่ได้จากการแบ่งกลุ่มค่าสังเกตต้องมีค่าใหญ่พอ กฎคร่าว ๆ คือ ในแต่ละเซลล์จะต้องมีความถี่คาดหวังมากกว่า 1 และอย่างน้อย 80 เปอร์เซ็นต์ของจำนวนเซลล์ทั้งหมดจะต้องมีความถี่คาดหวังมากกว่า 5 ดังนั้นในกรณีที่แบ่งกลุ่มแล้วผลที่ได้ไม่เป็นไปตามกฎดังกล่าวจึงมีวิธีแก้ไขโดยการลดจำนวนกลุ่มลงโดยการยุบกลุ่มที่มีความถี่คาดหวังน้อย ๆ ที่อยู่ติดกันรวมเข้าด้วยกัน และจำนวนกลุ่มไม่ควรน้อยกว่า 6 กลุ่มเนื่องจากถ้าจำนวนกลุ่มน้อยเกินไปจะทำให้ไม่เห็นความแตกต่างระหว่างค่าสังเกตกับความถี่คาดหวัง

Tsiatis (1980:250-251) เสนอตัวสถิติที่คล้ายกับ ตัวสถิติ $\hat{C}$ โดยในการสร้างตัวสถิติจะแบ่งกลุ่มค่าสังเกตออกเป็น G กลุ่ม โดยสร้างฟังก์ชันดัชนี (indicator functions) $I^{(j)}$, $j = 1,\ldots, G$ เพื่อใช้ในการระบุกลุ่มข้อมูล โดยที่ $I^{(j)} = 1$ ถ้า $(x_1,\ldots,x_p) \in R_j$ และ $I^{(j)} = 0$ กรณีอื่น ๆ เมื่อ R_j แทนกลุ่มที่ j, j=1,…,G. และ ตัวสถิติดังกล่าวคือ ตัวสถิติที่ใช้ทดสอบสัมประสิทธิ์ของตัวแปรที่กำหนดขึ้นใหม่ดังกล่าว Tsiatis แสดงให้เห็นว่าตัวสถิติสำหรับสกอร์เทสสามารถคำนวณในรูปควอดราติกฟอร์มระหว่างค่าสังเกตกับความถี่คาดหวังในแต่ละเซลล์และตัวสถิติดังกล่าวมีการแจกแจงโดยประมาณแบบไคสแควร์เมื่อใกล้อนันต์

Lipsitz et al. (1996:175-190) ศึกษาเกี่ยวกับการรักษาอาการปวดไขข้อกระดูก ซึ่งได้เปรียบเทียบระหว่างการใช้ยา ออราโนฟิน กับยาหลอก ตัวแปรตามประเมินโดยคนไข้ให้คะแนนความรู้สึกปวดไขข้อ 3 ระดับคือ poor(1) fair(2) และ good(3) จากตัวอย่างสุ่มคนไข้จำนวน 293 คน คนไข้แต่ละคนจะได้รับทรีทเมนท์แบบสุ่มจากระดับของทรีตเมนท์ 2 ระดับ คือให้ยาจริง กับยาหลอก การทดลองจะเก็บข้อมูลระดับของอาการปวดข้อของคนไข้แต่ละคนเป็นมาตรฐาน ตัวแบบที่ใช้ศึกษาคือ **proportional odds model** ตัวสถิติที่ใช้ทดสอบภาวะสารูปดีคือตัวสถิติ likelihood ratio, Wald และ score ซึ่งใช้ร่วมกับตัวแบบที่ใช้สำหรับการทดสอบภาวะสารูปดีที่ขยายจากตัวสถิติทดสอบของ Hosmer and Lemeshow การทดสอบดังกล่าวทำได้โดยแบ่งข้อมูลเป็น 10 กลุ่มตามเปอร์เซ็นต์ไทล์ของค่าพยากรณ์ของตัวแปรตอบสนองที่ได้จากตัวแบบตามสมมติฐานว่าง จากการแบ่งกลุ่มข้อมูล ตัวสถิติทดสอบภาวะสารูปดีถูกสร้างขึ้นโดยนิยาม อินดิเคเตอร์ (indicator) สำหรับ G-1 กลุ่ม ซึ่งเป็นฟังก์ชันดัชนีดังนี้

$$I_{ig} = \begin{cases} 1 & \textit{if } \hat{\mu}_i \textit{ is in region } g, \\ 0 & \textit{if otherwise}, \end{cases} \quad i=1, \ldots, n \text{ และ } g=1, \ldots, G-1$$

แล้วการประเมินภาวะสารูปดีของสมมติฐานว่าง $L_{ik} = L_{ik}(\mathbf{p}_i) = \alpha_k + x_i'\beta_1$

จะพิจารณากับ ตัวแบบทางเลือก

$$L_{ik} = \alpha_k + x_i'\beta_1 + \sum_{g=1}^{G-1} I_{ig}\gamma \quad k=1, \ldots, K-1$$

ถ้าสมมติฐานว่างเป็นจริงแล้ว

$$\gamma_{1,1} = \ldots = \gamma_{G-1,1} = \gamma_{1,2} = \ldots = \gamma_{G-1,2} = \gamma_{1,K-1} = \ldots = \gamma_{G-1,K-1} = 0$$

อย่างไรก็ตามถ้าเราสนใจสมมติฐานทางเลือกที่ไม่เป็นไปตามข้างต้นเราสามารถพิจารณาตัวแบบทางเลือก

$$L_{ik} = \alpha_{ik} + x_i'\beta_{1k} + \sum_{g=1}^{G-1} I_{ig}\gamma_{gk}$$

ตัวแบบนี้มีพารามิเตอร์ $\{\gamma_{gk}\}$ จำนวน(G-1)(K-1) ที่มากกว่าตัวแบบตามสมมติฐานว่าง อีกครั้งถ้าสมมติฐานว่างถูกต้อง แล้ว $\gamma_{gk}= 0$

$\gamma_{1,1} = \ldots = \gamma_{G-1,1} = \gamma_{1,2} = \ldots = \gamma_{G-1,2} = \gamma_{1,K-1} = \ldots = \gamma_{G-1,K-1} = 0$ จากการทดสอบภาวะสารูปดีแบบต่างๆ ดังที่กล่าวมาข้างต้นพบว่าตัวแบบที่เหมาะสมที่สุดคือ ตัวแบบที่มีแต่อิทธิพลหลัก และหลังจากการทดสอบพลังการทดสอบ(%)พบว่าตัวแบบทางเลือก

$$L_{ik} = \alpha_k + x_i'\beta_1 + \sum_{g=1}^{G-1} I_{ig}\gamma$$

เป็นตัวแบบที่มีความเสถียร (stable) มากที่สุดเมื่อกำหนดตัวแบบไม่ถูกต้องในการทดสอบภาวะสารูปดีของตัวแบบอิทธิพลหลัก

นอกจากนี้ใน statistical packages ต่าง ๆ ยังนิยมใช้ตัวสถิติวาลด์(Wald statistic) เป็นทางเลือกของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด ในการวิเคราะห์ตัวแบบเชิงเส้นที่วางนัยทั่วไปโดย Wald (1943) ได้แสดงผลลัพธ์ใกล้อนันต์ (asymptotics results) ของตัวประมาณภาวะน่าจะเป็นสูงสุดของพารามิเตอร์ในตัวแบบว่ามีการแจกแจงแบบปกติ และใช้ในการทดสอบเชิงสถิติด้วยตัวสถิติ Wald ซึ่งมีสูตรดังนี้

$$z = (\hat{\beta} - \beta_0) / SE(\hat{\beta})$$

หรือกรณีตัวแปรพหุคือ

$$W = \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right)' \left[\mathrm{cov}(\hat{\boldsymbol{\beta}})\right]^{-1} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right)$$

อนึ่งตัวสถิติ Wald ยังมีจุดด้อยอยู่บ้างเช่น ถ้าค่า $|\hat{\beta}| = \infty$ หรือ $SE(\hat{\beta}) = \infty$ จะไม่สามารถคำนวณค่าของตัวสถิติ z หรือ W ได้ และถ้าตัวอย่างมีขนาดใหญ่มาก ยังทำให้ค่า $Cov(\hat{\beta})$ ต่ำลงมากขึ้น ตลอดจน Agresti (2002) กล่าวว่า กรณีตัวอย่างมีขนาดเล็ก การแจกแจงของ $\hat{\beta}$ อาจเบี่ยงเบนไปจากการแจกแจงแบบปกติ ซึ่งกรณีนี้อาจเกิดขึ้นได้เช่นกันสำหรับกรณีที่ตัวอย่างมีขนาดปานกลางถึงขนาดใหญ่ เมื่อตัวแบบเชิงสถิติประกอบด้วยพารามิเตอร์ในตัวแบบจำนวนมาก

บทที่ 3

วิธีการดำเนินการวิจัย

การดำเนินงานวิจัยของงานวิจัยนี้ประกอบด้วยกรณีหลัก ๆ 4 กรณีดังต่อไปนี้

กรณีที่ 1 การเปรียบเทียบพลังการทดสอบของตัวสถิติภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **pairwise** เมื่อตัวแบบจริงเป็นตัวแบบ **pairwise**

ศึกษาเปรียบเทียบพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็น กับตัวสถิติวาลด์ สำหรับตัวแบบทางเลือก (5) **partition** เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **pairwise** ที่มีเทอม interaction ระหว่างตัวแปรอธิบาย x_1 และ x_2 โดยมีขั้นตอนดังนี้

1. สร้างแบบจำลองข้อมูลโดย X_1~Bernulli(0.5) และ X_2~Normal(0,1)

2. กำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **pairwise**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad k = 1,2 \qquad i = 1,2,\ldots, n$$

3. คำนวณค่าความน่าจะเป็นของตัวแปรตอบสนองที่ k (p_k), k = 1, 2, 3 โดยใช้เซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ดังนี้ { -0.6, 0.2, 0.707, -0.707} และ $\beta_{12} = 0$ สูตรที่ใช้คำนวณเป็นดังนี้

$$p_1 = \frac{\exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{12} x_{1i} x_{2i})}{1 + \exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i})}$$

$$p_2 = \frac{\exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{12} x_{1i} x_{2i})}{1 + \exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i})} - p_1$$

$$p_3 = 1 - p_1 - p_2$$

4. สร้างตัวแปรสุ่ม Y ให้ Y=1 ด้วยความน่าจะเป็น p_1 , Y=2 ด้วยความน่าจะเป็น p_2 และ Y=3 ด้วยความน่าจะเป็น p_3 โดยสุ่มข้อมูลจากการแจกแจงแบบ Uniform(0,1) 1 ค่าเขียนแทนด้วย u แล้วนำไปเปรียบเทียบกับค่า p_j ที่ได้จากข้อ 2. กล่าวคือ

ถ้า $u < p_1$ แล้ว Y=1

ถ้า $p_1 \le u < p_1 + p_2$ แล้ว Y=2

ถ้า $u \ge p_1 + p_2$ แล้ว Y=3 (ทำซ้ำตั้งแต่ข้อ 1-4 จำนวนเท่ากับขนาดตัวอย่าง 600)

5. นำ X_1,X_2 และ Y ที่ได้ไปสร้างตัวแบบ (model fitting) ตามสมมติฐานว่างตัวแบบ(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad k = 1,2 \qquad i = 1,2\ldots, n$$

6. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น(-2D_1)

7. สร้างตัวแบบจากข้อมูลที่จำลองขึ้นตามสมมติฐานทางเลือกตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2 \qquad i = 1,2,...,\ n$$

8. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น($-2D_2$) และตัวสถิติวาลด์(W(partition))

9. คำนวณ Change of deviances D(partition) = $-2(D_1-D_2)$ จากนั้นหา P-value ของค่าสถิติแต่ละตัวซึ่งใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 9

10.นำข้อมูลเดียวกันสร้างตัวแบบตามสมมติฐานทางเลือกตัวแบบ(6) **pairwise**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad k = 1,2 \qquad i = 1,2,...,\ n$$

11.บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น($-2D_3$) และตัวสถิติวาลด์(W(pairwise))

12. คำนวณ Change of deviances D(pairwise) = $-2(D_1-D_3)$ จากนั้นหา P-value ของค่าสถิติแต่ละตัวซึ่งใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 1

13. เปรียบเทียบค่า P-value ที่ได้ถ้ามีค่าน้อยกว่า 0.05 จะปฏิเสธสมมติฐานว่าง

14. ทำซ้ำตั้งแต่ข้อ 4-10 จำนวน 500 ครั้ง

15. ทำซ้ำตั้งแต่ข้อ 2-11 โดยเพิ่มค่า β_{12} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

16. พลอตกราฟระหว่าง β_{12} กับพลังการทดสอบ(%)แล้วประเมินผลการวิจัย

17. ทำซ้ำตั้งแต่ข้อ 1-16 โดยเปลี่ยนขนาดตัวอย่าง เป็น 1000 และ 1500 ตามลำดับ

กรณีที่ 2 การเปรียบเทียบพลังการทดสอบของตัวสถิติภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **pairwise** เมื่อตัวแบบจริงเป็นตัวแบบ **quadratic**

ศึกษาเปรียบเทียบพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็น กับตัวสถิติวาลด์ สำหรับตัวแบบทางเลือก (5) **partition** เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบ(7) **quadratic** ที่มีเทอมกำลังสองของตัวแปรอธิบาย x_2 โดยมีขั้นตอนดังนี้

1. สร้างแบบจำลองข้อมูลโดย X_1~Bernulli(0.5) และ X_2~Normal(0,1)

2. กำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **quadratic**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2 \qquad k = 1,2 \qquad i = 1,2,...,\ n$$

3. คำนวณค่าความน่าจะเป็นของตัวแปรตอบสนองที่ k (p_k), k = 1, 2, 3 โดยใช้เซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ดังนี้ { -0.6, 0.2, 0.707, -0.707} และ $\beta_{22} = 0$ สูตรที่ใช้คำนวณเป็นดังนี้

$$p_1 = \frac{\exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{22} x_{2i}^2)}{1 + \exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2)}$$

$$p_2 = \frac{\exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{22} x_{2i}^2)}{1 + \exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2)} - p_2$$

$$p_3 = 1 - p_1 - p_2$$

4. สร้างตัวแปรสุ่ม Y ให้ Y=1 ด้วยความน่าจะเป็น p_1 , Y=2 ด้วยความน่าจะเป็น p_2 และ Y=3 ด้วยความน่าจะเป็น p_3 โดยสุ่มข้อมูลจากการแจกแจงแบบ Uniform(0,1) 1 ค่าเขียนแทนด้วย u แล้วนำไปเปรียบเทียบกับค่า p_j ที่ได้จากข้อ 2. กล่าวคือ

ถ้า $u < p_1$ แล้ว Y=1

ถ้า $p_1 \le u < p_1 + p_2$ แล้ว Y=2

ถ้า $u \ge p_1 + p_2$ แล้ว Y=3 (ทำซ้ำตั้งแต่ข้อ 1-4 จำนวนเท่ากับขนาดตัวอย่าง 600)

5. นำ X_1,X_2 และ Y ที่ได้ไปสร้างตัวแบบตามสมมติฐานว่างตัวแบบ(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad k = 1,2 \qquad i = 1,2..., n$$

6. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น(-2D_1)

7. สร้างตัวแบบจากข้อมูลที่จำลองขึ้นตามสมมติฐานทางเลือกตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2 \qquad i = 1,2,..., n$$

8. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น(-2D_2) และตัวสถิติวาลด์(W(partition))

9. คำนวณ Change of deviances D(partition) = -2(D_1-D_2) จากนั้นหา P-value ของค่าสถิติแต่ละตัวซึ่งใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 9

10.นำข้อมูลเดียวกันสร้างตัวแบบตามสมมติฐานทางเลือกตัวแบบ(6) **pairwise**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad k = 1,2 \qquad i = 1,2,..., n$$

11.บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น(-2D_3) และตัวสถิติวาลด์(W(pairwise))

12. คำนวณ Change of deviances D(pairwise) = -2(D_1-D_3) จากนั้นหา P-value ของค่าสถิติแต่ละตัวซึ่งใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 1

13. เปรียบเทียบค่า P-value ที่ได้ถ้ามีค่าน้อยกว่า 0.05 จะปฏิเสธสมมติฐานว่าง

14. ทำซ้ำตั้งแต่ข้อ 4-10 จำนวน 500 ครั้ง

15. ทำซ้ำตั้งแต่ข้อ 2-11 โดยเพิ่มค่า β_{22} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

16. พลอตกราฟระหว่าง β_{22} กับพลังการทดสอบ(%)แล้วประเมินผลการวิจัย

17. ทำซ้ำตั้งแต่ข้อ 1-16 โดยเปลี่ยนขนาดตัวอย่าง เป็น 1000 และ 1500 ตามลำดับ

กรณีที่ 3 การเปรียบเทียบพลังการทดสอบของตัวสถิติอัตราส่วนภาวะน่าจะเป็นภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **baseline-partition** เมื่อตัวแบบจริงคือตัวแบบ **pairwise**

ศึกษาความแตกต่างของพลังการทดสอบ(%)ของการใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับ กับการใช้ตัวแบบสำหรับตัวแปรตอบสนองที่ไม่มีลำดับ โดยที่จริง ๆ แล้วตัวแปรตอบสนองที่ต้องการศึกษาเป็นตัวแปรแบบมีลำดับ เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบที่มีเทอม interaction ระหว่างตัวแปรอธิบาย x_1 และ x_2 โดยมีขั้นตอนดังนี้

1. สร้างแบบจำลองข้อมูลโดย X_1~Bernulli(0.5) และ X_2~Normal(0,1)

2. กำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **pairwise**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad k = 1,2 \qquad i = 1,2,\ldots,\ n$$

3. คำนวณค่าคาดหวังของตัวแปรตาม p_k , $k = 1,\ 2,\ 3$ โดยใช้เซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ดังนี้ { -0.6, 0.2, 0.707, -0.707} และ $\beta_{12} = 0$ สูตรที่ใช้คำนวณเป็นดังนี้

$$p_1 = \frac{\exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{12} x_{1i} x_{2i})}{1 + \exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i})}$$

$$p_2 = \frac{\exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{12} x_{1i} x_{2i})}{1 + \exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i})} - p_1$$

$$p_3 = 1 - p_1 - p_2$$

4. สร้างตัวแปรสุ่ม Y ให้ Y=1 ด้วยความน่าจะเป็น p_1 , Y=2 ด้วยความน่าจะเป็น p_2 และ Y=3 ด้วยความน่าจะเป็น p_3 โดยสุ่มข้อมูลจากการแจกแจงแบบ Uniform(0,1) 1 ค่าเขียนแทนด้วย u แล้วนำไปเปรียบเทียบกับค่า p_j ที่ได้จากข้อ 2. กล่าวคือ

