ว.วิทย. มข. 44(2) 408-417 (2559) KKU Sci. J. 44(2) 408-417 (2016)

# ผลลัพธ์เชิงเปรียบเทียบด้วยพื้นที่ใต้โค้ง ROC และกำลังการทดสอบระหว่างตัวแบบผสมเชิงเส้นนัยทั่วไปแบบลอจิตสะสมและแบบโพรบิตสะสม

# A Comparative Gain in Areas under ROC Curve and Power of Tests Between Cumulative Logit GLMMs and Cumulative Probit GLMMs

วีรานันท์ พงศาภักดี*

ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร วิทยาเขตพระราชวังสนามจันทร์ นครปฐม 73000

*E-mail: pongsapukdee_v@silpakorn.edu, veeranun@hotmail.com

## บทคัดย่อ

ตัวแบบเชิงเส้นนัยทั่วไป (GLMs) ขยายจากตัวแบบเชิงเส้นแบบธรรมดาที่มีอิทธิพลตรึง (fixed effects) โดยใช้ฟังก์ชันของค่าเฉลี่ยเป็นฟังก์ชันเชื่อมโยงตัวแปรตอบสนองกับผลรวมเชิงเส้นแทนฟังก์ชันเดิมที่เป็นฟังก์ชันเอกรูป (identity link function) และสามารถใช้กับการแจกแจงของตัวแปรตอบสนองในวงศ์เลขชี้กำลังด้วย ส่วนตัวแบบผสมเชิงเส้นนัยทั่วไป (GLMMs) ขยายจากตัวแบบเชิงเส้นนัยทั่วไปเพื่อให้มีทั้งอิทธิพลตรึงและอิทธิพลสุ่ม (random effects) ในตัวแบบ และใช้กับข้อมูลที่อาจมีความสัมพันธ์กันภายในกลุ่มอีกด้วย อย่างไรก็ตามการนำตัวแบบ GLMMs ไปใช้ยังมีปัญหาเรื่องความซับซ้อนและประสิทธิภาพของตัวแบบ งานวิจัยนี้จึงเปรียบเทียบประสิทธิภาพของตัวแบบผสมเชิงเส้นทั่วไปลอจิตสะสม (cumulative logit GLMMs) และตัวแบบผสมเชิงเส้นทั่วไปโพรบิตสะสม (cumulative probit GLMMs) ภายใต้พื้นที่ใต้โค้ง ROC และกำลังการทดสอบ ด้วยการจำลองแบบข้อมูลเชิงประจักษ์ 1,000 ชุดในแต่ละเงื่อนไขของพารามิเตอร์ ขนาดตัวอย่าง จำนวนกลุ่ม และค่าสหสัมพันธ์ภายในกลุ่ม โดยใช้โปรแกรมแมโครร่วมกับโปรแกม SAS version 9.1 ผลการวิจัยพบว่าทั้งสองตัวแบบให้พื้นที่ใต้โค้ง ROC ค่อนข้างสูงใกล้เคียงกัน แต่ตัวแบบผสมเชิงเส้นทั่วไปลอจิตสะสม มีประสิทธิภาพดีกว่าตัวแบบผสมเชิงเส้นทั่วไปโพรบิตสะสมเป็นส่วนใหญ่ โดยผลลัพธ์เชิงเปรียบเทียบให้พื้นที่ใต้โค้ง ROC มากกว่า ซึ่งมีค่าสูงสุด 0.8% ส่วนกำลังการทดสอบมีผลลัพธ์เชิงเปรียบเทียบมีค่าสูงกว่ามาก ค่าสูงสุดประมาณ 10% เมื่อค่าสหสัมพันธ์ภายในกลุ่มมีค่าน้อยถึงปานกลางตัวแบบผสมเชิงเส้นทั่วไปลอจิตสะสมให้ผลลัพธ์ดีกว่าทุกกรณี นอกนั้นเป็นกรณีที่มีความซับซ้อน ซึ่งตัวแบบทั้งสองแม้จะมีประสิทธิภาพน้อยลงบ้างแต่ยังคงภาวะสารูปดีอย่างมีนัยสำคัญทางสถิติที่ 0.05

และให้ผลลัพธ์เชิงเปรียบเทียบใกล้เคียงกัน จึงเป็นตัวแบบเชิงสถิติขั้นสูงสองตัวแบบที่อาจนำไปประยุกต์กับข้อมูลจริงต่อไปอย่างมีความน่าเชื่อถือในเชิงสถิติ