ถ้า $u < p_1$ แล้ว Y=1

ถ้า $p_1 \le u < p_1 + p_2$ แล้ว Y=2

ถ้า $u \ge p_1 + p_2$ แล้ว Y=3 (ทำซ้ำตั้งแต่ข้อ 1-4 จำนวนเท่ากับขนาดตัวอย่าง 600)

5. นำ X_1,X_2 และ Y ที่ได้ไปสร้างตัวแบบตามสมมติฐานว่างตัวแบบ(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad k = 1,2 \qquad i = 1,2,\ldots,\ n$$

6. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น(-2D_1)

7. นำข้อมูลเดิมสร้างตัวแบบตามสมมติฐานทางเลือกตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig} \gamma_g \qquad k = 1,2 \qquad i = 1,2,\ldots,\ n$$

8. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น(-2D_2) และตัวสถิติวาลด์ W(ordinal)

9. คำนวณ Change of deviances D(ordinal) = -2(D_1-D_2) จากนั้นหา P-value ของค่าสถิติแต่ละตัวซึ่งใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 9

10. เปรียบเทียบค่า P-value ที่ได้ถ้ามีค่าน้อยกว่า 0.05 จะปฏิเสธสมมติฐานว่าง

11. ทำซ้ำตั้งแต่ข้อ 4-10 จำนวน 500 ครั้ง

12. ทำซ้ำข้อ 5-10 โดยเปลี่ยนสมมติฐานว่างเป็นตัวแบบ(8) **baseline-category**

$$\log\frac{p_k(\mathbf{x})}{p_3(\mathbf{x})} = \alpha_k + \beta_{1k}x_{1i} + \beta_{2k}x_{2i} \qquad k = 1,2, i = 1,..., n$$

และเปลี่ยนสมมติฐานทางเลือกเป็นตัวแบบ(9) **baseline-partition**

$$\log\frac{p_k(\mathbf{x})}{p_3(\mathbf{x})} = \alpha_k + \beta_{1k}x_{1i} + \beta_{2k}x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_{gk} \qquad j = 1,2 \qquad i = 1,2,..., n$$

คำนวณค่าสถิติ D(nominal) และ W(nominal) ใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 18 สำหรับหา P-value

13. ทำซ้ำตั้งแต่ข้อ 2-11 โดยเพิ่มค่า β_{12} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

14. พลอตกราฟระหว่าง β_{12} กับพลังการทดสอบ(%)แล้วประเมินผลการวิจัย

15. ทำซ้ำตั้งแต่ข้อ 1-14 โดยเปลี่ยนขนาดตัวอย่าง เป็น 1000 และ 1500 ตามลำดับ

กรณีที่ 4 การเปรียบเทียบพลังการทดสอบของตัวสถิติอัตราส่วนภาวะน่าจะเป็นภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **baseline-partition** เมื่อตัวแบบจริงคือตัวแบบ **quadratic**

ศึกษาความแตกต่างของพลังการทดสอบ(%)ของการใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับ กับการใช้ตัวแบบสำหรับตัวแปรตอบสนองที่ไม่มีลำดับ โดยที่จริง ๆ แล้วตัวแปรตอบสนองที่ต้องการศึกษาเป็นตัวแปรแบบมีลำดับ (3) เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบที่มีเทอมกำลังสองของตัวแปรอธิบาย$(\boldsymbol{X_2})$ โดยมีขั้นตอนดังนี้

1. สร้างแบบจำลองข้อมูลโดย X_1~Bernulli(0.5) และ X_2~Normal(0,1)

2. กำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ(7) **quadratic**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22}x_{2i}^2 \qquad k = 1,2 \qquad i = 1,2,..., n$$

3. คำนวณค่าคาดหวังของตัวแปรตาม p_k, $k = 1, 2, 3$ โดยใช้เซตของพารามิเตอร์ $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ ดังนี้ { -0.6, 0.2, 0.707, -0.707} และ $\beta_{22} = 0$ สูตรที่ใช้คำนวณเป็นดังนี้

$$p_1 = \frac{\exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{22} x_{2i}^2)}{1+\exp(\alpha_1 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2)}$$

$$p_2 = \frac{\exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{i2} + \beta_{22} x_{2i}^2)}{1+\exp(\alpha_2 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2)} - p_2$$

$$p_3 = 1 - p_1 - p_2$$

4. สร้างตัวแปรสุ่ม Y ให้ Y=1 ด้วยความน่าจะเป็น p_1 , Y=2 ด้วยความน่าจะเป็น p_2 และ Y=3 ด้วยความน่าจะเป็น p_3 โดยสุ่มข้อมูลจากการแจกแจงแบบ Uniform(0,1) 1 ค่าเขียนแทนด้วย u แล้วนำไปเปรียบเทียบกับค่าp_j ที่ได้จากข้อ 2. กล่าวคือ

ถ้า $u < p_1$ แล้ว Y=1

ถ้า $p_1 \le u < p_1 + p_2$ แล้ว Y=2

ถ้า $u \ge p_1 + p_2$ แล้ว Y=3 (ทำซ้ำตั้งแต่ข้อ 1-4 จำนวนเท่ากับขนาดตัวอย่าง 600)

5. นำ X_1,X_2 และ Y ที่ได้ไปสร้างตัวแบบตามสมมติฐานว่างตัวแบบ(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad k = 1,2 \qquad i = 1,2,\ldots, n$$

6. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น($-2D_1$)

7. นำข้อมูลเดิมสร้างตัวแบบตามสมมติฐานทางเลือกตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2 \qquad i = 1,2,\ldots, n$$

8. บันทึกค่าสถิติอัตราส่วนภาวะน่าจะเป็น($-2D_2$) และตัวสถิติวาลด์ W(ordinal)

9. คำนวณ Change of deviances D(ordinal) = $-2(D_1-D_2)$ จากนั้นหา P-value ของค่าสถิติแต่ละตัวซึ่งใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 9

10. เปรียบเทียบค่า P-value ที่ได้ถ้ามีค่าน้อยกว่า 0.05 จะปฏิเสธสมมติฐานว่าง

11. ทำซ้ำตั้งแต่ข้อ 4-10 จำนวน 500 ครั้ง

12. ทำซ้ำข้อ 5-10 โดยเปลี่ยนสมมติฐานว่างเป็นตัวแบบ(8) **baseline-category**

$$\log\frac{p_k(\mathbf{x})}{p_3(\mathbf{x})} = \alpha_k + \beta_{1k} x_{1i} + \beta_{2k} x_{2i} \qquad k = 1,2, i = 1,\ldots, n$$

และเปลี่ยนสมมติฐานทางเลือกเป็นตัวแบบ(9) **baseline-partition**

$$\log\frac{p_k(\mathbf{x})}{p_3(\mathbf{x})} = \alpha_k + \beta_{1k} x_{1i} + \beta_{2k} x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_{gk} \qquad j = 1,2 \qquad i = 1,2,\ldots, n$$

คำนวณค่าสถิติ D(nominal) และ W(nominal) ใช้การแจกแจงแบบไคสแควร์ องศาอิสระ 18 สำหรับหา P-value

13. ทำซ้ำตั้งแต่ข้อ 2-11 โดยเพิ่มค่า β_{22} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

14. พลอตกราฟระหว่าง β_{22} กับพลังการทดสอบ(%)แล้วประเมินผลการวิจัย

15. ทำซ้ำตั้งแต่ข้อ 1-14 โดยเปลี่ยนขนาดตัวอย่าง เป็น 1000 และ 1500 ตามลำดับ ในแต่ละกรณี พลังการทดสอบ(%)เท่ากับจำนวนครั้งที่ปฏิเสธสมมติฐานหารด้วย 5 ขั้นตอนต่าง ๆ ของการวิจัยทั้ง 4 กรณีได้เขียนเป็น macro program และประมวลผลร่วมกับโปรแกรม Minitab version. 11

บทที่ 4

ผลการวิจัย

ในงานวิจัยนี้มุ่งพิจารณาพลังการทดสอบ(%) ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ ของตัวแบบทางเลือกตัวแบบ(5) **partition**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad , k = 1,2 \qquad i = 1,2,..., n$$

โดยใช้ข้อมูลที่จำลองแบบขึ้นมาเพื่อศึกษาพฤติกรรมของตัวแบบทางเลือกดังกล่าวในกรณีต่าง ๆ ที่มีตัวแปรตอบสนองแบบมีลำดับและไม่มีลำดับ ผลของงานวิจัยนี้แบ่งออกเป็น 4 ตอนดังนี้

กรณีที่ 1 การเปรียบเทียบพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** กับตัวแบบ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **pairwise**

กรณีที่ 2 การเปรียบเทียบพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** กับตัวแบบ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบ(7) **quadratic**

กรณีที่ 3 การเปรียบเทียบพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** กับตัวแบบ(9) **baseline-partition** สำหรับตัวแปรตอบสนองแบบมีลำดับและไม่มีลำดับ เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **pairwise**

กรณีที่ 4 การเปรียบเทียบพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** กับตัวแบบ(9) **baseline-partition** สำหรับตัวแปรตอบสนองแบบมีลำดับและไม่มีลำดับ เมื่อตัวแบบที่ถูกต้องเป็นตัวแบบ(7) **quadratic**

กรณีที่ 1 การเปรียบเทียบพลังการทดสอบของตัวสถิติภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **pairwise** เมื่อตัวแบบจริงเป็นตัวแบบ **pairwise**

กรณีที่ 1. ศึกษาพลังการทดสอบ(%)ของตัวแบบทางเลือก(5) **partititon** เปรียบเทียบกับตัวแบบทางเลือก(6) **pairwise** เมื่อตัวแบบที่ถูกต้องเป็น ตัวแบบ(6) **pairwise** โดยใช้ตัวแบบตามสมมติฐานว่างเป็นตัวแบบ(4) **minimal**

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} \qquad k = 1,2 \qquad i = 1,2..., n$$

ตัวแบบ(5) **partition** คือ

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{g=1}^{G-1} I_{ig}\gamma_g \qquad k = 1,2 \qquad i = 1,2,..., n$$

และตัวแบบ(6) **pairwise** คือ

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} \qquad k = 1,2 \qquad i = 1,2,..., n$$

ตารางที่ 1 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด(LR) และตัวสถิติวาลด์ (Wald) ภายใต้ตัวแบบ(5) **partition** และตัวแบบ(6) **pairwise** เมื่อเพิ่มค่า β_{12} กรณีตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **pairwise** จำแนกตามขนาดตัวอย่าง

β_{12}	พลังการทดสอบ(%)											
	ขนาดตัวอย่าง 600				ขนาดตัวอย่าง 1000				ขนาดตัวอย่าง 1500			
	partition_LR	Pairwise_LR	partition_wald	Pairwise_wald	partition_LR	Pairwise_LR	partition_wald	Pairwise_wald	partition_LR	Pairwise_LR	partition_wald	Pairwise_wald
0.0	4.6	4.8	3.4	4.8	4.6	4.6	4.6	4.8	6.2	5.4	5.4	5.4
0.1	5.8	6.8	5.0	6.6	5.6	10.8	5.8	10.8	5.0	13.0	4.2	12.8
0.2	6.8	23.0	6.4	22.8	7.8	35.0	7.6	35.0	8.2	49.0	7.6	49.0
0.3	9.6	43.6	8.8	43.0	13.4	68.4	12.2	68.4	25.0	82.6	24.4	82.6
0.4	18.4	72.0	17.4	71.8	22.4	88.4	21.6	88.6	43.6	96.8	42.2	96.8
0.5	20.6	90.0	18.6	89.6	45.0	97.4	43.6	97.4	65.2	100.0	64.6	100.0
0.6	49.2	95.6	46.8	95.6	67.0	99.8	65.8	99.8	87.4	100.0	86.8	100.0
0.7	67.6	99.6	65.2	99.8	91.8	99.8	91.6	99.8	96.4	100.0	96.2	100.0
0.8	83.0	100.0	-	100.0	97.4	100.0	-	100.0	100.0	100.0	-	100.0
0.9	96.6	100.0	-	100.0	99.8	100.0	-	100.0	100.0	100.0	-	100.0
1.0	99.0	100.0	-	100.0	100.0	100.0	-	100.0	100.0	100.0	-	100.0
1.1	99.4	100.0	-	100.0	100.0	100.0	-	100.0	100.0	100.0	-	100.0
1.2	99.8	100.0	-	100.0	100.0	100.0	-	100.0	100.0	100.0	-	100.0
1.3	100.0	100.0	-	100.0	100.0	100.0	-	100.0	100.0	100.0	-	100.0
1.4	100.0	100.0	-	100.0	100.0	100.0	-	100.0	100.0	100.0	-	100.0
Mean	57.36	75.69	-	75.60	63.65	80.28	-	80.31	69.13	83.12	-	83.11
S.D.	41.80	36.74	-	36.83	41.75	34.42	-	34.40	39.77	32.90	-	32.94

- หาค่าไม่ได้เนื่องจากตัวประมาณลู่ออก

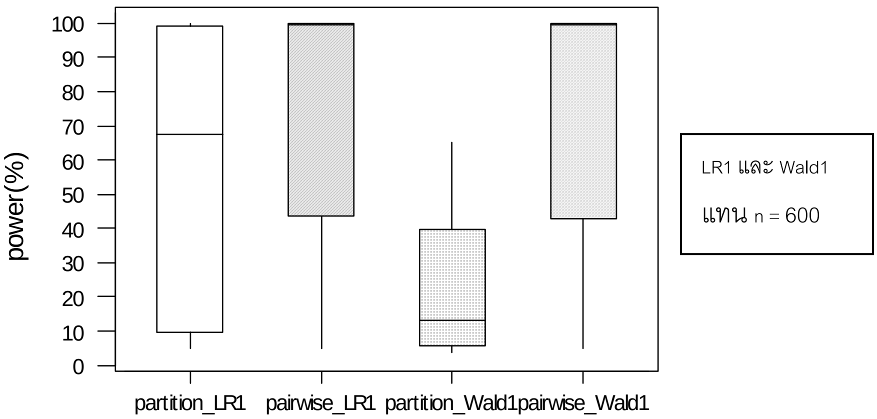
จากตารางที่ 1. การศึกษาพลังการทดสอบ(%)เมื่อกำหนดขนาดตัวอย่าง 600 พบว่าพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุดของตัวแบบ(5) **partition_LR** มีค่าเข้าใกล้ 100% เมื่อ β_{12} มีค่ามากกว่า 1.2 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของตัวแบบ(6) **pairwise_LR** และ **pairwise_wald** มีค่าเข้าใกล้ 100% เมื่อ β_{12} มีค่ามากกว่า 0.7

เมื่อกำหนดขนาดตัวอย่าง 1000 พบว่าพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุดของตัวแบบ(5)) **partition_LR** มีค่าเข้าใกล้ 100% เมื่อ β_{12} มีค่ามากกว่า 0.9 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของตัวแบบ(6) **pairwise_LR** และ **pairwise_Wald** มีค่าเข้าใกล้ 100% เมื่อ β_{12} มีค่ามากกว่า 0.7

เมื่อกำหนดขนาดตัวอย่าง 1500 พบว่าพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุดของตัวแบบ(5)) **partition_LR** มีค่าเข้าใกล้ 100% เมื่อ β_{12} มีค่ามากกว่า 0.7 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของตัวแบบ(6) **pairwise_LR** และ **pairwise_Wald** มีค่าเข้าใกล้ 100% เมื่อ β_{12} มีค่ามากกว่า 0.4

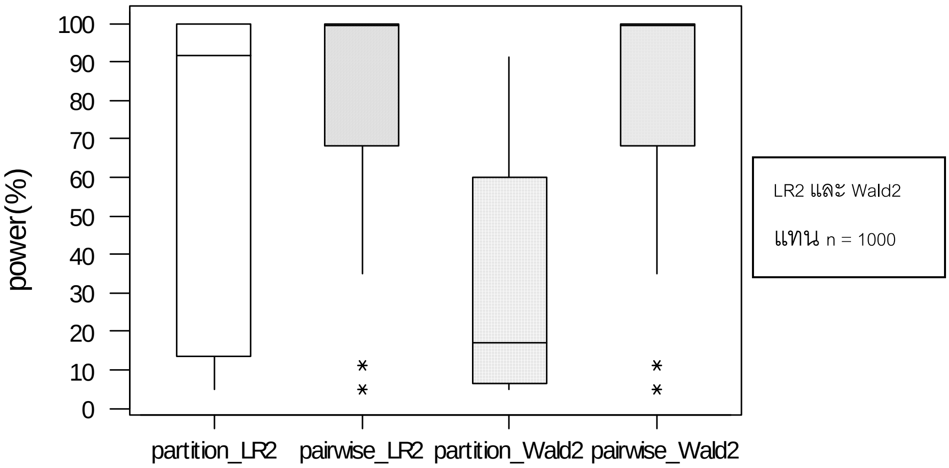
ในการศึกษาพลังการทดสอบ(%)ของตัวสถิติวาลด์ของตัวแบบ(5) **partition_Wald** ไม่สามารถหาค่าได้เมื่อ β_{12} มีค่ามากกว่า 0.7 เนื่องจากการลู่ออกของตัวประมาณพารามิเตอร์ของตัวแบบ(5) ซึ่งค่าพารามิเตอร์ดังกล่าวมีความจำเป็นในการคำนวณค่าสถิติวาลด์

สรุปทั้ง 3 ขนาดตัวอย่างพบว่า ตัวแบบ(6) **pairwise-LR** และ **pairwise_Wald** ให้ค่าพลังการทดสอบ(%)สูงสอดคล้องกับตัวแบบที่ถูกต้องคือตัวแบบ **pairwise** โดยให้ค่าเฉลี่ยพลังการทดสอบ(%) (ภายใต้ค่าพารามิเตอร์ 0-1.4) มากกว่าตัวแบบ(5) **partition** แต่ส่วนเบี่ยงเบนมาตรฐานน้อยกว่าจึงสอดคล้องกับทฤษฎีและความเป็นจริง (รายละเอียดเพิ่มเติมแสดงไว้ใน รูปที่ 1- รูปที่ 10)

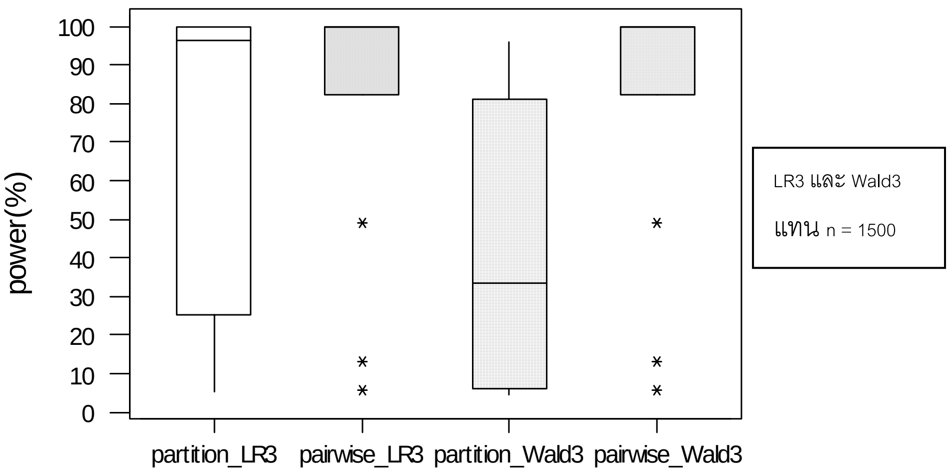


รูปที่ 1. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องคือ (6) **pairwise** และ n = 600

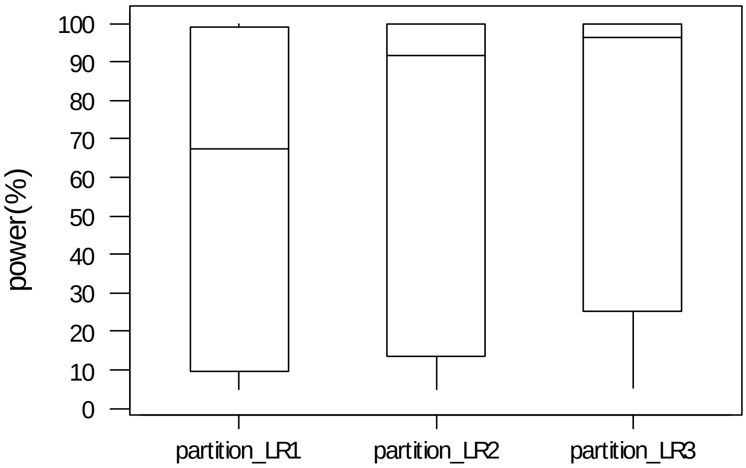
หมายเหตุ รูปที่ 1- รูปที่ 7 : ตัวสถิติวาลด์ภายใต้ตัวแบบ **partition** หาค่าไม่ได้เมื่อ $\beta_{12} > 0.7$



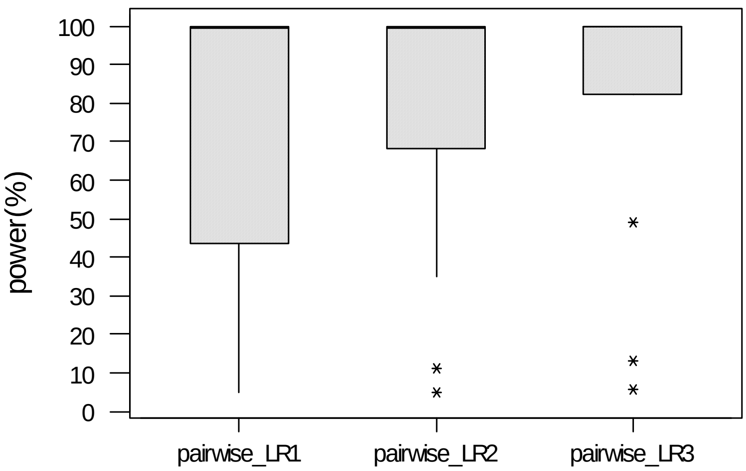
รูปที่ 2. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องคือ (6) **pairwise** และ n = 1000



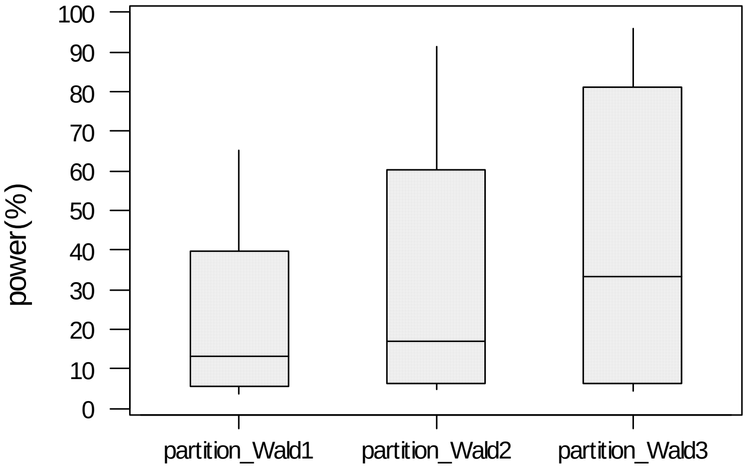
รูปที่ 3. Boxplot ของ พลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องคือ (6) **pairwise** และ n = 1500



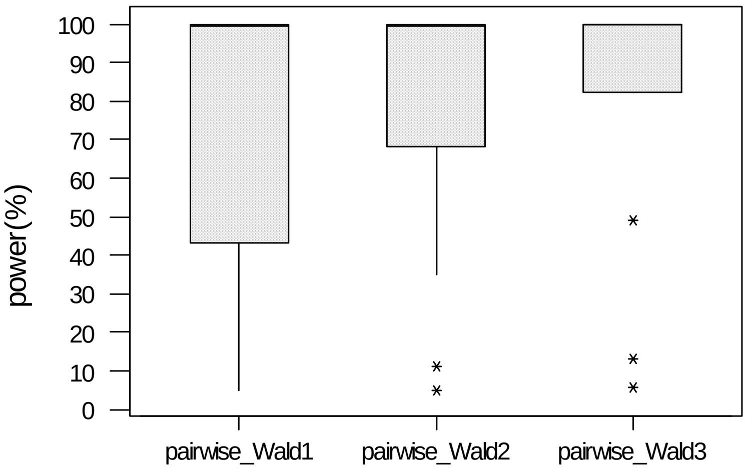
รูปที่ 4. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(5) **partition** ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



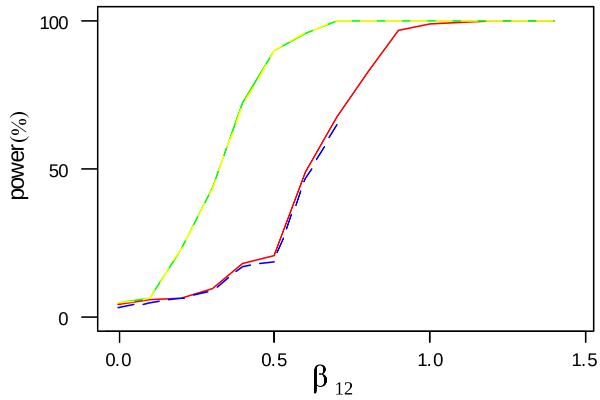
รูปที่ 5. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(6) **pairwise** ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



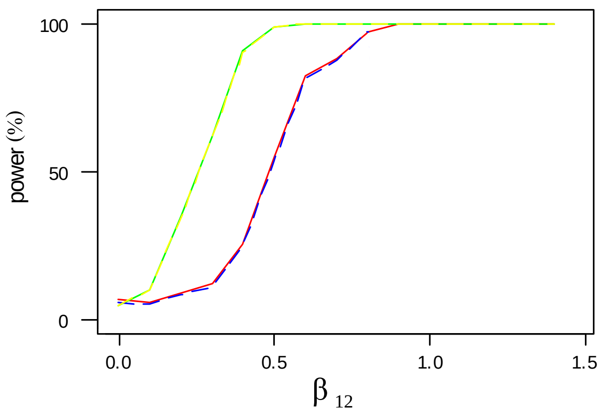
รูปที่ 6.Boxplot ของพลังการทดสอบ(%)ของตัวสถิติวาลด์จากตัวแบบ(5) **partition**
ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



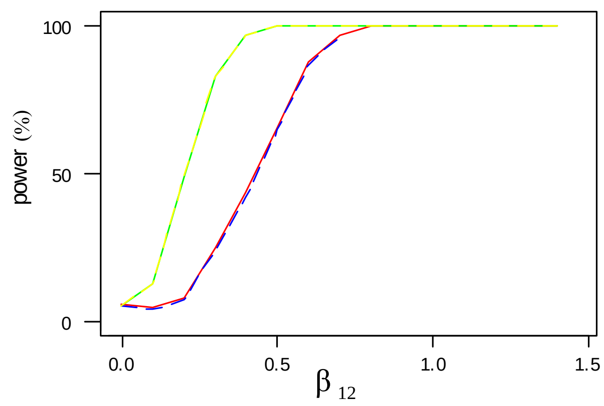
รูปที่ 7.Boxplot ของพลังการทดสอบ(%)ของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise**
ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



รูปที่ 8 กราฟแสดงความสัมพันธ์ระหว่าง β_{12} กับพลังการทดสอบ(%)ภายใต้ตัวแบบต่าง ๆ ขนาดตัวอย่าง 600



รูปที่ 9 กราฟแสดงความสัมพันธ์ระหว่าง β_{12} กับพลังการทดสอบ(%)ภายใต้ตัวแบบต่าง ๆ ขนาดตัวอย่าง 1000



รูปที่ 10 กราฟแสดงความสัมพันธ์ระหว่าง β_{12} กับพลังการทดสอบ(%) ภายใต้ตัวแบบต่าง ๆ ขนาดตัวอย่าง 1500

partition_LR ——— partition_Wald - - - - pairwise_LR - - - - pairwise_Wald - - - -

จากรูปที่ 8-รูปที่ 10 แต่ละกราฟศึกษาที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับ จากกราฟพบว่าพลังการทดสอบ(%)ของตัวสถิติ อัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ ให้พลังการทดสอบ(%)เหมือนกัน เมื่อตัวแบบเป็นตัวแบบ(6) **pairwise_LR** และ **pairwise_Wald** อย่างไรก็ตามค่าสถิติทั้งสองมีความแตกต่างกันเล็กน้อยเมื่อตัวแบบเป็นตัวแบบ(5) **partition_LR** และ **partition_Wald** แต่ความแตกต่างดังกล่าวลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น จากกราฟเมื่อ β_{12} มีค่าตั้งแต่ 0.7 ขึ้นไปไม่สามารถแสดงเส้นกราฟพลังการทดสอบ(%)ได้ เนื่องจาก ในการคำนวณค่าสถิติวาลด์จำเป็นจะต้องใช้ค่าประมาณพารามิเตอร์ทุกตัวในตัวแบบแต่สำหรับตัวแบบ(5) เมื่อ β_{12} มากกว่า 0.7 แบบจำลองที่สร้างขึ้นบางตัวแบบไม่สามารถหาค่าประมาณของพารามิเตอร์บางตัวได้เนื่องจากค่าพารามิเตอร์ดังกล่าวไม่ลู่เข้า

จากกราฟพบว่าพลังการทดสอบ(%)ทั้งตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิตวาลด์ของตัวแบบ(6) **pairwise_LR** และ **pairwise_Wald** แสดงเส้นกราฟสูงกว่าพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของตัวแบบ(5) **partitiion_LR** และ **partition_Wald** อย่างชัดเจน แต่ความแตกต่างดังกล่าวลดลงเมื่อขนาดตัวอย่างมากขึ้น

กรณีที่ 2 การเปรียบเทียบพลังการทดสอบของตัวสถิติภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **pairwise** เมื่อตัวแบบจริงเป็นตัวแบบ **quadratic**

กรณีที่ 2. ศึกษาพลังการทดสอบ(%)ของตัวแบบทางเลือก(5) เปรียบเทียบกับตัวแบบทางเลือก(6) **pairwise** เมื่อตัวแบบที่ถูกต้องเป็น ตัวแบบ(7) ซึ่งมีเทอมควอดราติกของตัวแปรอธิบาย X_2

$$L_{ik} = \alpha_k + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{22} x_{2i}^2 \qquad k = 1,2 \qquad i = 1,2,\ldots,\ n$$

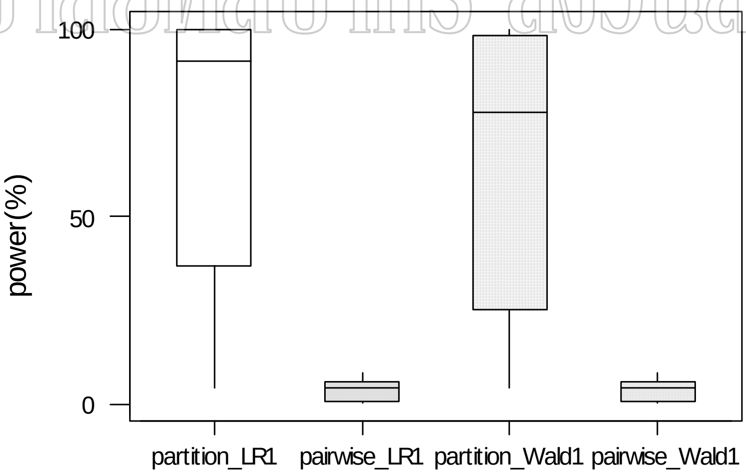
ตารางที่ 2 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด(LR) และตัวสถิติวาลด์ (Wald) ภายใต้ตัวแบบ(5) **partition** และตัวแบบ(6) **pairwise** เมื่อเพิ่มค่า β_{22} กรณีตัวแบบที่ถูกต้องเป็นตัวแบบ(7) **quadratic** จำแนกตามขนาดตัวอย่าง

β_{22}	พลังการทดสอบ(%)											
	ขนาดตัวอย่าง 600				ขนาดตัวอย่าง 1000				ขนาดตัวอย่าง 1500			
	partition_LR	Pairwise_LR	partition_wald	Pairwise_wald	partition_LR	Pairwise_LR	partition_wald	Pairwise_wald	partition_LR	Pairwise_LR	partition_wald	Pairwise_wald
0.0	4.2	6.2	4.2	6.0	5.6	4.2	5.4	4.0	5.0	6.0	4.6	6.0
0.1	10.8	5.0	8.6	4.8	12.6	5.6	11.8	5.6	17.0	4.6	16.2	4.6
0.2	24.4	4.8	22.6	4.6	34.0	4.6	32.2	4.6	58.8	6.8	57.0	6.8
0.3	36.8	5.8	31.8	5.8	55.8	18.0	53.4	18.2	82.4	3.8	80.4	3.8
0.4	69.2	4.2	65.4	4.2	84.4	9.6	80.8	9.8	92.8	5.4	92.4	5.2
0.5	76.6	7.6	73.6	7.6	98.0	7.0	98.2	7.4	100.0	8.8	100.0	9.0
0.6	91.8	3.4	89.8	3.4	99.4	4.6	99.4	5.0	100.0	5.6	99.8	5.8
0.7	86.2	5.2	82.2	5.6	99.8	1.6	99.8	1.6	100.0	12.6	100.0	13.2
0.8	98.6	2.4	97.2	2.4	100.0	1.2	100.0	1.6	100.0	5.8	100.0	6.2
0.9	100.0	8.0	100.0	8.2	100.0	14.6	100.0	15.0	100.0	9.2	100.0	9.8
1.0	99.8	0.6	99.4	0.6	100.0	1.0	100.0	1.4	100.0	0.4	100.0	0.6
1.1	99.6	1.8	98.8	2.0	100.0	0.2	100.0	0.6	100.0	0.6	100.0	0.6
1.2	100.0	0.2	-	0.2	100.0	0.0	-	0.0	100.0	17.4	-	18.4
1.3	100.0	0.2	-	0.2	100.0	3.2	-	3.6	100.0	0.2	-	0.2
1.4	100.0	0.4	-	0.4	100.0	0.2	-	0.2	100.0	19.8	-	20.8
Mean	73.20	3.72	-	3.73	79.31	5.04	-	5.24	83.73	7.13	-	7.40
S.D.	35.68	2.67	-	2.68	34.52	5.37	-	5.41	31.63	5.76	-	6.09

- หาค่าไม่ได้เนื่องจากตัวประมาณลู่ออก

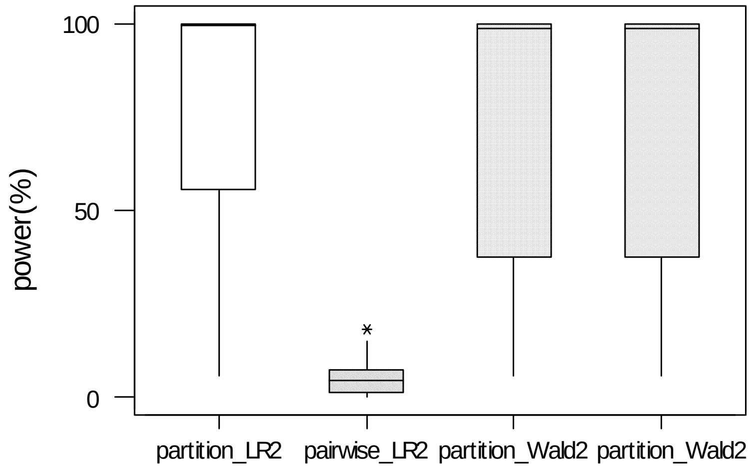
จากตารางที่ 2. การศึกษาพลังการทดสอบ(%)เมื่อกำหนดขนาดตัวอย่าง 600, 1000 และ 1500 พบว่าพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุดของตัวแบบ(5) **partition_LR** มีค่าเข้าใกล้ 100% เมื่อ β_{22} มีค่ามากกว่า 1.1, 0.9 และ 0.4 ตามลำดับ พลังการทดสอบ(%)ของตัวสถิติวาลด์ของตัวแบบ(5) **partition_wald** มีค่าแตกต่างจากตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุดของตัวแบบ(5)อย่างชัดเจนที่ขนาดตัวอย่าง 600 และใกล้เคียงกันเมื่อขนาดตัวอย่างเป็น 1000 และ 1500 อย่างไรก็ตามพลังการทดสอบ(%)ของตัวสถิติวาลด์ของตัวแบบ(5) หาค่าไม่ได้เมื่อ β_{22} มีค่ามากกว่า 1.1 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของตัวแบบ(6) **pairwise_LR** และ **pairwise_wald** มีค่าน้อยมากอย่างเห็นได้ชัดเจนในทุกขนาดตัวอย่างและไม่มีแนวโน้มเข้าใกล้ 100% ไม่ว่าจะเพิ่มค่า β_{22} เท่าใดก็ตาม