## ABSTRACT

Generalized Linear Models (GLMs) extend ordinary general linear model with fixed effects in the linear predictor by allowing non-normal responses and a link function of the mean. The Generalized Linear Mixed Models (GLMMs) are further extensions of GLMs that permits both the fixed effect and the random effect in models and account for the dependency inherent in data. However, limitation of few applications was found due to the complexity of the models and their efficiency. In this article, a further comparative gain in area under ROC curves and power of tests between the cumulative logit GLMMs and the cumulative probit GLMMs are investigated and discussed. The 1,000 empirical datasets for each condition of parameters and the number of clusters and cluster sizes are simulated using SAS rewritten macro program. The results reveal that the cumulative logit GLMM is superior (0.8%) to the cumulative probit GLMM. As the number of clusters and the cluster sizes are increased, the sensitivity and the precision through the AUC and power of the tests are better fitted. Overall, the maximum absolute percentage gain power between the two models is approximately 10% with satisfactorily high area under ROC curve values. It is clear that, for small and moderate intra-cluster correlation, the cumulative logit GLMMs are also preferred; otherwise, for more complicated cases; even if the two GLMMs are less efficiency than that before but still are closely adequate of fits significantly at 0.05. Hence the two GLMMs probably are used for further implementation in real data.

**คำสำคัญ:** การประเมินตัวแบบ ตัวแบบลอจิตสะสม GLMMS ตัวแบบโพรบิตสะสม GLMMS ค่าตอบสนองอันดับที่สัมพันธ์กัน

**Keywords:** Assessing of model, Logit Generalized Linear Mixed Models, Probit Generalized Linear Mixed Models, Correlated ordinal responses

## บทนำ

การวิเคราะห์ข้อมูลจำแนกประเภท เช่น การวิเคราะห์ตารางการจร การพยากรณ์ค่าคาดหมาย การประเมินผลจากตัวแบบเชิงสถิติเพื่อศึกษาขนาดของอิทธิพลที่ส่งผลกรทบต่อระดับหรือประเภทที่สนใจศึกษา และการพยากรณ์ความน่าจะเป็นของระดับต่าง ๆ ได้รับความสนใจและนิยมใช้อย่างแพร่หลายในปัจจุบัน ชุดข้อมูลอาจเป็นข้อมูลทางการแพทย์ การควบคุมคุณภาพ พันธุกรรม การตลาด สังคมศาสตร์ สิ่งแวดล้อมและนิเวศวิทยา ฯลฯ ซึ่งสนใจค่าตอบสนองจำแนกประเภทที่อาจแบ่งเป็น 2 ประเภทหรือมากกว่า (dichotomous or polychromous)

อาจเป็นข้อมูลนามบัญญัติ อันดับ หรือจำนวนนับไม่ต่อเนื่อง เช่น จำนวนคนจำแนกตามกลุ่มของระยะเวลาที่รอดชีพ (discrete counts or grouped survival times) นอกจากนี้โดยทั่วไปมีการรวบรวมข้อมูลสำหรับการอธิบายตัวแปรตอบสนองด้วย ข้อมูลของตัวแปรอธิบายอาจเป็นข้อมูลจำแนกประเภท เชิงกลุ่ม หรือต่อเนื่อง ข้อมูลควรมาจากตัวแทนของประชากร เช่น การสำรวจตัวอย่าง การทดลอง หรือทั้งสองแบบ หรือมาจากงานการจำลองแบบเชิงสุ่ม ซึ่งนำมาสู่กระบวนการวิเคราะห์และการสร้างตัวแบบเชิงสถิติต่อไป เช่น ตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Models or GLMs, Nelder and Wededderburn, 1972; McCullagh, 1983, McCullagh and Nelder, 1989) ตัวแบบลอจิตสะสม (cumulative logiit model) ตัวแบบโพรบิตสะสม (cumulative probit model) (Agresti, 2002; วีรานันท์ พงศาภักดี, 2555; Pongsapukdee and Sukgumphaphan, 2008) และตัวแบบผสมเชิงเส้นนัยทั่วไป (Generalized Linear Mixed Models or GLMMs, McCulloch and Searle, 2001)

โดยทั่วไปเมื่อตัวแปรตอบสนองเป็นแบบอันดับ ตัวแบบลอจิตสะสมและตัวแบบโพรบิตสะสมจะมีประสิทธิภาพหรือมีกำลังการทดสอบที่ดีกว่าตัวแบบลอจิตและตัวแบบโพรบิตแบบธรรมดา (Agresti, 2013) และเป็นตัวแบบที่มีใช้ในโปรแกรมสำเร็จรูปขั้นสูงเช่น SAS ได้โดยสะดวก ในหลายสถานการณ์ข้อมูลที่สนใจศึกษาอาจมีการวัดซ้ำหลายครั้ง ข้อมูลไม่อิสระต่อกัน การใช้ตัวแบบเชิงเส้นนัยทั่วไปจะไม่เพียงพอหรือไม่สอดคล้องกับสมบัติและข้อสมมติของตัวแบบ ทำให้มีงานวิจัยศึกษาที่เสนอ ตัวแบบผสมเชิงเส้นนัยทั่วไป (Generalized Linear Mixed Models or GLMMs, Liang and Zeger, 1986; Goldstein, 1991 และ McCulloch and Searle, 2001) เพื่อใช้แทนตัวแบบ GLMs อย่างไรก็ตามตัวแบบ GLMMs ที่พัฒนาขึ้นมีรูปแบบที่ซับซ้อนและมีข้อจำกัดเรื่องประสิทธิภาพและวิธีประมาณพารามิเตอร์ จึงมีวิธีการประมาณพารามิเตอร์ของตัวแบบวิธีใหม่ ๆ ที่มีประสิทธิภาพมากขึ้นหลายวิธี เช่น วิธี Generalized Estimating Equation or GEE (Liang and Zeger, 1986; Goldstein, 1991,1995); วิธี Marginal Quasi-Likelihood or MQL (Goldstein, 1991, 1995); วิธี Penalized Quasi-Likelihood or PQL (Green, 1987) และวิธี Adaptive Gaussian Quadrature or AGQ (Liu and Pierce, 1994) ซึ่งพบว่าวิธี AGQ ใช้สำหรับการประมาณพารามิเตอร์ของตัวแบบ GLMMS ได้อย่างมีประสิทธิภาพและมีความเชื่อถือได้ โดยสรุปดังนี้