สรุปทั้ง 3 ขนาดตัวอย่างพบว่าตัวแบบ(5) **partition** กลับให้ค่าเฉลี่ย (ภายใต้ค่าพารามิเตอร์ 0-1.4) ของพลังการทดสอบ(%)มากกว่าตัวแบบ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องคือตัวแบบ(7) **quadratic** (รายละเอียดเพิ่มเติมแสดงไว้ในรูปที่ 11- รูปที่ 20)

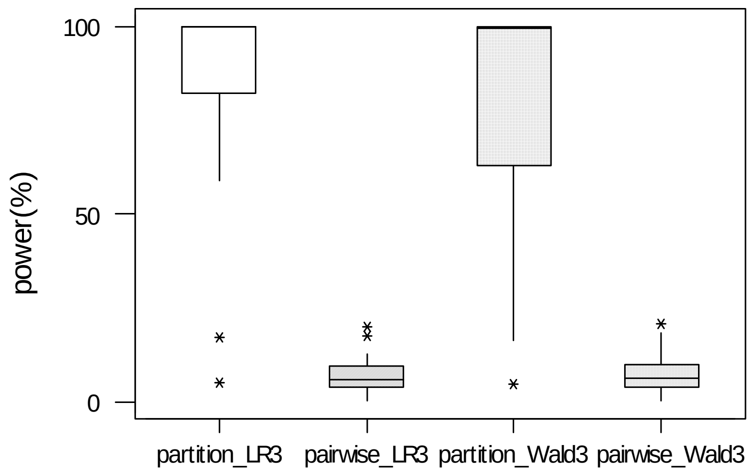


รูปที่ 11. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องคือ (7) **quadratic** และ n = 600

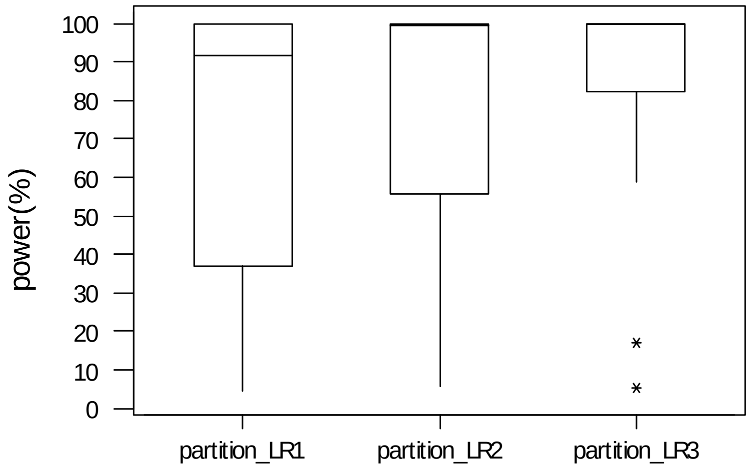
หมายเหตุ รูปที่ 1- รูปที่ 7 : ตัวสถิติวาล์ภายใต้ตัวแบบ **partition** หาค่าไม่ได้เมื่อ $\beta_{22} > 1.1$



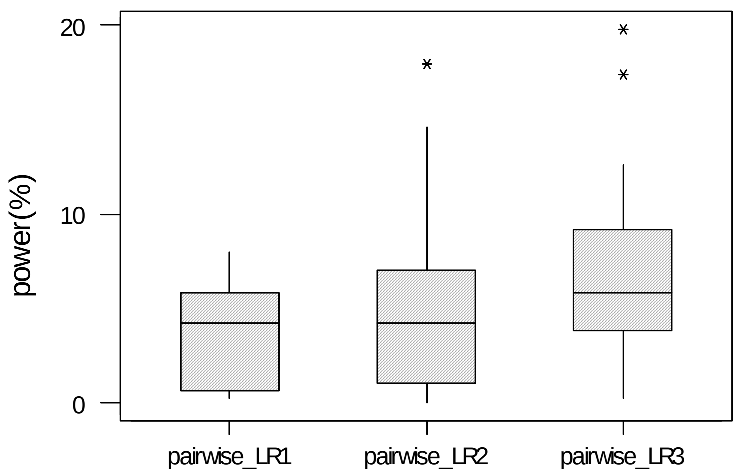
รูปที่ 12. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องคือ (7) **quadratic** และ n = 1000



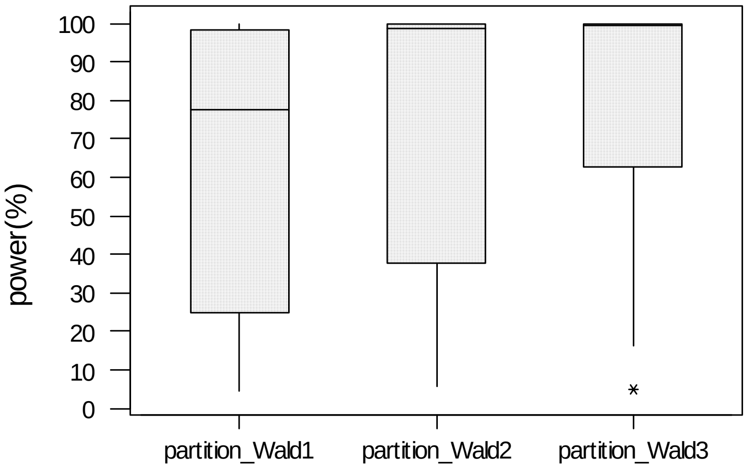
รูปที่ 13. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(6) **pairwise** เมื่อตัวแบบที่ถูกต้องคือ (7) **quadratic** และ n = 1500



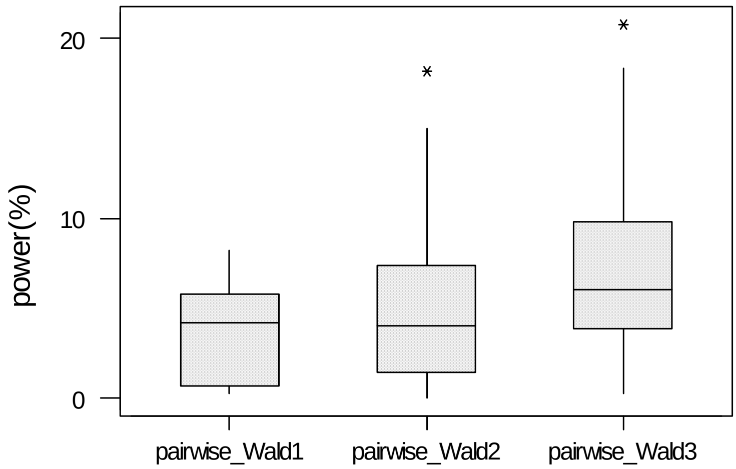
รูปที่ 14. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(5) **partition** ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



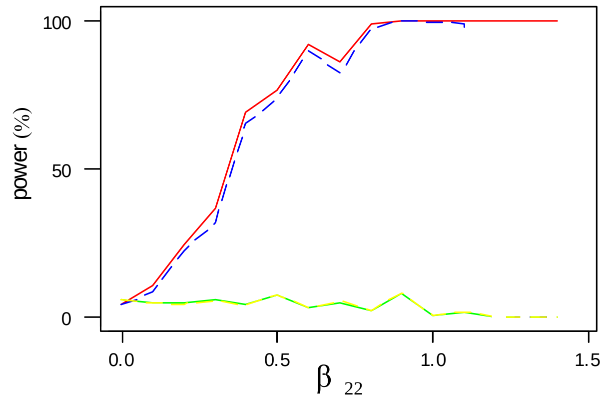
รูปที่ 15. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(6) **pairwise** ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



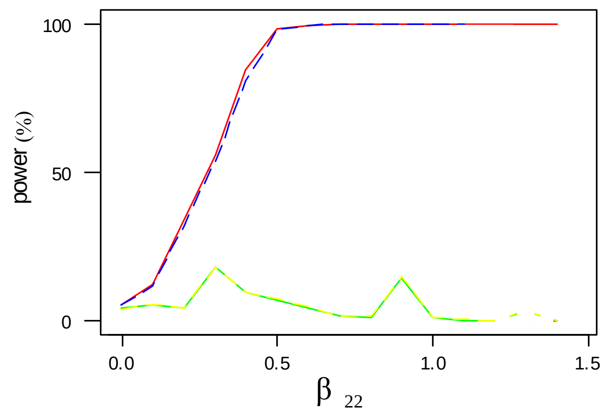
รูปที่ 16. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติวาลด์จากตัวแบบ(5) **partition** ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



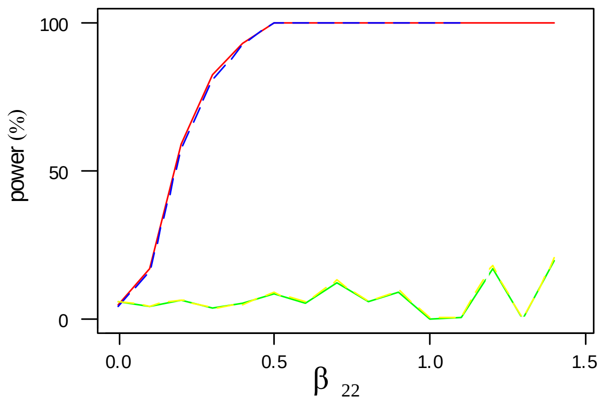
รูปที่ 17. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise** ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



รูปที่ 18. กราฟแสดงความสัมพันธ์ระหว่าง β_{22} กับพลังการทดสอบ(%)ภายใต้ตัวแบบต่าง ๆ ขนาดตัวอย่าง 600



รูปที่ 19. กราฟแสดงความสัมพันธ์ระหว่าง β_{22} กับพลังการทดสอบ(%)ภายใต้ตัวแบบต่าง ๆ ขนาดตัวอย่าง 1000



รูปที่ 20. กราฟแสดงความสัมพันธ์ระหว่าง β_{22} กับพลังการทดสอบ(%)ภายใต้ตัวแบบต่าง ๆ ขนาดตัวอย่าง1500

partition_LR ——— partition_Wald - - - - pairwise_LR - - - - pairwise_Wald - - - -

จากรูปที่ 18-รูปที่ 20 ซึ่งแต่ละกราฟศึกษาที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับ จากกราฟพบว่าพลังการทดสอบ(%)ของตัวสถิติ อัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของแต่ละตัวแบบ ให้พลังการทดสอบ(%)ที่แตกต่างกันเมื่อตัวอย่างมีขนาด 600 และใกล้เคียงกันเมื่อขนาดตัวอย่างเพิ่มเป็น 1000 และ 1500 ตามลำดับ จากกราฟพบว่าพลังการทดสอบ(%)ทั้งตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิตวาลด์ ของตัวแบบ(6) **pairwise** แสดงเส้นกราฟต่ำกว่าตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ของตัวแบบ(5) **partition** อย่างชัดเจนในการศึกษาทุกขนาดตัวอย่าง นอกจากนี้แล้วยังเห็นได้ชัดว่าพลังการทดสอบ(%)ดังกล่าวของตัวแบบ(6) **pairwise**ไม่มีแนวโน้มลู่เข้าสู่ 100% ในทุกขนาดตัวอย่าง

กรณีที่ 3 การเปรียบเทียบพลังการทดสอบของตัวสถิติอัตราส่วนภาวะน่าจะเป็นภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **baseline-partition** เมื่อตัวแบบจริงคือตัวแบบ **pairwise**

กรณีที่ 3. ศึกษาความแตกต่างของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด เมื่อใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับและไม่มีลำดับ โดยตัวแปรตอบสนองที่ต้องการศึกษาเป็นตัวแปรแบบมีลำดับ กำหนดตัวแบบที่ถูกต้องเป็น ตัวแบบ(6) **pairwise** ซึ่งเป็นตัวแบบที่ใช้ในกรณีที่ตัวแปรตอบสนองแบบมีลำดับ

ในการศึกษาพลังการทดสอบ(%)ในกรณีที่ใช้ตัวแบบสำหรับตัวตอบสนองแบบมีลำดับนั้น ใช้ตัวแบบตอมสมมติฐานว่างเป็นตัวแบบ(4) **minimal** และตัวแบบตามสมมติฐานทางเลือกเป็นตัวแบบ(5) **partition**

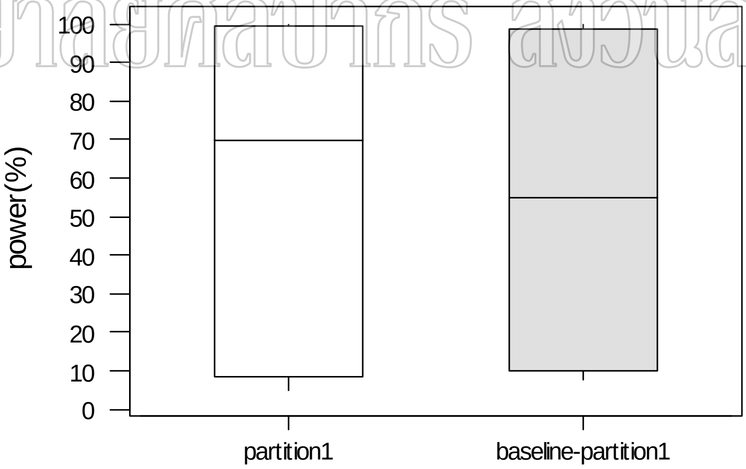
ในการศึกษาพลังการทดสอบ(%)ในกรณีที่ใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบไม่มีลำดับนั้นใช้ตัวแบบตามสมมติฐานว่างเป็นตัวแบบ(8) **baseline category** และตัวแบบทางเลือกคือตัวแบบ(9) **baseline-partition**

ตารางที่ 3 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด(LR) ภายใต้ตัวแบบ(5) **partition** และตัวแบบ(9) **baseline-partition** เมื่อเพิ่มค่า β_{12} กรณีตัวแบบที่ถูกต้องเป็นตัวแบบ(6) **pairwise** จำแนกตามขนาดตัวอย่าง

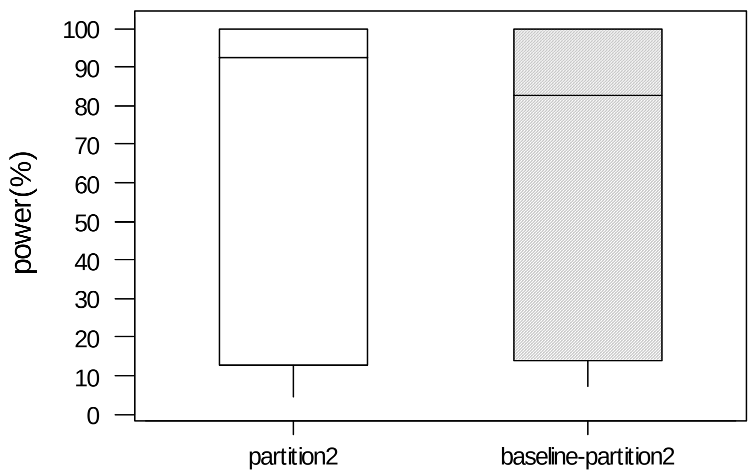
β_{12}	พลังการทดสอบ					
	ขนาดตัวอย่าง 600		ขนาดตัวอย่าง 1000		ขนาดตัวอย่าง 1500	
	Ordinal	Nominal	Ordinal	nominal	Ordinal	nominal
0.0	4.8	7.6	6.0	7.2	4.6	7.2
0.1	6.6	8.0	4.4	7.0	5.0	8.2
0.2	7.8	7.4	8.8	9.8	9.0	10.2
0.3	8.2	9.8	12.6	13.6	17.6	16.0
0.4	16.2	14.6	28.2	19.4	40.6	31.8
0.5	30.0	21.6	49.4	34.6	67.0	54.4
0.6	46.8	30.8	73.0	59.2	90.8	80.4
0.7	70.0	55.0	92.6	83.0	97.0	95.2
0.8	89.4	75.8	98.6	95.2	99.8	99.2
0.9	96.0	88.6	99.6	98.0	100.0	100.0
1.0	99.4	97.8	100.0	99.8	100.0	100.0
1.1	99.8	99.0	100.0	100.0	100.0	100.0
1.2	100.0	100.0	100.0	100.0	100.0	100.0
1.3	100.0	99.8	100.0	100.0	100.0	100.0
1.4	100.0	99.8	100.0	100.0	100.0	100.0
Mean	58.33	54.37	64.88	61.79	68.76	66.84
S.D.	41.79	40.94	41.35	41.15	40.77	40.36

จากตารางที่ 3. ผลการศึกษาพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุดของตัวแบบตามสมมติฐานทางเลือก(5) **partition** และ ตัวแบบตามสมมติฐานทางเลือก(9) **baseline-partition** ตารางที่ 3 พบว่าเมื่อกำหนด β_{12} =0 พลังการทดสอบ(%)ที่ได้จากตัวแบบ(5) **partition** มีค่า 4.8, 6.0 และ 4.6 ที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับ ซึ่งใกล้เคียงระดับนัยสำคัญ(5%) ในขณะที่พลังการทดสอบ(%)จากตัวแบบ(9) **baseline-partition** มีค่า 7.6 ,7.2 และ 7.2 ที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับซึ่งสูงกว่าระดับนัยสำคัญอย่างชัดเจน

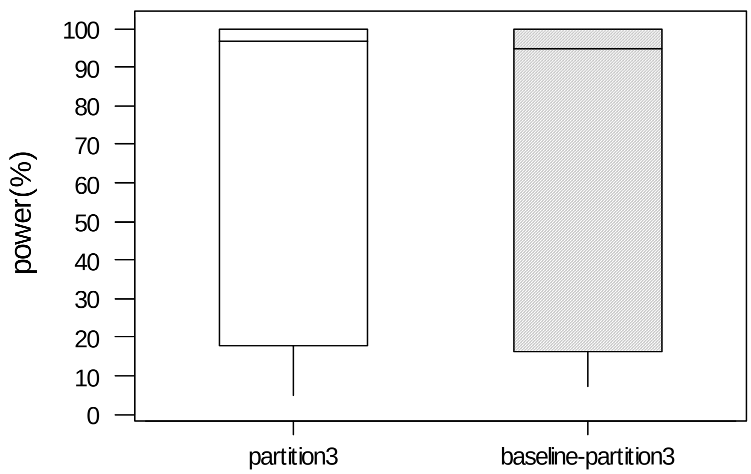
สรุปทั้ง 3 ตัวแบบพบว่าตัวแบบ(5) **partition** ให้ค่าเฉลี่ย (ภายใต้ค่าพารามิเตอร์ 0-1.4) ของพลังการทดสอบ(%)มากกว่าตัวแบบ(9) **baseline-partition** ที่ใช้ตัวแปรตอบสนองเชิงกลุ่มแบบไม่มีลำดับอย่างชัดเจน ภายใต้ตัวแบบที่ถูกต้อง(6) **pairwise** โดยพลังการทดสอบของทั้ง 2 ตัวแบบเพิ่มขึ้นจนถึง 100% เมื่อขนาดตัวอย่างเพิ่มสูงขึ้น แต่ของตัวแบบ(9) **baseline-partition** เพิ่มขึ้นได้ช้ากว่าและน้อยกว่า (รายละเอียดเพิ่มเติมแสดงไว้ใน รูปที่ 21-รูปที่ 28)



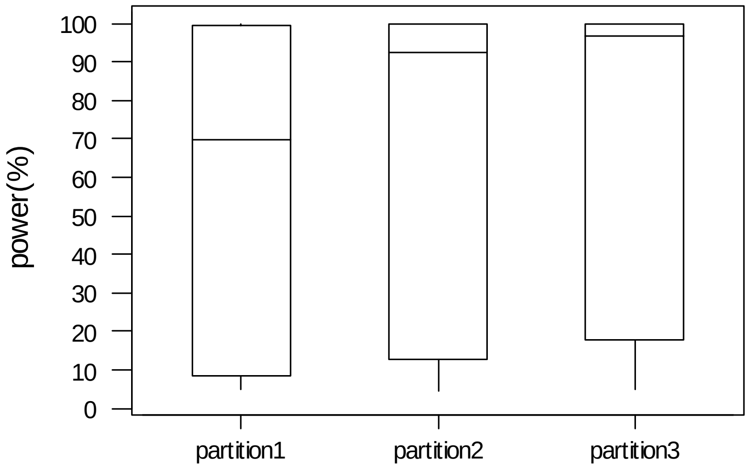
รูปที่ 21. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(9) **baseline-partition** เมื่อตัวแบบที่ถูกต้องคือ (6) **pairwise** และ n = 600



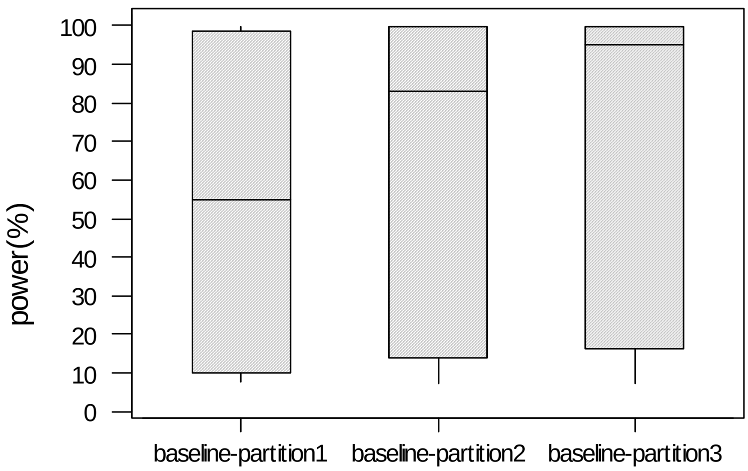
รูปที่ 22. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(9) **baseline-partition** เมื่อตัวแบบที่ถูกต้องคือ (6) **pairwise** และ n = 1000



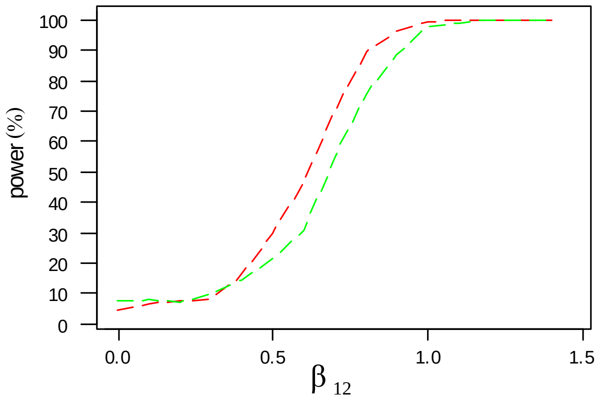
รูปที่ 23. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(9) **baseline-partition** เมื่อตัวแบบที่ถูกต้องคือ (6) **pairwise** และ n = 1500



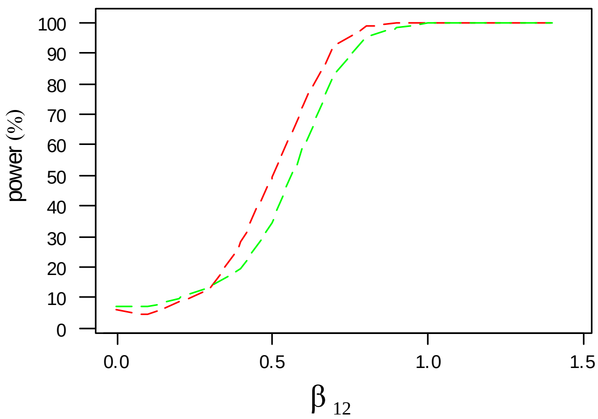
รูปที่ 24. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(5) **partition** ภายใต้ขนาดตัวอย่าง 600, 1000และ 15000



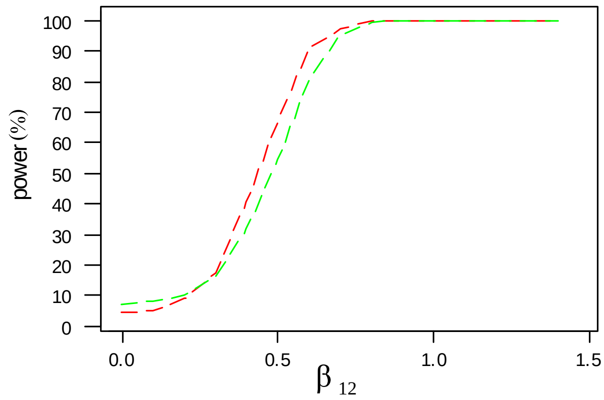
รูปที่ 25. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(9) **baseline-partition** ภายใต้ขนาดตัวอย่าง 600, 1000และ 15000



รูปที่ 26. กราฟแสดงความสัมพันธ์ระหว่าง β_{12} กับพลังการทดสอบ(%) ขนาดตัวอย่าง 600



รูปที่ 27. กราฟแสดงความสัมพันธ์ระหว่าง β_{12} กับพลังการทดสอบ(%) ขนาดตัวอย่าง 1000



รูปที่ 28. กราฟแสดงความสัมพันธ์ระหว่าง β_{12} กับพลังการทดสอบ(%) ขนาดตัวอย่าง 1500

partition ——— baseline-partition - - - -

จากรูปที่ 26-รูปที่ 28 พบว่า เมื่อ β_{12} น้อยกว่า 0.3 พลังการทดสอบ(%)จากตัวแบบ(9) **baseline-partition** มีค่าสูงกว่าพลังการทดสอบ(%)จากตัวแบบ(5) **partition** แต่เมื่อ β_{12} มากกว่า 0.3 พบว่าพลังการทดสอบ(%)จากตัวแบบ(5) **partition** มีค่าสูงกว่าพลังการทดสอบ(%)จากตัวแบบ(9) **baseline-partition** อย่างไรก็ตามความแตกต่างดังกล่าวนี้ลดน้อยลงเมื่อขนาดตัวอย่างเพิ่มขึ้น และพลังการทดสอบ(%)จากตัวแบบทั้งคู่ลู่เข้าสู่ 100% เมื่อ β_{12} มากกว่า 1.1, 1.0 และ 0.8 ที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับ

กรณีที่ 4 การเปรียบเทียบพลังการทดสอบของตัวสถิติอัตราส่วนภาวะน่าจะเป็นภายใต้ตัวแบบตามสมมติฐานว่างตัวแบบ **partition** และตัวแบบ **baseline-partition** เมื่อตัวแบบจริงคือตัวแบบ **quadratic**

กรณีที่ 4. เป็นศึกษาความแตกต่างของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด เมื่อใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับและไม่มีลำดับ โดยตัวแปรตอบสนองที่ต้องการศึกษาเป็นตัวแปรแบบมีลำดับ กำหนดตัวแบบที่ถูกต้องเป็น ตัวแบบ(7) **quadratic** ซึ่งเป็นตัวแบบที่ใช้ในกรณีที่ตัวแปรตอบสนองแบบมีลำดับ

ในการศึกษาพลังการทดสอบ(%)ในกรณีที่ใช้ตัวแบบสำหรับตัวตอบสนองแบบมีลำดับนั้น ใช้ตัวแบบตอมสมมติฐานว่างเป็นตัวแบบ(4) **minimal** และตัวแบบตามสมมติฐานทางเลือกเป็นตัวแบบ (5) **partition**

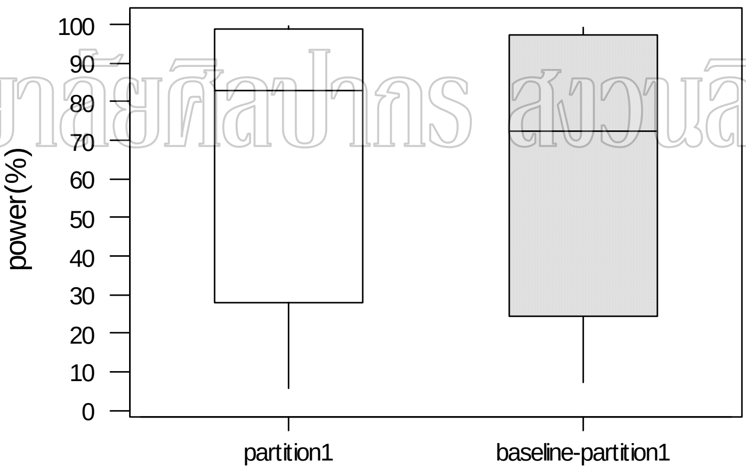
ในการศึกษาพลังการทดสอบ(%)ในกรณีที่ใช้ตัวแบบสำหรับตัวแปรตอบสนองแบบไม่มีลำดับนั้นใช้ตัวแบบตามสมมติฐานว่างเป็นตัวแบบ(8) **baseline-category** และตัวแบบทางเลือกคือตัวแบบ(9) **baseline-partition**

ตารางที่ 4 พลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด(LR ภายใต้ตัวแบบ(5) **partition** และตัวแบบ(9) **baseline-partition** เมื่อเพิ่มค่า β_{22} กรณีตัวแบบที่ถูกต้องคือตัวแบบ(7) **quadratic** จำแนกตามขนาดตัวอย่าง

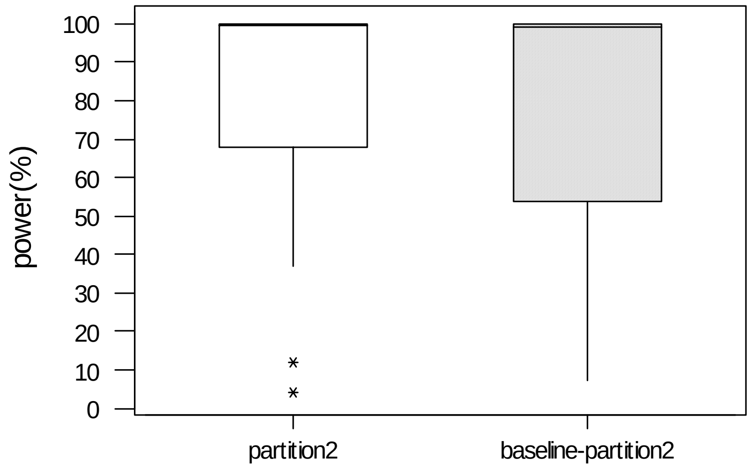
β_{22}	พลังการทดสอบ					
	ขนาดตัวอย่าง 600		ขนาดตัวอย่าง 1000		ขนาดตัวอย่าง 1500	
	ordinal	Nominal	ordinal	Nominal	Ordinal	nominal
0.0	5.4	7.2	3.8	7.2	5.6	7.6
0.1	6.8	7.0	11.6	12.0	16.0	13.8
0.2	14.4	12.8	37.0	26.2	50.8	39.8
0.3	28.0	24.4	67.8	53.8	85.4	68.8
0.4	51.0	38.4	88.2	77.6	97.0	91.4
0.5	67.4	51.4	96.2	91.4	99.2	98.8
0.6	75.0	61.6	99.6	98.4	100.0	99.6
0.7	83.2	72.4	99.8	99.4	100.0	100.0
0.8	91.8	86.0	100.0	100.0	100.0	100.0
0.9	93.2	89.6	100.0	100.0	100.0	100.0
1.0	96.8	93.8	100.0	100.0	100.0	100.0
1.1	99.0	97.4	100.0	100.0	100.0	100.0
1.2	99.6	98.2	100.0	100.0	100.0	100.0
1.3	99.4	98.8	100.0	100.0	100.0	100.0
1.4	99.8	99.4	100.0	100.0	100.0	100.0
Mean	67.39	62.56	80.27	77.73	83.60	81.32
S.D.	36.54	36.16	34.21	34.93	32.27	33.17

จากตารางที่ 4. ผลการศึกษาพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุดของตัวแบบตามสมมติฐานทางเลือก(5) **partition** และ ตัวแบบตามสมมติฐานทางเลือก(9) **baseline-partition** ตารางที่ 4 พบว่าเมื่อกำหนด β_{22} =0 พบพลังการทดสอบ(%)ที่ได้จากตัวแบบ(5) มีค่า 5.4, 3.8 และ 5.6 ที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับ ซึ่งใกล้เคียงระดับนัยสำคัญ(5%) ในขณะที่พลังการทดสอบ(%)จากตัวแบบ(9) มีค่า 7.2 ,7.2 และ 7.2 ที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับซึ่งสูงกว่าระดับนัยสำคัญอย่างชัดเจน

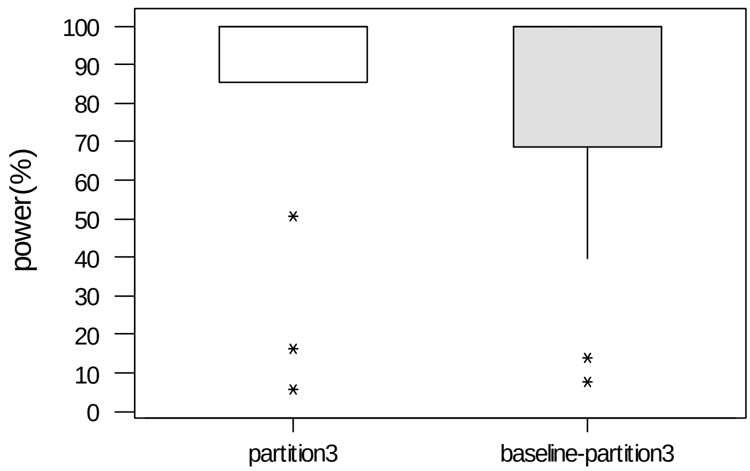
สรุปผลจากตารางที่ 4. ได้ทำนองเดียวกับผลลัพธ์ของตารางที่ 3. โดยทั้ง 3 ขนาดตัวอย่างพบว่า ตัวแบบ(5) **partition** ให้ค่าเฉลี่ย (ภายใต้ค่าพารามิเตอร์ 0-1.4) ของพลังการทดสอบ(%)มากกว่าตัวแบบ(9) **baseline-partition** อย่างชัดเจนเช่นกัน ภายใต้ตัวแบบที่ถูกต้อง(7) **quadratic** (รายละเอียดเพิ่มเติมแสดงไว้ใน รูปที่ 29-รูปที่ 36)



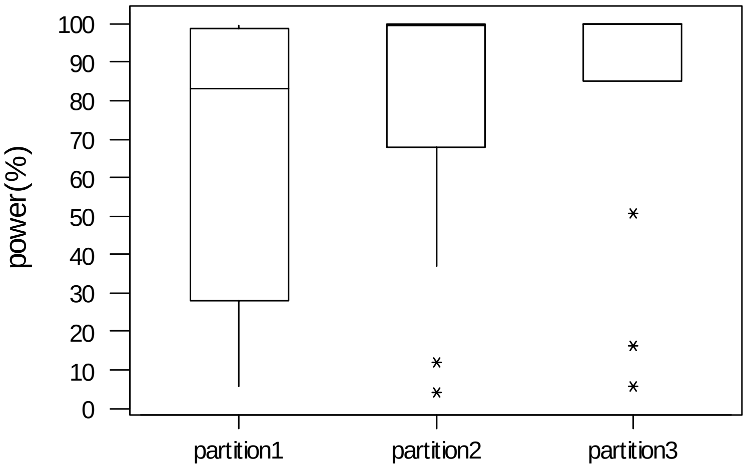
รูปที่ 29. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(9) **baseline-partition** เมื่อตัวแบบที่ถูกต้องคือ (7) **quadratic** และ n = 600



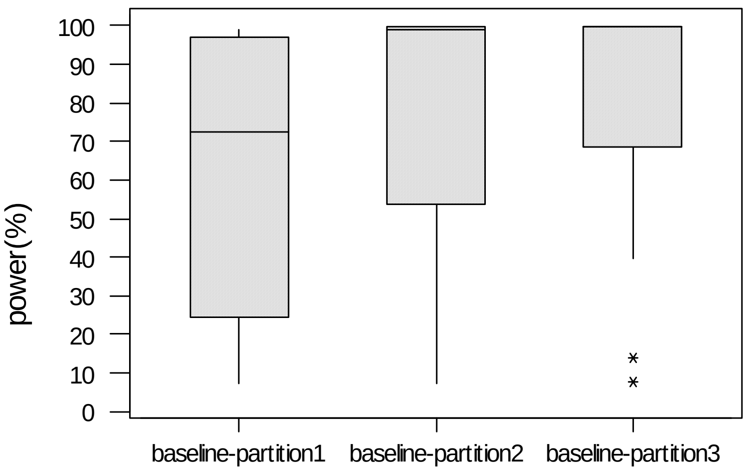
รูปที่ 30. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(9) **baseline-partition** เมื่อตัวแบบที่ถูกต้องคือ (7) **quadratic** และ n = 1000



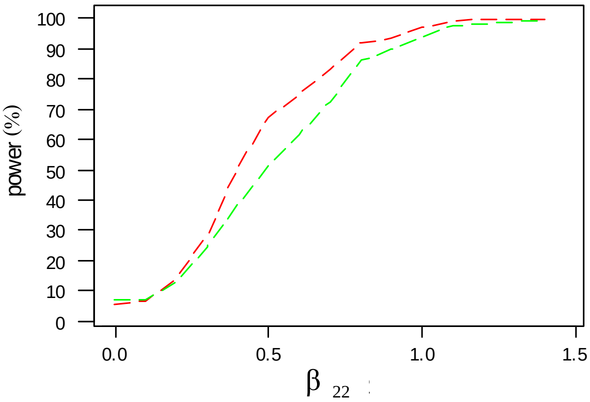
รูปที่ 31. Boxplot ของพลังการทดสอบ(%)ภายใต้ตัวแบบ(5) **partition** และ(9) **baseline-partition** เมื่อตัวแบบที่ถูกต้องคือ (7) **quadratic** และ n = 1500