สำหรับตัวแปรตอบสนองแบบ 2 กลุ่มที่มีการวัดซ้ำนั้น Liang and Zeger (1986) เสนอระบบสมการสำหรับการประมาณพารามิเตอร์ของตัวแบบ GLMs และ GLMMs เรียกว่า Generalized Estimating Equations (GEEs) พบว่าตัวประมาณพารามิเตอร์และความแปรปรวนจากวิธี GEE ให้ตัวประมาณที่คงเส้นคงวาและมีการแจกแจงปกติเชิงเส้นกำกับ (asymptotic normal) ต่อมา Green (1987) เสนอทางเลือกใหม่เรียกว่าวิธี Penalized Quasi-Likelihood (PQL) โดยใช้ Laplace approximation พบว่าวิธี PQL ให้ตัวประมาณพารามิเตอร์ที่เอนเอียงเมื่อขนาดตัวอย่างมีขนาดเล็กและสหสัมพันธ์ของข้อมูลในกลุ่มมีค่าสูง Goldstein (1991) เสนออีกวิธีหนึ่งที่เรียกว่า Marginal Quasi-Likelihood (MQL) และพบว่าวิธี MQL เป็นตัวประมาณพารามิเตอร์ที่เอนเอียงเมื่อขนาดตัวอย่างมีขนาดเล็กเช่นกัน Breslow and Clayton (1993) ได้เปรียบเทียบประสิทธิภาพของตัวแบบข้างต้นใหม่สำหรับตัวแปรตอบสนองแบบ 2 กลุ่ม เมื่อค่าตอบสนองมีความสัมพันธ์กันโดยใช้วิธีประมาณพารามิเตอร์ 2 วิธี

คือ Penalized Quasi -Likelihood (PQL) และ Marginal Quasi-Likelihood (MQL) พบว่าวิธี PQL ให้ค่าประมาณสัมประสิทธิ์ของตัวแบบและความคลาดเคลื่อนมาตรฐานที่ต่ำกว่าค่าจริง และให้ค่าประมาณความแปรปรวนที่ต่ำกว่าค่าจริงมาก ส่วนวิธี MQL ให้ค่าประมาณความแปรปรวนที่สูงกว่าค่าจริง Liu and Pierce (1994) เสนอวิธีใหม่เรียกว่า Gauss-Hermite quadrature (GH) หรือ Gaussian Quadrature (GQ) โดยการหาผลรวมแบบถ่วงน้ำหนัก และกำหนดจำนวน Quadrature point ในการหาผลรวมก่อนที่จะหาค่าประมาณพารามิเตอร์ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นมีค่าสูงสุด วิธีนี้พบว่าจะมีความถูกต้องยิ่งขึ้นถ้าจำนวน Quadrature point มากขึ้น แต่ต้องใช้เวลาการคำนวณมากขึ้นด้วย ต่อมา Pinheiro and Bates (1995) ได้นำวิธี Gaussian Quadrature มาปรับตรง Location และ Dispersion ของการแจกแจง เพื่อให้ใช้จำนวน Quadrature point ในแต่ละมิติของการอินทิเกรตน้อยลงแต่ยังคงความถูกต้องเรียกว่า “Adaptive Gaussian Quadrature ” หรือ AGQ และพบว่าให้ผลลัพธ์ดีขึ้น

ดังกล่าวข้างต้นตัวแบบเชิงเส้นนัยทั่วไปหรือ GLMs และตัวแบบผสมเชิงเส้นนัยทั่วไปหรือ GLMMs เป็นตัวแบบหลักและตัวแบบสำคัญในการวิเคราะห์ข้อมูลจำแนกประเภท จึงยังมีการศึกษาวิจัยด้านความเหมาะสมและประสิทธิภาพของตัวแบบในกลุ่มของตัวแบบ GLMs และ GLMMs อย่างต่อเนื่อง ซึ่งรวมถึงวิธีประมาณพารามิเตอร์ของตัวแบบด้วย