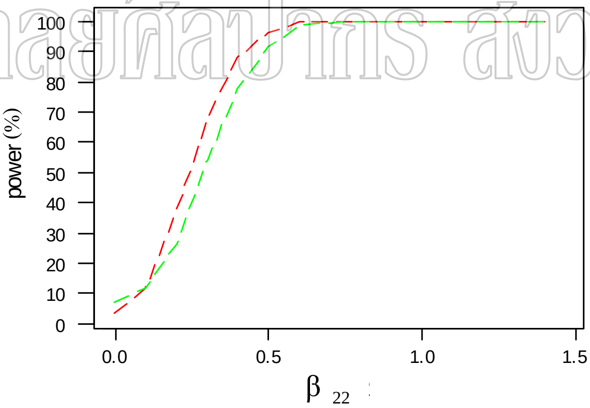
รูปที่ 32. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(5) **partition** ภายใต้ขนาดตัวอย่าง 600, 1000และ 1500



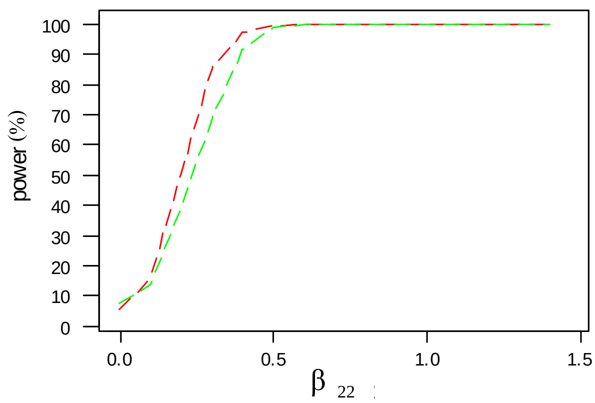
รูปที่ 33. Boxplot ของพลังการทดสอบ(%)ของตัวสถิติอัตราส่วนภาวะน่าจะเป็นจากตัวแบบ(9) **baseline-partition** ภายใต้ขนาดตัวอย่าง 600, 1000และ 15000



รูปที่ 34. กราฟแสดงความสัมพันธ์ระหว่าง β_{22} กับพลังการทดสอบ(%) ขนาดตัวอย่าง 600



รูปที่ 35. กราฟแสดงความสัมพันธ์ระหว่าง β_{22} กับพลังการทดสอบ(%) ขนาดตัวอย่าง 1000



รูปที่ 36. กราฟแสดงความสัมพันธ์ระหว่าง β_{22} กับพลังการทดสอบ(%) ขนาดตัวอย่าง 1500

partition ——— baseline-partition – – –

จากรูปที่ 34-36 พบว่า เมื่อ β_{22} น้อยกว่า 0.2 พลังการทดสอบ(%)จากตัวแบบ(9) **baseline-partition** มีค่าสูงกว่าพลังการทดสอบ(%)จากตัวแบบ(5) แต่เมื่อ β_{22} มากกว่า 0.2 พบว่าพลังการทดสอบ(%)จากตัวแบบ(5) **partition** มีค่าสูงกว่าพลังการทดสอบ(%)จากตัวแบบ(9) **baseline-partition** อย่างไรก็ตามความแตกต่างดังกล่าวนี้ลดน้อยลงเมื่อขนาดตัวอย่างเพิ่มขึ้น และพลังการทดสอบ(%)จากตัวแบบทั้งคู่ลู่เข้าสู่ 100% เมื่อ β_{22} มากกว่า 1.4, 0.7 และ 0.6 ที่ขนาดตัวอย่าง 600, 1000 และ 1500 ตามลำดับ

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

สรุปและอภิปรายผลการวิจัย

การเปรียบเทียบพลังการทดสอบของการวัดภาวะสารูปดีเมื่อกำหนดตัวแบบไม่ถูกต้อง สำหรับตัวแปรตอบสนองเชิงกลุ่มแบบมีลำดับ โดยใช้ตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ ภายใต้ตัวแบบ **minimal** และตัวแบบทางเลือกต่าง ๆ ที่อยู่ในรูปแบบของตัวแบบเชิงเส้นที่วางนัยทั่วไปหลายตัวแบบ ข้อมูลของการวิจัยใช้การจำลองแบบภายใต้เงื่อนไขของพารามิเตอร์และลักษณะการแจกแจงของตัวแปรอธิบาย การประมวลผลในแต่ละเงื่อนไขทำซ้ำ 500 ครั้ง

จากผลการศึกษาพลังการทดสอบของตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ ภายใต้ตัวแบบทางเลือกตัวแบบ(5) **partition** และตัวแบบ(6) **pairwise** ขนาดตัวอย่าง 600, 1000 และ 1500 พบว่าในการศึกษาการทดสอบภาวะสารูปดีของตัวแบบ(4) **minimal** กรณีที่ตัวแบบที่ถูกต้องเป็นตัวแบบ (6) **pairwise** ซึ่งเป็นตัวแบบที่มีเทอมผลคูณของตัวแปร X_1 และตัวแปร X_2 การใช้ตัวแบบทางเลือก (6) **pairwise** ซึ่งคือตัวแบบที่ถูกต้องนั่นเอง กรณีนี้พลังการทดสอบ(%)สูงกว่ากรณีที่ใช้ตัวแบบตามสมมติฐานทางเลือกเป็นตัวแบบ(5) **partition** ในทุกขนาดตัวอย่าง แต่ในกรณีที่กำหนดตัวแบบที่ถูกต้องเป็นตัวแบบ (7) **quadratic** ซึ่งเป็นตัวแบบที่มีเทอมควอดราติก ของตัวแปร X_2 กรณีนี้พบว่าการใช้ตัวแบบตามสมมติฐานทางเลือก (5) **partition** ในการทดสอบภาวะสารูปดีของตัวแบบ (4) **minimal** นั้นพลังการทดสอบ(%)สูงกว่าในกรณีที่ใช้ตัวแบบตามสมมติฐานทางเลือกเป็นตัวแบบ (6) **pairwise** อยู่มาก และพบว่าเมื่อขนาดตัวอย่างมากขึ้นการเพิ่มค่าสัมประสิทธิของตัวแปรอธิบายเพียงเล็กน้อย ทำให้พลังการทดสอบ(%)เพิ่มขึ้นอย่างรวดเร็ว จากผลที่ได้ดังกล่าวนี้แสดงให้เห็นว่าตัวแบบตามสมมติฐานทางเลือก (5) **partition** เป็นตัวแบบที่เหมาะสมสำหรับใช้ในการทดสอบภาวะสารูปดีของตัวแบบต่าง ๆ เนื่องจากในทางปฏิบัติเราไม่ทราบว่าตัวแบบที่ถูกต้องหรือตัวแบบจริงนั้นเป็นตัวแบบใด จึงเหมาะที่จะใช้ตัวประเมินหรือตัวแบบทางเลือกที่มีความเป็นทั่วไป มากกว่าการใช้ตัวแบบที่มีความเฉพาะเจาะจงในการทดสอบภาวะสารูปดีบางตัวแบบเพียงเท่านั้น ในงานวิจัยนี้ได้ศึกษาโดยใช้ตัวสถิติอัตราส่วนภาวะน่าจะเป็นและตัวสถิติวาลด์ จากผลการศึกษาพบว่าตัวสถิติทั้งสองตัวมีพลังการทดสอบ(%)แตกต่างกันเล็กน้อยเมื่อตัวอย่างมีขนาดใหญ่ แต่ตัวสถิติวาลด์นั้นมักจะมีปัญหาในการคำนวณเนื่องจากในการคำนวณค่าสถิติดังกล่าวจำเป็นจะต้องใช้ค่าประมาณของพารามิเตอร์ที่

ต้องการจะทดสอบในตัวแบบ ซึ่งในกรณีที่ใช้ตัวแบบที่มีจำนวนพารามิเตอร์มาก ๆ เช่นตัวแบบ(5) **partition** ในบางแบบจำลองข้อมูลค่าประมาณพารามิเตอร์ของตัวแบบลู่ออก($|\hat{\beta}| = \infty$) จึงเป็นเหตุให้ไม่สามารถคำนวณค่าสถิติวาลด์ได้ ปัญหาดังกล่าวนี้ไม่ส่งผลกระทบในการคำนวณค่าสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด (Agresti 2002) เพราะพิจารณาการลู่เข้าของผลต่างระหว่างตัวประมาณพารามิเตอร์ของตัวแบบ เมื่อ $\hat{\beta}$ ลู่ออกยังสามารถนำค่าประมาณที่พอเพียง (sufficient statistic) ต่าง ๆ มาใช้ในเทอมของ log-likelihood function ในสูตรของ Deviance ต่อไปได้

ในการศึกษาพลังการทดสอบของตัวสถิติอัตราส่วนภาวะน่าจะเป็นภายใต้ตัวแบบตามสมมติฐานทางเลือกตัวแบบ(5) **partition** และตัวแบบ(9) **baseline-partition** ขนาดตัวอย่าง 600, 1000 และ 1500 พบว่ากรณีที่ตัวแปรตอบสนองเป็นตัวแปรแบบมีลำดับ ซึ่งในทางปฏิบัติควรใช้ตัวแบบ **proportional odds models** แต่กลับใช้ตัวแบบสำหรับข้อมูลที่ตัวแปรตอบสนองเป็นแบบเชิงกลุ่ม (**baseline-category logit model**) จากผลของการวิจัยพบว่านอกจากจะสูญเสียสารสนเทศของข้อมูลแล้วยังพบว่าสูญเสียพลังการทดสอบ(%)อีกด้วย แสดงดังรูป ที่ 21-รูปที่ 36

ข้อเสนอแนะจากงานวิจัย

เนื่องจากการประเมินภาวะสารูปดีของตัวแบบสำหรับตัวแปรตอบสนองแบบมีลำดับ เมื่อกำหนดตัวแบบไม่ถูกต้องโดยใช้ตัวสถิติ LR และ ตัวสถิติ Wald พบว่าถึงแม้ตัวแบบที่รวมอิทธิพลการแบ่งกลุ่ม(partition) นั้นจะมีพลังการทดสอบ(%)ลดลงแต่ยังให้พลังการทดสอบ(%)ค่อนข้างเสถียร(stable) และยังมีพลังการทดสอบสูงขึ้นเมื่อใช้ร่วมกับตัวสถิติอัตราส่วนภาวะน่าจะเป็นสูงสุด(LR) ดังนั้นในการทดสอบภาวะสารูปดีของตัวแบบควรเลือกใช้ตัวแบบที่สอดคล้องกับประเภทของข้อมูล และในการนำตัวแบบที่รวมเทอม partitionไปใช้เราสามารถเพิ่มเติมตัวแปรอธิบายที่สนใจตรวจสอบอิทธิพลที่มีต่อตัวแปรตอบสนองไว้ในตัวแบบตามสมมติฐานว่างได้

ข้อเสนอแนะในการวิจัยครั้งต่อไป

เนื่องจากงานวิจัยนี้ทำการตรวจสอบภาวะสารูปดีโดยศึกษาพลังการทดสอบ(%)ของตัวแบบ GLMs ที่มีตัวแปรตอบสนองเชิงกลุ่มแบบมีลำดับจำนวนหลายตัวแบบ อย่างไรก็ตามยังมีตัวแบบสำหรับตัวแปรตอบสนองเชิงกลุ่มแบบมีลำดับที่น่าสนใจศึกษาเพิ่มเติม เช่น ตัวแบบ Generalized cumulative logit models และตัวแบบ Adjacent-category logit models ฯลฯ นอกจากนี้ถ้าตัวแบบมีความซับซ้อนมากขึ้นควรใช้จำนวน computer simulation runs และขนาดตัวอย่างมากขึ้นด้วย

จากการวิจัยชี้ให้เห็นว่าตัวแบบ(5) **partition** เป็นตัวแบบที่ดีในการทดสอบภาวะสารูปดีของตัวแบบ(4) **minimal** ซึ่งการทดสอบภาวะสารูปดีวิธีนี้มีพื้นฐานมาจากการแบ่งส่วนของค่าสังเกตออกเป็นกลุ่ม ๆ โดยให้ภายในแต่ละกลุ่มนั้นมีค่าพยากรณ์ความน่าจะเป็นของตัวแปรตอบสนองที่ได้จากการสร้างตัวแบบ(4) **minimal** มีค่าใกล้เคียงกัน ดังนั้นถ้าเราสามารถแบ่งกลุ่มค่าสังเกตให้แต่ละกลุ่มมีค่าพยากรณ์ของตัวแปรตอบสนองใกล้เคียงกันมากที่สุดโดยอาจใช้วิธีการ Cluster Analysis กับค่าพยากรณ์ของตัวแปรตอบสนอง ในการแบ่งกลุ่มค่าสังเกตดังกล่าว ซึ่งไม่จำเป็นต้องใช้ G =10 เพื่อให้การทดสอบภาวะสารูปดีมีประสิทธิภาพมากขึ้น

บรรณานุกรม

ภาษาไทย

วีรานันท์ พงศาภักดี. การวิเคราะห์ข้อมูลเชิงกลุ่ม. นครปฐม : โรงพิมพ์มหาวิทยาลัยศิลปากร, 2541.

ภาษาต่างประเทศ

Agresti, A. Analysis of Ordinal Categorical Data. New York : Wiley, 1984.

________. Categorical Data Analysis. New York : Wiley, 2002.

Clogg, C.C., and Shihadeh, E.S. Statistical Models for Ordinal variables. London : Sage Publication, 1994.

Dobson, A. An Introduction to Statistical Modelling. New York : Chapman Hall, 1983.

Fienberg, S.E. and Mason, W. M. "Identification and estimation of age-period-cohort models in the analysis of discrete archival data." Sociol. Methodol 1979 : 1-67.

Haberman, S.J. "Log-linear models for Frequency tables with ordered classification." Biometrics 30(1974) : 589-600.

Hosmer, D.W., and Lemeshow, S. "A goodness-of-fit test for the multiple logistic regression model." Communs Statis. Theory Meth 10(1980) : 1043-1069.

________. Applied logistic Regressions. New York : Wiley, 1989.

Hosmer, D. W., Lemeshow, S., and Klar, J. "Goodness-of-fit testing for multiple logistic regression analysis when the estimated probabilities are small." Biometrical Journal 30(1988) : 911-924.

Lemeshow, S., and Hosmer, D. W. "The use of goodness-of –fit statistics in the development of logistic regression models." American Journal of Epidemiology 115(1982) : 92-106.

McCullagh, P. "Regression models for ordinal data (with discussion)." J. R. Statist. Soc. B 42(1980) : 109-142.

McCullagh, P., and Nelder, J.A. Generalized Linear Models. London: Chapman Hall, 1983.

McFadden, D. Conditional logit analysis of quantitative choice behavior. New York : Academic Press, 1974 : 105-142.

Moore, D. S., and Spruill, M. C. "Unified large-sample theory of general chi-squared statistics for tests of fit." Ann. Statist 3(1975) : 599-616.

Lipsitz, S., Fitzmaurice, G. M., and Molenberghs G. “Goodness-of-fit Test for Ordinal Response Regression Models.” Applied Statistics 45(1996) : 175-190.

Tsiatis, A. A. “A note on the goodness of fit the logistic regression model.” Biometrika 67(1990) : 250-251.

Williams, O. D., and Grizzle, J. E. “Analysis of contingency tables having ordered response catetories.” J. Am. Statist. Ass. 67(1972) : 55-63.

ภาคผนวก

ภาคผนวก

โปรแกรมที่ใช้ในงานวิจัย

ตอนที่ 1

1. ความหมายของตัวแปรในโปรแกรม

MCONSTANT

N	ขนาดตัวอย่าง
B01	ค่าพารามิเตอร์ α_1 ของตัวแบบที่ถูกต้อง
B02	ค่าพารามิเตอร์ α_2 ของตัวแบบที่ถูกต้อง
B1	ค่าพารามิเตอร์ β_1 ของตัวแบบที่ถูกต้อง
B2	ค่าพารามิเตอร์ β_2 ของตัวแบบที่ถูกต้อง
B12	ค่าพารามิเตอร์ β_{12} ของตัวแบบที่ถูกต้อง
LOGL1	ค่า log-likelihood จากตัวแบบ(4) **minimal**
LOGL2	ค่า log-likelihood จากตัวแบบ(5) **partition**
LOGL3	ค่า log-likelihood จากตัวแบบ(6) **pairwise**
PV1	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
PV2	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(6) **pairwise**
PVWAL1	P-value ของตัวสถิติวาลด์จากตัวแบบ(5) **partition**
PVWAL2	P-value ของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise**
POWER1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
POWER2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(6) **pairwise**
PWAL1	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(5) **partition**
PWAL2	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise**

MCOLUMN

B12C	ค่า β_{12}
POWERC1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**

POWERC2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(6) **pairwise**
PWALC1	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(5) **partition**
PWALC2	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise**

2. ขั้นตอนการทำงานของโปรแกรม

2.1 จำลองแบบข้อมูลตามตัวแบบที่ถูกต้องตัวแบบ(6) **pairwise** กำหนด $\beta_{12} = 0$

2.2 นำข้อมูลที่ได้จากข้อ 2.1 ฟิตตัวแบบ(4) **minimal** จากนั้นเรียงข้อมูลใหม่โดยเรียงตามค่าประมาณความน่าจะเป็นของตัวแปรตอบสนองระดับที่ 1 ($\hat{p}_1$)

2.3 นำข้อมูลที่ได้จากข้อ 2.2 ฟิตตัวแบบ(5) **partition** และตัวแบบ(6) **pairwise**

2.4 คำนวณค่าสถิติอัตราส่วนความน่าจะเป็นสูงสุดและค่าสถิติวาลด์ที่ได้จากตัวแบบในข้อ 2.3

2.5 ทำการทดสอบสมมติฐานว่างตัวแบบ(4) **minimal**

2.6 ทำซ้ำตั้งแต่ข้อ 2.1-2.5 จำนวน 500 ครั้ง แล้วนับจำนวนครั้งที่ปฏิเสธสมมติฐานว่างจากตัวสถิติภายใต้เงื่อนไขต่าง ๆ

2.7 พลังการทดสอบคำนวณจากจำนวนครั้งที่ปฏิเสธสมมติฐานว่างจากข้อ 2.6 หารด้วย 5

2.8 ทำซ้ำตั้งแต่ข้อ 2.1-2.7 โดยเพิ่มค่า β_{12} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

3. แสดงรายละเอียดของโปรแกรม

```
MACRO
STATISTIC  n  B01 B02  B1  B2  B12C POWERC1 POWERC2 PWALC1 PWALC2
MCOLUMN  X1  X2  P  P1  P2  P3  U  Y A1 A2 A3 A4 A5 A6 A7 A8 A9 A10
           COEF1 COEF2 CO1 CO2
MCOLUMN    P02   EPRO1  EPRO2  EPRO3  INDICAT  B12C  POWERC1
MCOLUMN   POWERC2 PWALC1 PWALC2
MCONSTANT B01 B02  B1  B2 B12  J  n  LOGL1   LOGL2  LOGL3  I  K1  K2
           PV1  PV2  Q
MCONSTANT   DIV1  DIV2 DIV3  D1 D2 CUT2  POWER1 POWER2  REJ1  REJ2
           G WALD1 WALD2
MCONSTANT  KW1 KW2 PVWAL1 PVWAL2 REJW1 REJW2  PWAL1 PWAL2
           CON O
MMATRIX       BETA1 BETA2 BETA1T BETA2T XPWX1 XPWX2  XPWX1V
           XPWX2V MUL1 MUL2
MMATRIX  WAL1  WAL2
LET CON = N/2
LET B12 = 0
LET G = N/10
```

```
 SET INDICAT
 1( 1 2 3 4 5 6 7 8 9 10 )G
END.
          Indicator INDICAT   A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 .

DO Q = 1:15
         LET POWER1 = 0
         LET POWER2 = 0
         LET PWAL1 = 0
         LET PWAL2 = 0
         LET O = CON
         WHILE O = CON
         RANDOM n X1;
                  Bernulli .5.
          Random n X2;
                  Normal 0 1.
         SUM X1 O
         ENDWHILE
LET  P1 = EXPO(B01+(B2*X2)+(B1*X1)  +(B12*X1*X2)   ) /
(1+EXPO(B01+(B2*X2)+(B1*X1) +(B12*X1*X2)  ))
LET P02 = EXPO(B02+(B2*X2)+(B1*X1) +(B12*X1*X2)   ) /
(1+EXPO(B02+(B2*X2)+(B1*X1) +(B12*X1*X2)  ))
                           LET P2 = P02 - P1
                           LET P3 = 1-P1-P2
                           LET  P = P1 + P2 + P3
                  DO I = 1:500
                  Random n U;
                           Uniform 0 1.
                  DO J=1:n
                  LET CUT2 = P1(J) + P2(J)
                  IF U(J) <  P1(J)
                           LET Y(J) = 0
                  ELSE  IF U(J) >= CUT2
                            LET  Y(J) = 2
                  ELSE
                           LET  Y(J) = 1
                  ENDIF
                  ENDDO
         OLogistic Y = X1   X2 ;
         Logit;
          Eprobability EPRO1  EPRO2  EPRO3;
         Loglikelihood LOGL1;
         Brief 2.

SORT  X1  X2  Y  P1  P2  P3  P02  EPRO1  X1  X2  Y P1  P2  P3  P02  EPRO1;
         BY EPRO1 .
         IF B12 < 0.8
```

```
OLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
Logit;
 Coefficients COEF1;
 XPWXinverse XPWX1;
Loglikelihood LOGL2;
Brief 2.
table INDICAT Y;
COUNTS.
Copy COEF1  BETA1;
omit 1:4.
TRANSPOSE BETA1  BETA1T.
COPY XPWX1 XPWX1;
OMIT 1:4.
TRANSPOSE XPWX1 XPWX1.
COPY XPWX1 XPWX1;
OMIT 1:4.
INVERT XPWX1 XPWX1V.
MULTIPLY BETA1T XPWX1V MUL1.
MULTIPLY MUL1 BETA1  WAL1.
COPY WAL1  CO1.
LET WALD1 = CO1  ## DF 9  ##
OLogistic Y = X1   X2  X1*X2 ;
Logit;
 Coefficients COEF2;
 XPWXinverse XPWX2;
Loglikelihood LOGL3;
Brief 2.
Copy COEF2  BETA2;
omit 1:4.
TRANSPOSE BETA2  BETA2T.
COPY XPWX2 XPWX2;
OMIT 1:4.
TRANSPOSE XPWX2 XPWX2.
COPY XPWX2 XPWX2;
OMIT 1:4.
INVERT XPWX2 XPWX2V.
MULTIPLY BETA2T XPWX2V MUL2.
MULTIPLY MUL2 BETA2  WAL2.
COPY WAL2  CO2.
LET WALD2 = CO2  ##  DF 1 ####
LET DIV1 = (-2)*LOGL1
LET DIV2 = (-2)*LOGL2
LET DIV3 = (-2)*LOGL3
LET D1 = DIV1-DIV2
LET D2 = DIV1-DIV3
 CDF D1  K1;
 Chisquare 9.
```

```
 CDF D2  K2;
 Chisquare 1.
CDF WALD1  KW1;
 Chisquare 9.
CDF WALD2  KW2;
 Chisquare 1.
LET PV1 = 1-K1
LET PV2 = 1-K2
LET PVWAL1 = 1-KW1
LET PVWAL2 = 1-KW2
        IF PV1 < .05
                LET  REJ1 = 1
        ELSE
                 LET  REJ1 = 0
        ENDIF
        IF PV2 < .05
                LET  REJ2 = 1
        ELSE
                 LET  REJ2 = 0
        ENDIF
        IF PVWAL1 < .05
                LET  REJW1 = 1
        ELSE
                 LET  REJW1 = 0
        ENDIF
        IF PVWAL2 < .05
                LET  REJW2 = 1
        ELSE
                 LET  REJW2 = 0
        ENDIF
        LET  POWER1 = POWER1 + REJ1
        LET  POWER2 = POWER2 + REJ2
        LET  PWAL1 = PWAL1 + REJW1
        LET  PWAL2 = PWAL2 + REJW2
PRINT POWER1
PRINT POWER2
PRINT PWAL1
PRINT PWAL2
PRINT B12
ENDDO
LET POWER1 = POWER1 / 5
LET POWER2 = POWER2 / 5
LET PWAL1 = PWAL1 / 5
LET PWAL2 = PWAL2 / 5
LET B12C(Q) =  B12
LET POWERC1(Q) = POWER1
LET POWERC2(Q) = POWER2
```

```
        LET PWALC1(Q) = PWAL1
        LET PWALC2(Q) = PWAL2
ELSE
OLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
        Logit;
         Coefficients COEF1;
         XPWXinverse XPWX1;
        Loglikelihood LOGL2;
        Brief 2.
        table INDICAT Y;
        COUNTS.
        OLogistic Y = X1   X2  X1*X2 ;
        Logit;
         Coefficients COEF2;
         XPWXinverse XPWX2;
        Loglikelihood LOGL3;
        Brief 2.
        Copy COEF2  BETA2;
        omit 1:4.
        TRANSPOSE BETA2  BETA2T.
        COPY XPWX2 XPWX2;
        OMIT 1:4.
        TRANSPOSE XPWX2 XPWX2.
        COPY XPWX2 XPWX2;
        OMIT 1:4.
        INVERT XPWX2 XPWX2V.
        MULTIPLY BETA2T XPWX2V MUL2.
        MULTIPLY MUL2 BETA2  WAL2.
        COPY WAL2  CO2.
        LET WALD2 = CO2  ##  DF 1 ####
        LET DIV1 = (-2)*LOGL1
        LET DIV2 = (-2)*LOGL2
        LET DIV3 = (-2)*LOGL3
        LET D1 = DIV1-DIV2
        LET D2 = DIV1-DIV3
         CDF D1  K1;
         Chisquare 9.
         CDF D2  K2;
         Chisquare 1.
        CDF WALD2  KW2;
         Chisquare 1.
        LET PV1 = 1-K1
        LET PV2 = 1-K2
        LET PVWAL2 = 1-KW2
                IF PV1 < .05
                        LET  REJ1 = 1
                ELSE
```

```
                LET  REJ1 = 0
            ENDIF
            IF PV2 < .05
                LET  REJ2 = 1
            ELSE
                LET  REJ2 = 0
            ENDIF
            IF PVWAL2 < .05
                LET  REJW2 = 1
            ELSE
                LET  REJW2 = 0
            ENDIF
            LET  POWER1 = POWER1 + REJ1
            LET  POWER2 = POWER2 + REJ2
            LET  PWAL2 = PWAL2 + REJW2
        PRINT POWER1
        PRINT POWER2
        PRINT PWAL2
        PRINT B12
        ENDDO
        LET POWER1 = POWER1 / 5
        LET POWER2 = POWER2 / 5
        LET PWAL2 = PWAL2 / 5
        LET B12C(Q) =  B12
        LET POWERC1(Q) = POWER1
        LET POWERC2(Q) = POWER2
        LET PWALC2(Q) = PWAL2
ENDIF
        LET B12 = B12 + 0.1

ENDDO
ENDMACRO
```


ตอนที่ 2

1. ความหมายของตัวแปรในโปรแกรม

MCONSTANT

N	ขนาดตัวอย่าง
B01	ค่าพารามิเตอร์ α_1 ของตัวแบบที่ถูกต้อง
B02	ค่าพารามิเตอร์ α_2 ของตัวแบบที่ถูกต้อง
B1	ค่าพารามิเตอร์ β_1 ของตัวแบบที่ถูกต้อง
B2	ค่าพารามิเตอร์ β_2 ของตัวแบบที่ถูกต้อง
B22	ค่าพารามิเตอร์ β_{22} ของตัวแบบที่ถูกต้อง

LOGL1	ค่า log-likelihood จากตัวแบบ(4) **minimal**
LOGL2	ค่า log-likelihood จากตัวแบบ(5) **partition**
LOGL3	ค่า log-likelihood จากตัวแบบ(6) **pairwise**
PV1	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
PV2	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(6) **pairwise**
PVWAL1	P-value ของตัวสถิติวาลด์จากตัวแบบ(5) **partition**
PVWAL2	P-value ของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise**
POWER1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
POWER2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(6) **pairwise**
PWAL1	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(5) **partition**
PWAL2	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise**
MCOLUMN	
B22C	ค่า β_{22}
POWERC1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
POWERC2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(6) **pairwise**
PWALC1	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(5) **partition**
PWALC2	พลังการทดสอบของตัวสถิติวาลด์จากตัวแบบ(6) **pairwise**

2. ขั้นตอนการทำงานของโปรแกรม

2.1 จำลองแบบข้อมูลตามตัวแบบที่ถูกต้องตัวแบบ(7) **quadratic** กำหนด $\beta_{22} = 0$

2.2 นำข้อมูลที่ได้จากข้อ 2.1 ฟิตตัวแบบ(4) **minimal** จากนั้นเรียงข้อมูลใหม่โดยเรียงตามค่าประมาณความน่าจะเป็นของตัวแปรตอบสนองระดับที่ 1 ($\hat{p}_1$)

2.3 นำข้อมูลที่ได้จากข้อ 2.2 ฟิตตัวแบบ(5) **partition** และตัวแบบ(6) **pairwise**

2.4 คำนวณค่าสถิติอัตราส่วนความน่าจะเป็นสูงสุดและค่าสถิติวาลด์ที่ได้จากตัวแบบในข้อ 2.3

2.5 ทำการทดสอบสมมติฐานว่างตัวแบบ(4) **minimal**

2.6 ทำซ้ำตั้งแต่ข้อ 2.1-2.5 จำนวน 500 ครั้ง แล้วนับจำนวนครั้งที่ปฏิเสธสมมติฐานว่างจากตัวสถิติภายใต้เงื่อนไขต่าง ๆ

2.7 พลังการทดสอบคำนวณจากจำนวนครั้งที่ปฏิเสธสมมติฐานว่างจากข้อ 2.6 หารด้วย 5

2.8 ทำซ้ำตั้งแต่ข้อ 2.1-2.7 โดยเพิ่มค่า β_{22} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

3 แสดงรายละเอียดของโปรแกรม

```
MACRO
STATISTIC n B01 B02 B1 B2 B22C POWERC1 POWERC2 PWALC1 PWALC2
MCOLUMN X1 X2 P P1 P2 P3 U Y A1 A2 A3 A4 A5 A6 A7 A8 A9 A10
        COEF1 COEF2 CO1 CO2
MCOLUMN  P02  EPRO1  EPRO2  EPRO3  INDICAT  B22C  POWERC1
MCOLUMN  POWERC2 PWALC1 PWALC2
MCONSTANT B01 B02  B1  B2 B22  J  n  LOGL1   LOGL2  LOGL3  I  K1  K2
          PV1  PV2  Q
MCONSTANT   DIV1  DIV2 DIV3  D1 D2 CUT2  POWER1 POWER2  REJ1  REJ2
          G WALD1 WALD2
MCONSTANT  KW1 KW2 PVWAL1 PVWAL2 REJW1 REJW2  PWAL1 PWAL2
          CON O
MMATRIX    BETA1 BETA2 BETA1T BETA2T XPWX1 XPWX2  XPWX1V
          XPWX2V MUL1 MUL2
MMATRIX  WAL1  WAL2
LET CON = N/2
LET B22 = 0
LET G = N/10
 SET INDICAT
 1( 1 2 3 4 5 6 7 8 9 10 )G
END.
     Indicator INDICAT   A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 .

DO Q = 1:15
     LET POWER1 = 0
     LET POWER2 = 0
     LET PWAL1 = 0
```

```
          LET PWAL2 = 0
          LET O = CON
          WHILE O = CON
          RANDOM n X1;
                    Bernulli .5.
           Random n X2;
                    Normal 0 1.
          SUM X1 O
          ENDWHILE
LET  P1 = EXPO(B01+(B2*X2)+(B1*X1)  +(B12*X2**2)   ) /
(1+EXPO(B01+(B2*X2)+(B1*X1) +(B12*X2**2)  ))
LET P02 = EXPO(B02+(B2*X2)+(B1*X1) +(B12*X2**2)   ) /
(1+EXPO(B02+(B2*X2)+(B1*X1) +(B12*X2**2)  ))
                              LET P2 = P02 - P1
                              LET P3 = 1-P1-P2
                              LET  P = P1 + P2 + P3
                    DO I = 1:500
                    Random n U;
                              Uniform 0 1.
                    DO J=1:n
                    LET CUT2 = P1(J) + P2(J)
                    IF U(J) <  P1(J)
                              LET Y(J) = 0
                    ELSE  IF U(J) >= CUT2
                               LET  Y(J) = 2
                    ELSE
                              LET  Y(J) = 1
                    ENDIF
                    ENDDO
          OLogistic Y = X1   X2 ;
          Logit;
           Eprobability EPRO1  EPRO2  EPRO3;
          Loglikelihood LOGL1;
          Brief 2.

SORT  X1  X2  Y  P1  P2  P3  P02  EPRO1  X1  X2  Y P1  P2  P3  P02  EPRO1;
          BY EPRO1 .
                    IF B22 < 1.3
          OLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
          Logit;
           Coefficients COEF1;
           XPWXinverse XPWX1;
          Loglikelihood LOGL2;
          Brief 2.
          table INDICAT Y;
          COUNTS.
          Copy COEF1  BETA1;
```

```
omit 1:4.
TRANSPOSE BETA1  BETA1T.
COPY XPWX1 XPWX1;
OMIT 1:4.
TRANSPOSE XPWX1 XPWX1.
COPY XPWX1 XPWX1;
OMIT 1:4.
INVERT XPWX1 XPWX1V.
MULTIPLY BETA1T XPWX1V MUL1.
MULTIPLY MUL1 BETA1  WAL1.
COPY WAL1  CO1.
LET WALD1 = CO1  ## DF 9  ##
OLogistic Y = X1   X2  X1*X2 ;
Logit;
 Coefficients COEF2;
 XPWXinverse XPWX2;
Loglikelihood LOGL3;
Brief 2.
Copy COEF2  BETA2;
omit 1:4.
TRANSPOSE BETA2  BETA2T.
COPY XPWX2 XPWX2;
OMIT 1:4.
TRANSPOSE XPWX2 XPWX2.
COPY XPWX2 XPWX2;
OMIT 1:4.
INVERT XPWX2 XPWX2V.
MULTIPLY BETA2T XPWX2V MUL2.
MULTIPLY MUL2 BETA2  WAL2.
COPY WAL2  CO2.
LET WALD2 = CO2  ##  DF 1 ####
LET DIV1 = (-2)*LOGL1
LET DIV2 = (-2)*LOGL2
LET DIV3 = (-2)*LOGL3
LET D1 = DIV1-DIV2
LET D2 = DIV1-DIV3
 CDF D1  K1;
 Chisquare 9.
 CDF D2  K2;
 Chisquare 1.
CDF WALD1  KW1;
 Chisquare 9.
CDF WALD2  KW2;
 Chisquare 1.
LET PV1 = 1-K1
LET PV2 = 1-K2
LET PVWAL1 = 1-KW1
```

```
          LET PVWAL2 = 1-KW2
                    IF PV1 < .05
                              LET  REJ1 = 1
                    ELSE
                               LET  REJ1 = 0
                    ENDIF
                    IF PV2 < .05
                              LET  REJ2 = 1
                    ELSE
                               LET  REJ2 = 0
                    ENDIF
                    IF PVWAL1 < .05
                              LET  REJW1 = 1
                    ELSE
                               LET  REJW1 = 0
                    ENDIF
                    IF PVWAL2 < .05
                              LET  REJW2 = 1
                    ELSE
                               LET  REJW2 = 0
                    ENDIF
                    LET  POWER1 = POWER1 + REJ1
                    LET  POWER2 = POWER2 + REJ2
                    LET  PWAL1 = PWAL1 + REJW1
                    LET  PWAL2 = PWAL2 + REJW2
          PRINT POWER1
          PRINT POWER2
          PRINT PWAL1
          PRINT PWAL2
          PRINT B12
          ENDDO
          LET POWER1 = POWER1 / 5
          LET POWER2 = POWER2 / 5
          LET PWAL1 = PWAL1 / 5
          LET PWAL2 = PWAL2 / 5
          LET B12C(Q) =  B12
          LET POWERC1(Q) = POWER1
          LET POWERC2(Q) = POWER2
          LET PWALC1(Q) = PWAL1
          LET PWALC2(Q) = PWAL2
ELSE
OLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
          Logit;
           Coefficients COEF1;
           XPWXinverse XPWX1;
          Loglikelihood LOGL2;
          Brief 2.
```

```
table INDICAT Y;
COUNTS.
OLogistic Y = X1   X2  X1*X2 ;
Logit;
 Coefficients COEF2;
 XPWXinverse XPWX2;
Loglikelihood LOGL3;
Brief 2.
Copy COEF2  BETA2;
omit 1:4.
TRANSPOSE BETA2  BETA2T.
COPY XPWX2 XPWX2;
OMIT 1:4.
TRANSPOSE XPWX2 XPWX2.
COPY XPWX2 XPWX2;
OMIT 1:4.
INVERT XPWX2 XPWX2V.
MULTIPLY BETA2T XPWX2V MUL2.
MULTIPLY MUL2 BETA2  WAL2.
COPY WAL2  CO2.
LET WALD2 = CO2  ##  DF 1 ####
LET DIV1 = (-2)*LOGL1
LET DIV2 = (-2)*LOGL2
LET DIV3 = (-2)*LOGL3
LET D1 = DIV1-DIV2
LET D2 = DIV1-DIV3
 CDF D1  K1;
 Chisquare 9.
 CDF D2  K2;
 Chisquare 1.
CDF WALD2  KW2;
 Chisquare 1.
LET PV1 = 1-K1
LET PV2 = 1-K2
LET PVWAL2 = 1-KW2
        IF PV1 < .05
                LET  REJ1 = 1
        ELSE
                 LET  REJ1 = 0
        ENDIF
        IF PV2 < .05
                LET  REJ2 = 1
        ELSE
                 LET  REJ2 = 0
        ENDIF
        IF PVWAL2 < .05
                LET  REJW2 = 1
```

```
            ELSE
                LET  REJW2 = 0
            ENDIF
            LET  POWER1 = POWER1 + REJ1
            LET  POWER2 = POWER2 + REJ2
            LET  PWAL2 = PWAL2 + REJW2
        PRINT POWER1
        PRINT POWER2
        PRINT PWAL2
        PRINT B12
        ENDDO
        LET POWER1 = POWER1 / 5
        LET POWER2 = POWER2 / 5
        LET PWAL2 = PWAL2 / 5
        LET B12C(Q) =  B12
        LET POWERC1(Q) = POWER1
        LET POWERC2(Q) = POWER2
        LET PWALC2(Q) = PWAL2
ENDIF
        LET B12 = B12 + 0.1
        LET B22 = B22 + 0.1

ENDDO
ENDMACRO
```