งานวิจัยนี้ได้เลือกวิธี AGQ (Liu and Pierce, 1994) สำหรับการประมาณพารามิเตอร์ของตัวแบบ GLMMs และสนใจศึกษาผลลัพธ์เชิงเปรียบเทียบในเรื่องพื้นที่ใต้โค้ง ROC และกำลังการทดสอบระหว่างตัวแบบ GLMMs แบบลอจิตสะสม (the cumulative logit GLMMs) กับตัวแบบ GLMMs แบบโพรบิตสะสม (the cumulative probit GLMMs) ภายใต้เงื่อนไขของพารามิเตอร์ที่เกี่ยวข้อง โดยคำนวณผลลัพธ์เชิงเปรียบเทียบด้วยค่าเฉลี่ย ค่าผลต่างของค่าเฉลี่ยใน 1,000 ชุด รวมทั้งร้อยละของผลต่างของพื้นที่ใต้โค้ง ROC และค่ากำลังการทดสอบ จำแนกตามแต่ละเงื่อนไขของจำนวนกลุ่ม (number of clusters, N = 50, 100, 250), ขนาดของกลุ่ม cluster sizes, $n_i$ = 5, 7, 10), และสหสัมพันธ์ภายในกลุ่ม (intra-cluster correlation, ICC = 0.05, 0.20, 0.40) ซึ่งเป็นพารามิเตอร์ของการจำลองแบบภายใต้ตัวแบบ GLMMs ทั้งสองตัวแบบ

## วิธีการดำเนินการวิจัย

### ตัวแบบผสมเชิงเส้นนัยทั่วไปลอจิตสะสม โพรบิตสะสม และการประมาณพารามิเตอร์

ตัวแบบลอจิตสะสม เมื่อเพิ่มเทอมของอิทธิพลสุ่ม $\boldsymbol{u}_i$ เข้าไปในผลรวมเชิงเส้น ( $\boldsymbol{x}'\boldsymbol{\beta}$ ) เรียกว่า ตัวแบบลอจิตสะสมแบบอิทธิพลสุ่ม หรือตัวแบบสะสม GLMMs หรือ ตัวแบบผสมเชิงเส้นนัยทั่วไปลอจิตสะสม ที่ตัวแปรตอบสนองอันดับมีจำนวน J ระดับหรือกลุ่ม ดังต่อไปนี้

$$\text{logit}[P(Y \le j \mid \boldsymbol{x}, \boldsymbol{u}_i)] = \alpha_j + \boldsymbol{x}'\boldsymbol{\beta} + \boldsymbol{z}'\boldsymbol{u}_i, \quad j = 1, \ldots, J-1 \qquad (1)$$

$$\text{probit}[P(Y \le j \mid \boldsymbol{x}, \boldsymbol{u}_i)] = \alpha_j + \boldsymbol{x}'\boldsymbol{\beta} + \boldsymbol{z}'\boldsymbol{u}_i, \quad j = 1, \ldots, J-1. \qquad (2)$$

เมื่อ $\boldsymbol{u}_i$ แทนเวกเตอร์อิทธิพลสุ่มของกลุ่ม $i$ ที่เป็นอิสระต่อกันและมีการแจกแจงแบบปกติด้วยค่าเฉลี่ย $\boldsymbol{0}$ และเมทริกซ์ความแปรปรวนร่วม $\boldsymbol{\Sigma}_u$ และ $\alpha_j$ แทน จุดตัด (intercept) ที่ j, $\boldsymbol{x}$ แทนเวกเตอร์ตัวแปรอธิบาย

สำหรับอิทธิพลตรึง $\boldsymbol{z}$ แทนเวกเตอร์ตัวแปรอธิบายสำหรับอิทธิพลสุ่ม และตัวแบบอิทธิพลสุ่มพื้นฐาน (เมื่อ $\boldsymbol{z}'_{ij} = 1$) ลดรูปแบบเป็น

$$\operatorname{logit}[P(Y \le j \mid \boldsymbol{x}, u_i)] = \alpha_j + \boldsymbol{x}'\boldsymbol{\beta} + u_i, \quad j = 1, \ldots, J-1$$
$$\operatorname{probit}[P(Y \le j \mid \boldsymbol{x}, u_i)] = \alpha_j + \boldsymbol{x}'\boldsymbol{\beta} + u_i, \quad j = 1, \ldots, J-1$$

เมื่อ $\boldsymbol{u}_i$ แทนอิทธิพลสุ่มของกลุ่ม *i* ที่เป็นอิสระต่อกันและมีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และความแปรปรวน $\sigma_u^2$

โดยที่ ความน่าจะเป็นสะสมแบบมีเงื่อนไข สามารถคำนวณจาก

$$P(Y_{ij} \le j \mid \boldsymbol{x}_{ij}, u_i) = P(Y_{ij}^* \le c_j \mid \boldsymbol{x}_{ij}, u_i) \;= P(-\boldsymbol{x}'_{ij}\boldsymbol{\beta} - u_i + \varepsilon_{ij} \le c_j)$$

สำหรับภาวะน่าจะเป็นแบบมีเงื่อนไขของกลุ่ม *i* มีแนวคิดที่ว่าตัวแปรตอบสนองแบบอันดับมาจากการจัดกลุ่มของตัวแปรแฝงแบบต่อเนื่อง $Y^*$ ซึ่งเป็นตัวแปรที่ไม่ได้สังเกตค่า ซึ่งในกรณีนี้ $Y^*$ มีความสัมพันธ์กับตัวแปรอธิบายในรูปแบบของ $Y_{ij}^* = -\boldsymbol{x}'\boldsymbol{\beta} - u_i + \varepsilon_{ij}$ เมื่อ $\varepsilon_{ij}$ คือความคลาดเคลื่อนของตัวแปรแฝง และ $Y_{ij} = j$ เมื่อ $c_{j-1} < Y_{ij}^* \le c_j$

การประมาณพารามิเตอร์ของตัวแบบ GLMMs อาจใช้วิธีภาวะน่าจะเป็นสูงสุดสำหรับตัวแปรตอบสนองแบบอันดับภายใต้ฟังก์ชันภาวะน่าจะเป็นแบบมีเงื่อนไขของกลุ่ม *i* โดยมีข้อสมมติที่ว่าค่าตอบสนองของหน่วยที่ *i* เป็นอิสระต่อกันเมื่อกำหนดอิทธิพลสุ่ม เรียกข้อสมมตินี้ว่า ความเป็นอิสระกันแบบมีเงื่อนไข และภาวะน่าจะเป็นแบบไม่มีเงื่อนไขหรือภาวะน่าจะเป็นแบบมาร์จินัลของ $\mathrm{Y_i}$ แสดงในรูปแบบของอินทิกรัลของภาวะน่าจะเป็นแบบมีเงื่อนไข $f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta})$ คือ

$$L_i(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = \int f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta}) \times g(\boldsymbol{u}_i; \boldsymbol{\Sigma}_u) d\boldsymbol{u}_i$$

เมื่อ $g(\boldsymbol{u}_i; \boldsymbol{\Sigma}_u)$ แทนฟังก์ชันความหนาแน่นของอิทธิพลสุ่ม ซึ่งมักจะสมมติว่าเป็นฟังก์ชันความหนาแน่นแบบปกติหลายตัวแปร ในขณะที่ $f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta})$ แทนภาวะน่าจะเป็นแบบมีเงื่อนไขของกลุ่ม *i* และ $L_i(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u)$ แทนภาวะน่าจะเป็นแบบไม่มีเงื่อนไขหรือแบบมาร์จินัล สำหรับเวกเตอร์ค่าตอบสนองของกลุ่ม *i* แล้ว The log-likelihood function ของตัวอย่าง N กลุ่มแสดงไว้ใน (3) คือ

$$l(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = \log L(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = \log \prod_{i=1}^{N} L_i(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = \sum_{i=1}^{N} \log L_i(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) \qquad (3)$$

เนื่องจากฟังก์ชันภาวะน่าจะเป็นมีเทอมของอินทิกรัลของฟังก์ชันข้างต้นมีความซับซ้อน จึงประมาณฟังก์ชันภาวะน่าจะเป็นขึ้นมาก่อนที่จะใช้วิธีภาวะน่าจะเป็นสูงสุดในการประมาณพารามิเตอร์ (Agresti, 2002; Pongsapukdee and Phonork, 2011) และใช้วิธีที่ได้รับความนิยมสำหรับตัวแบบ GLMMs คือ Adaptive Gaussian Quadrature Approach (AGQ) ในการประมาณค่าพารามิเตอร์ของการจำลองแบบต่อไป

วิธี AGQ ซึ่งปรับปรุงจากวิธี GQ (Gauss quadrature) หรือวิธี Gauss-Hermite quadrature (GH) เสนอโดย Liu and Pierce (1994) เป็นวิธีประมาณอินทิกรัล โดยการหาผลรวมแบบถ่วงน้ำหนัก เมื่อกำหนดจำนวน Quadrature point ในแต่ละมิติของการอินทิเกรต แต่มีจุดอ่อนคือผลรวมตลอดจำนวน Quadrature point มีค่ามาก โดยเฉพาะเมื่ออิทธิพลสุ่มมีจำนวนมากขึ้น ในการแก้ปัญหานี้จึงมีการพัฒนาเป็นวิธี AGQ

(Adaptive Gaussian Quadrature โดย Pinhero & Bate (1995) ซึ่งใช้จำนวน Quadrature point ในแต่ละมิติน้อยลงด้วยการปรับ Location และ Dispersion ของการแจกแจงที่จะอินทิเกรต ในการประมาณพารามิเตอร์ที่ทำให้ภาวะน่าจะเป็นมีค่าสูงสุดอาศัยวิธีนิวตัน-รัฟสัน (Newton-Raphson) หาค่าประมาณ $\hat{\boldsymbol{\beta}}$ และ $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}$ (Agresti, 2007) ในการประมาณความคลาดเคลื่อนมาตรฐาน อาจใช้จำนวน Quadrature points มากกว่าการประมาณค่า $\hat{\boldsymbol{\beta}}$ โดยเพิ่มจำนวน Quadrature points ขึ้นเรื่อย ๆ จนกระทั่งค่าเปลี่ยนแปลงลดน้อยลงมากทั้งตัวประมาณสัมประสิทธิ์ของตัวแปรและความคลาดเคลื่อนมาตรฐาน จะได้ค่าประมาณพารามิเตอร์ของตัวแบบ GLMMs ที่สนใจ

**การจำลองแบบ**

ในการสร้างตัวแบบ (1) และ (2) ประกอบด้วย ตัวแปรตอบสนองอันดับ (Y) มี 3 อันดับหรือ J=3, ภายใต้อิทธิพลสุ่มของ (Y, U) เมื่อ Y จำลองจากตัวแบบจริง และ U จำลองจาก $\mathrm{U} \sim N(0, \sigma_u^2)$ ตัวแปรอธิบาย 2 ตัวภายใต้อิทธิพลตรึงประกอบด้วย $X_1 \sim$Bernoulli (0.5) and $X_2 \sim$ Bernoulli (0.5) โดยที่ตัวแบบ (1) - (2) สร้างจากการจำลองแบบของข้อมูลเชตเดียวกัน ตัวแบบละ 1,000 ชุดในแต่ละเงื่อนไขของพารามิเตอร์ต่าง ๆ ได้แก่ จำนวนกลุ่ม (number of clusters, N = 50, 100, 250), ขนาดของกลุ่ม cluster sizes, $n_i$ = 5, 7, 10), และสหสัมพันธ์ภายในกลุ่ม (intra-cluster correlation, $ICC = \sigma_u^2 / \left(\sigma_u^2 + \sigma_\varepsilon^2\right)$) ซึ่งในงานวิจัยนี้คำนวณค่า ICC = 0.05, 0.20, 0.40 ที่สอดคล้องกับ $\sigma_u^2$ = 0.17, 0.82, 2.19 และ $\sigma_\varepsilon^2 = \pi^2 / 3$ ตามลำดับ ภายใต้ค่าพารามิเตอร์ของตัวแบบคือ $\alpha_1$ = -1, $\alpha_2$ =1, $\beta_1$ = 0.5 และ $\beta_2$ = 0.5 ซึ่งมีค่าใกล้เคียงกันกับของ Spiess and Hamerle (2000) การจำลองแบบทั้งหมดทำพร้อมกันภายใต้โปรแกรมแมโครที่เขียนให้สอดคล้องและประมวลผลร่วมกับโปรแกรม SAS version 9.1.

การประเมินผลและการวิเคราะห์ตัวแบบ พิจารณาจากผลต่างของค่าเฉลี่ยความไวของตัวแบบที่วัดด้วยพื้นที่ใต้โค้ง ROC ($\leq 1$) และกำลังการทดสอบ ($\leq 1$) ทีละอย่าง โดยค่าบวกหมายถึงตัวแบบผสมเชิงเส้นนัยทั่วไปลอจิตสะสมดีกว่าตัวแบบผสมเชิงเส้นนัยทั่วไปโพรบิตสะสม และค่าสูงหมายความว่าตัวแบบแรกมีประสิทธิภาพสูงกว่าและให้ค่าพยากรณ์แม่นกว่าด้วย (Agresti, 2013) ส่วนการทดสอบภาวะสารูปดีของตัวแบบที่ระดับนัยสำคัญ 0.05 อาศัยตัวสถิติทดสอบอัตราส่วนภาวะน่าจะเป็น (likelihood ratio statistic) ของฟังก์ชันภาวะน่าจะเป็น (3) ภายใต้ตัวแบบที่สนใจศึกษา ดังนั้นผลลัพธ์เชิงเปรียบเทียบระหว่างตัวแบบผสมเชิงเส้นนัยทั่วไปแบบลอจิตสะสมกับตัวแบบผสมเชิงเส้นนัยทั่วไปแบบโพรบิตสะสมจึงสรุปจาก 3 ประเด็นได้แก่ พื้นที่ใต้โค้ง ROC กำลังการทดสอบและภาวะสารูปดีเชิงสถิติด้วย

## ผลการวิจัย

ผลลัพธ์เชิงเปรียบเทียบด้วยพื้นที่ใต้โค้ง ROC ระหว่างตัวแบบผสมเชิงเส้นนัยทั่วไปลอจิตสะสมและตัวแบบผสมเชิงเส้นนัยทั่วไปโพรบิตสะสม อาศัยผลต่างค่าเฉลี่ยพื้นที่ใต้โค้ง และผลต่างค่าเฉลี่ยกำลังการทดสอบระหว่างทั้งสองตัวแบบที่ละอย่าง โดยจำแนกตามพารามิเตอร์สหสัมพันธ์ภายในกลุ่ม (ICC = 0.05, 0.20, 0.40); จำนวนกลุ่ม (N = 50, 100, 250) และขนาดของกลุ่ม ($n_i$ = 5, 7, 10), ตามลำดับ พบว่ามีค่าเป็นบวกเกือบทุกค่า

โดยค่าสูงมาจากตัวแบบผสมเชิงเส้นนัยทั่วไปลอจิตสะสมเกือบทั้งหมด โดยผลลัพธ์เชิงเปรียบเทียบพื้นที่ใต้โค้ง ROC ที่มากว่านั้นมีมากกว่า 1% อย่างมีเหตุและผลและสอดคล้องกัน (max 0.8%, min 0.1% ตารางที่ 1) ค่าเฉลี่ยพื้นที่ใต้โค้ง ROC เพิ่มขึ้นจากน้อยไปมากและเข้าใกล้ 1 เมื่อจำนวนกลุ่ม (N) และสหสัมพันธ์ภายในกลุ่ม (ICC) มีค่ามากขึ้น และมีค่าลดลงเมื่อขนาดของกลุ่ม(n) มากขึ้น โดยค่าเฉลี่ยพื้นที่ใต้โค้ง ROC อยู่ในช่วง (0.6830, 0.898) รายละเอียดเพิ่มเติมแสดงไว้ในตารางที่ 1

ผลลัพธ์เชิงเปรียบเทียบด้วยผลต่างค่าเฉลี่ยกำลังการทดสอบ ให้ผลสอดคล้องกับผลลัพธ์เชิงเปรียบเทียบด้วยผลต่างค่าเฉลี่ยพื้นที่ใต้โค้ง และสามารถให้ความชัดเจนมากขึ้นได้โดยพบว่าค่าเฉลี่ยกำลังการทดสอบส่วนใหญ่เข้าใกล้ 1 ได้มากกว่า เมื่อจำแนกตามค่าของ ICC, N และ n ยกเว้นเฉพาะกรณีเมื่อค่า N และ n มีค่าเล็ก ๆ ผลลัพธ์เชิงเปรียบเทียบค่ากำลังการทดสอบที่มากว่านั้นมีช่วงระหว่าง 9.7% ถึง 0.0% อย่างมีเหตุและผลและสอดคล้องกัน (max (9.7%), min (0.0%) ตารางที่ 1)

**ตารางที่ 1** ผลลัพธ์เชิงเปรียบเทียบด้วยพื้นที่ใต้โค้ง ROC (AUC) และกำลังการทดสอบ (Power) จำแนกตามสหสัมพันธ์ภายในกลุ่ม (ICC), จำนวนกลุ่ม (N), และขนาดของกลุ่ม (n)

| Intra-cluster correlation (ICC) | Number of Clusters (N) | Cluster size (n) | Gain in Area under ROC Curve | | Gain in Power of the tests | |
|---|---|---|---|---|---|---|
| | | | Average AUC Logit&Probit GLMMs | Effective Gain in AUC (%) | Average power Logit&Probit GLMMs | Effective Gain in Power (%) |
| 0.05 | 50 | 3 | 0.706 | 0.007(0.7) | 0.452 | 0.047(4.7%) |
| | | 5 | 0.686 | 0.007(0.7) | 0.700 | 0.057(5.7%) |
| | | 7 | 0.683 | 0.005(0.5) | 0.850 | 0.045(4.5%) |
| | 100 | 3 | 0.706 | 0.008(0.8) | 0.757 | 0.097(9.7%) |
| | | 5 | 0.684 | 0.007(0.7) | 0.941 | 0.046(4.6%) |
| | | 7 | 0.685 | 0.006(0.6) | 0.989 | 0.017(1.7%) |
| | 250 | 3 | 0.707 | 0.010(1.0) | 0.963 | 0.071(7.1%) |
| | | 5 | 0.695 | 0.008(0.8) | 0.990 | 0.020(2.0%) |
| | | 7 | 0.690 | 0.003(0.3) | 1.00 | 0.000(0.0%) |
| 0.2 | 50 | 3 | 0.821 | 0.004(0.4) | 0.427 | 0.002(0.2%) |
| | | 5 | 0.803 | 0.003(0.3) | 0.680 | 0.007(0.7%) |
| | | 7 | 0.788 | 0.002(0.2) | 0.839 | 0.001(0.1%) |
| | 100 | 3 | 0.832 | 0.004(0.4) | 0.736 | 0.009(0.9%) |
| | | 5 | 0.805 | 0.003(0.3) | 0.951 | 0.003(0.3%) |
| | | 7 | 0.789 | 0.002(0.2) | 0.987 | 0.001(0.1%) |
| | 250 | 3 | 0.840 | 0.003(0.3) | 0.990 | 0.002(0.2%) |
| | | 5 | 0.808 | 0.003(0.3) | 1.000 | 0.000(0.0%) |
| | | 7 | 0.791 | 0.002 (0.2) | 1.000 | 0.000(0.0%) |

**ตารางที่ 1** ผลลัพธ์เชิงเปรียบเทียบด้วยพื้นที่ใต้โค้ง ROC (AUC) และกำลังการทดสอบ (Power) จำแนกตามสหสัมพันธ์ภายในกลุ่ม (ICC), จำนวนกลุ่ม (N), และขนาดของกลุ่ม (n) (ต่อ)

| Intra-cluster correlation (ICC) | Number of Clusters (N) | Cluster size (n) | Gain in Area under ROC Curve | | Gain in Power of the tests | |
|---|---|---|---|---|---|---|
| | | | Average AUC Logit&Probit GLMMs | Effective Gain in AUC (%) | Average power Logit&Probit GLMMs | Effective Gain in Power (%) |
| 0.4 | 50 | 3 | 0.896 | 0.002(0.2) | 0.355 | -0.001 (-0.1%) |
| | | 5 | 0.868 | 0.001(0.1) | 0.594 | 0.001(0.1%) |
| | | 7 | 0.851 | 0.001(0.1) | 0.762 | 0.001(0.1%) |
| | 100 | 3 | 0.898 | 0.002(0.2) | 0.645 | -0.006 (-0.6%) |
| | | 5 | 0.868 | 0.001(0.1) | 0.903 | 0.00 (0.0%) |
| | | 7 | 0.852 | 0.001(0.1) | 0.975 | 0.00 (0.0%) |
| | 250 | 3 | 0.898 | 0.002(0.2) | 0.970 | 0.003(0.3%) |
| | | 5 | 0.868 | 0.001(0.1) | 1.000 | 0.00(0.0%) |
| | | 7 | 0.854 | 0.001(0.1) | 1.000 | 0.00 (0.0%) |
| | | | max(0.898) | max(0.8%) | max (1.00) | max(9.7%) |
| | | | min(0.683) | min (0.1%) | min(0.355) | min(0.0%) |

## วิจารณ์ผลการวิจัย

จากผลการวิจัยพบว่า ภายใต้ตัวแบบ GLMMs เมื่อตัวแปรตอบสนองเป็นแบบอันดับมากกว่าสองกลุ่ม และเมื่อใช้ตัวแปรอธิบายแบบจำแนกประเภทและเชิงกลุ่มทั้งหมด พบว่าตัวแบบลอจิตสะสม GLMMs และตัวแบบโพรบิตสะสม GLMMs ต่างมีความไวหรือพื้นที่ใต้โค้ง ROC สูง และตัวแบบลอจิตสะสม GLMMs มีประสิทธิภาพหรือมีกำลังการทดสอบที่ดีกว่าตัวแบบโพรบิตแบบ GLMMs อย่างชัดเจน ซึ่งสอดคล้องกับกรณีของตัวแบบ GLMs เมื่อใช้ตัวแปรอธิบายแบบจำแนกประเภทและเชิงกลุ่มทั้งหมด ตัวแบบลอจิตสะสม GLMs มีประสิทธิภาพสูงกว่าตัวแบบโพรบิตสะสม GLMs (Lipsitz, KIM, and Zhao, 1994; Agresti, 2007) รวมถึงเมื่อใช้ตัวแปรอธิบายแบบผสมที่มีทั้งแบบจำแนกประเภท เชิงกลุ่ม และแบบต่อเนื่อง ก็พบว่า ตัวแบบลอจิต GLMs มีประสิทธิภาพมากกว่าตัวแบบฟัสสิ (fuzzy logic models, Tang and Chi, 2005). อย่างไรก็ตาม ในกรณีถ้าใช้ตัวแปรอธิบายแบบต่อเนื่องทั้งหมด หรือเป็นแบบผสมที่มีแบบจำแนกประเภทบ้าง พบว่าส่วนใหญ่ตัวแบบโพรบิต GLMs มีประสิทธิภาพมากกว่าตัวแบบลอจิต GLMs (Hosmer and Lemeshow, 1989; Pongsapukdee and Sukgumphaphan, 2008, Agresti, 2013) ดังนั้นการนำตัวแบบ GLMs และ GLMMs ไปใช้ จึงควรเลือกตัวแบบตามความเหมาะสมในแต่ละกรณีต่อไปด้วย

## สรุปผลการวิจัย

ผลลัพธ์