ตอนที่ 3

1. ความหมายของตัวแปรในโปรแกรม

MCONSTANT	
N	ขนาดตัวอย่าง
B01	ค่าพารามิเตอร์ α_1 ของตัวแบบที่ถูกต้อง
B02	ค่าพารามิเตอร์ α_2 ของตัวแบบที่ถูกต้อง
B1	ค่าพารามิเตอร์ β_1 ของตัวแบบที่ถูกต้อง
B2	ค่าพารามิเตอร์ β_2 ของตัวแบบที่ถูกต้อง
B12	ค่าพารามิเตอร์ β_{12} ของตัวแบบที่ถูกต้อง
LOGL1	ค่า log-likelihood จากตัวแบบ(4) **minimal**
LOGL2	ค่า log-likelihood จากตัวแบบ(5) **partition**
LOGL3	ค่า log-likelihood จากตัวแบบ(8) **baseline-category logit**

LOGL4	ค่า log-likelihood จากตัวแบบ(9) **baseline-partition**
PV1	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
PV2	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(9) **baseline-partition**
POWER1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
POWER2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(9) **baseline-partition**
MCOLUMN	
B12C	ค่า β_{12}
POWERC1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
POWERC2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(9) **baseline-partition**

2. ขั้นตอนการทำงานของโปรแกรม

2.1 จำลองแบบข้อมูลตามตัวแบบที่ถูกต้องตัวแบบ(6) **pairwise** กำหนด $\beta_{12} = 0$

2.2 นำข้อมูลที่ได้จากข้อ 2.1 ฟิตตัวแบบ(4) **minimal** จากนั้นเรียงข้อมูลใหม่โดยเรียงตามค่าประมาณความน่าจะเป็นของตัวแปรตอบสนองระดับที่ 1 ($\hat{p}_1$)

2.3 นำข้อมูลที่ได้จากข้อ 2.2 ฟิตตัวแบบ(5) **partition**

2.4 คำนวณค่าสถิติอัตราส่วนความน่าจะเป็นสูงสุดที่ได้จากตัวแบบในข้อ 2.3

2.5 ทำการทดสอบสมมติฐานว่างตัวแบบ(4) **minimal**

2.6 นำข้อมูลที่ได้จากข้อ 2.2 ฟิตตัวแบบ(8) **baseline-category logit** จากนั้นเรียงข้อมูลใหม่โดยเรียงตามค่าประมาณความน่าจะเป็นของตัวแปรตอบสนองระดับที่ 1 ($\hat{p}_1$)

2.7 นำข้อมูลที่ได้จากข้อ 2.6 ฟิตตัวแบบ(9) **baseline-partition**

2.8 คำนวณค่าสถิติอัตราส่วนความน่าจะเป็นสูงสุดที่ได้จากตัวแบบในข้อ 2.7

2.9 ทำการทดสอบสมมติฐานว่างตัวแบบ(8) **baseline-category logit**

2.10 ทำซ้ำตั้งแต่ข้อ 2.1-2.9 จำนวน 500 ครั้ง แล้วบันทึกจำนนวนครั้งที่ปฏิเสธสมมติฐานว่าง จากตัวสถิติภายใต้เงื่อนไขต่าง ๆ

2.11 พลังการทดสอบคำนวณจากจำนวนครั้งที่ปฏิเสธสมมติฐานว่างจากข้อ 2.10 หารด้วย 5

2.12 ทำซ้ำตั้งแต่ข้อ 2.1-2.7 โดยเพิ่มค่า β_{12} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

3. แสดงรายละเอียดของโปรแกรม

```
MACRO
Y3              n  B01 B02  B1  B2   INDICAT   X1 X2 Y P1 P2 P3 P B12C
                POWERC1 POWERC2
MCOLUMN         X1  X2  P  P1  P2  P3  U  Y A1 A2 A3 A4 A5 A6 A7 A8 A9 A10
MCOLUMN         P02   EPRO1  EPRO2  EPRO3  EPRO4  EPRO5  EPRO6  INDICAT
                B12C
MCOLUMN         POWERC1 POWERC2
MCONSTANT B01 B02  B1 B2 B12  J  n  LOGL1  POWER1 REJ1   I  K1 K2  PV1
                PV2  Q G O
MCONSTANT  LOGL2 LOGL3 LOGL4 DIV1 DIV2 DIV3 DIV4  D1 D2  CUT2
                POWER2 REJ2
MCONSTANT  CON

LET CON = N/2
LET O = CON
WHILE O = CON
        RANDOM n X1;
        Bernulli .5.
         Random n X2;
        Normal 0 1.
         SUM X1 O
ENDWHILE
LET B12 = 0
LET G = N/10
 SET INDICAT
 1( 1 2 3 4 5 6 7 8 9 10 )G
END.
         Indicator INDICAT   A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 .

DO Q = 1:15

        LET POWER1 = 0
        LET POWER2 = 0
LET  P1 = EXPO(B01+(B2*X2)+(B1*X1)  +(B12*X1*X2)   ) /
(1+EXPO(B01+(B2*X2)+(B1*X1) +(B12*X1*X2)  ))
LET P02 = EXPO(B02+(B2*X2)+(B1*X1) +(B12*X1*X2)   ) /
(1+EXPO(B02+(B2*X2)+(B1*X1) +(B12*X1*X2)  ))
                        LET P2 = P02 - P1
```

```
            LET P3 = 1-P1-P2
            LET  P = P1 + P2 + P3
        DO I = 1:500

        Random n U;
            Uniform 0 1.
        DO J=1:n
        LET CUT2 = P1(J) + P2(J)
        IF U(J) <  P1(J)
            LET Y(J) = 0
        ELSE  IF U(J) >= CUT2
             LET  Y(J) = 2
        ELSE
            LET  Y(J) = 1
        ENDIF
        ENDDO
OLogistic Y = X1   X2 ;
Logit;
 Eprobability EPRO1  EPRO2  EPRO3;
Loglikelihood LOGL1;
Brief 2.
SORT  X1  X2  Y  P1  P2  P3  P02  X1  X2  Y P1  P2  P3  P02 ;
BY EPRO1 .
OLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
Logit;
Loglikelihood LOGL2;
Brief 2.
NLogistic Y = X1   X2 ;
 Eprobability EPRO4  EPRO5  EPRO6;
Loglikelihood LOGL3;
Brief 2.
SORT  X1  X2  Y  P1  P2  P3  P02    X1  X2  Y P1  P2  P3  P02  ;
BY EPRO4 .
NLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
Loglikelihood LOGL4;
Brief 2.
LET DIV1 = (-2)*LOGL1
LET DIV2 = (-2)*LOGL2
LET DIV3 = (-2)*LOGL3
LET DIV4 = (-2)*LOGL4
LET D1 = DIV1-DIV2
LET D2 = DIV3-DIV4
 CDF D1  K1;
 Chisquare 9.
 CDF D2  K2;
 Chisquare 18.
LET PV1 = 1-K1
```

```
        LET PV2 = 1-K2
                IF PV1 < .05
                        LET  REJ1 = 1
                ELSE
                         LET  REJ1 = 0
                ENDIF
                IF PV2 < .05
                        LET  REJ2 = 1
                ELSE
                         LET  REJ2 = 0
                ENDIF
                LET  POWER1 = POWER1 + REJ1
                LET  POWER2 = POWER2 + REJ2
        PRINT POWER1
        #PRINT PV1#
        PRINT POWER2
        #PRINT PV2#
        PRINT B12
        ENDDO
        LET POWER1 = POWER1 / 5
        LET POWER2 = POWER2 / 5
        LET B12C(Q) =  B12
        LET POWERC1(Q) = POWER1
        LET POWERC2(Q) = POWER2
        LET B12 = B12 + 0.1
ENDDO
ENDMACRO
```


ตอนที่ 4

1. ความหมายของตัวแปรในโปรแกรม

MCONSTANT	
N	ขนาดตัวอย่าง
B01	ค่าพารามิเตอร์ α_1 ของตัวแบบที่ถูกต้อง
B02	ค่าพารามิเตอร์ α_2 ของตัวแบบที่ถูกต้อง
B1	ค่าพารามิเตอร์ β_1 ของตัวแบบที่ถูกต้อง
B2	ค่าพารามิเตอร์ β_2 ของตัวแบบที่ถูกต้อง
B22	ค่าพารามิเตอร์ β_{22} ของตัวแบบที่ถูกต้อง
LOGL1	ค่า log-likelihood จากตัวแบบ(4) **minimal**
LOGL2	ค่า log-likelihood จากตัวแบบ(5) **partition**

LOGL3	ค่า log-likelihood จากตัวแบบ(8) **baseline-category logit**
LOGL4	ค่า log-likelihood จากตัวแบบ(9) **baseline-partition**
PV1	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
PV2	P-value ของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(9) **baseline-partition**
POWER1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
POWER2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(9) **baseline-partition**
MCOLUMN	
B22C	ค่า β_{22}
POWERC1	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(5) **partition**
POWERC2	พลังการทดสอบของตัวสถิติอัตราส่วนความน่าจะเป็นสูงสุดจากตัวแบบ(9) **baseline-partition**

2. **ขั้นตอนการทำงานของโปรแกรม**

2.1 จำลองแบบข้อมูลตามตัวแบบที่ถูกต้องตัวแบบ(6) **pairwise** กำหนด $\beta_{22} = 0$

2.2 นำข้อมูลที่ได้จากข้อ 2.1 ฟิตตัวแบบ(4) **minimal** จากนั้นเรียงข้อมูลใหม่โดยเรียงตามค่าประมาณความน่าจะเป็นของตัวแปรตอบสนองระดับที่ 1 ($\hat{p}_1$)

2.3 นำข้อมูลที่ได้จากข้อ 2.2 ฟิตตัวแบบ(5) **partition**

2.4 คำนวณค่าสถิติอัตราส่วนความน่าจะเป็นสูงสุดที่ได้จากตัวแบบในข้อ 2.3

2.5 ทำการทดสอบสมมติฐานว่างตัวแบบ(4) **minimal**

2.6 นำข้อมูลที่ได้จากข้อ 2.2 ฟิตตัวแบบ(8) **baseline-category logit** จากนั้นเรียงข้อมูลใหม่โดยเรียงตามค่าประมาณความน่าจะเป็นของตัวแปรตอบสนองระดับที่ 1 ($\hat{p}_1$)

2.7 นำข้อมูลที่ได้จากข้อ 2.6 ฟิตตัวแบบ(9) **baseline-partition**

2.8 คำนวณค่าสถิติอัตราส่วนความน่าจะเป็นสูงสุดที่ได้จากตัวแบบในข้อ 2.7

2.9 ทำการทดสอบสมมติฐานว่างตัวแบบ(8) **baseline-category logit**

2.10 ทำซ้ำตั้งแต่ข้อ 2.1-2.9 จำนวน 500 ครั้ง แล้วบันทึกจำนนวนครั้งที่ปฏิเสธสมมติฐานว่างจากตัวสถิติภายใต้เงื่อนไขต่าง ๆ

2.11 พลังการทดสอบคำนวณจากจำนวนครั้งที่ปฏิเสธสมมติฐานว่างจากข้อ 2.10 หารด้วย 5

2.12 ทำซ้ำตั้งแต่ข้อ 2.1-2.7 โดยเพิ่มค่า β_{22} จาก 0-1.4 โดยเพิ่มครั้งละ 0.1

3 แสดงรายละเอียดของโปรแกรม

```
MACRO
Y3              n  B01 B02  B1  B2   INDICAT   X1 X2 Y P1 P2 P3 P B22C
                POWERC1 POWERC2
MCOLUMN     X1  X2  P  P1  P2  P3  U  Y A1 A2 A3 A4 A5 A6 A7 A8 A9 A10
MCOLUMN     P02   EPRO1  EPRO2  EPRO3  EPRO4  EPRO5  EPRO6  INDICAT
            B22C
MCOLUMN     POWERC1 POWERC2
MCONSTANT B01 B02  B1 B2 B22  J  n  LOGL1  POWER1 REJ1   I  K1 K2  PV1
            PV2  Q G O
MCONSTANT  LOGL2 LOGL3 LOGL4 DIV1 DIV2 DIV3 DIV4  D1 D2  CUT2
            POWER2 REJ2
MCONSTANT  CON

LET CON = N/2
LET O = CON
WHILE O = CON
        RANDOM n X1;
        Bernulli .5.
         Random n X2;
        Normal 0 1.
         SUM X1 O
ENDWHILE
LET B22 = 0
LET G = N/10
 SET INDICAT
 1( 1 2 3 4 5 6 7 8 9 10 )G
END.
         Indicator INDICAT   A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 .

DO Q = 1:15

        LET POWER1 = 0
        LET POWER2 = 0
```

```
LET  P1 = EXPO(B01+(B2*X2)+(B1*X1)  +(B12*X1*X2)   ) /
(1+EXPO(B01+(B2*X2)+(B1*X1) +(B12*X1*X2)  ))
LET P02 = EXPO(B02+(B2*X2)+(B1*X1) +(B12*X1*X2)   ) /
(1+EXPO(B02+(B2*X2)+(B1*X1) +(B12*X1*X2)  ))
                    LET P2 = P02 - P1
                    LET P3 = 1-P1-P2
                    LET  P = P1 + P2 + P3
      DO I = 1:500
            Random n U;
                    Uniform 0 1.
            DO J=1:n
            LET CUT2 = P1(J) + P2(J)
            IF U(J) <  P1(J)
                    LET Y(J) = 0
            ELSE  IF U(J) >= CUT2
                     LET  Y(J) = 2
            ELSE
                    LET  Y(J) = 1
            ENDIF
            ENDDO
      OLogistic Y = X1   X2 ;
      Logit;
       Eprobability EPRO1  EPRO2  EPRO3;
      Loglikelihood LOGL1;
      Brief 2.
      SORT  X1  X2  Y  P1  P2  P3  P02  X1  X2  Y P1  P2  P3  P02 ;
      BY EPRO1 .
      OLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
      Logit;
      Loglikelihood LOGL2;
      Brief 2.
      NLogistic Y = X1   X2 ;
       Eprobability EPRO4  EPRO5  EPRO6;
      Loglikelihood LOGL3;
      Brief 2.
      SORT  X1  X2  Y  P1  P2  P3  P02    X1  X2  Y P1  P2  P3  P02  ;
      BY EPRO4 .
      NLogistic Y = X1  X2  A1  A2  A3  A4  A5  A6  A7  A8  A9   ;
      Loglikelihood LOGL4;
      Brief 2.
      LET DIV1 = (-2)*LOGL1
      LET DIV2 = (-2)*LOGL2
      LET DIV3 = (-2)*LOGL3
      LET DIV4 = (-2)*LOGL4
      LET D1 = DIV1-DIV2
      LET D2 = DIV3-DIV4
       CDF D1  K1;
```

```
 Chisquare 9.
 CDF D2  K2;
 Chisquare 18.
LET PV1 = 1-K1
LET PV2 = 1-K2
        IF PV1 < .05
                LET  REJ1 = 1
        ELSE
                 LET  REJ1 = 0
        ENDIF
        IF PV2 < .05
                LET  REJ2 = 1
        ELSE
                 LET  REJ2 = 0
        ENDIF
        LET  POWER1 = POWER1 + REJ1
        LET  POWER2 = POWER2 + REJ2
PRINT POWER1
#PRINT PV1#
PRINT POWER2
#PRINT PV2#
PRINT B22
ENDDO
```



```
LET POWER1 = POWER1 / 5
LET POWER2 = POWER2 / 5
LET B22C(Q) =  B22
LET POWERC1(Q) = POWER1
LET POWERC2(Q) = POWER2
LET B22 = B22 + 0.1
ENDDO
ENDMACRO
```

ประวัติผู้วิจัย

ชื่อ	นายสุจินต์ สุขกุมภาพันธ์
	ที่อยู่ 91 หมู่ 9 ตำบลหนองงูเหลือม อำเภอเมือง จังหวัดนครปฐม รหัสไปรษณีย์ 73000
ประวัติการศึกษา	
พ.ศ. 2544	สำเร็จการศึกษาวิทยาศาสตร์บัณฑิต สาขาเทคโนโลยีปิโตรเคมีและวัสดุพอลิเมอร์ จากมหาวิทยาลัยศิลปากร
พ.ศ. 2546	ศึกษาต่อระดับปริญญามหาบัณฑิต สาขาสถิติประยุกต์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร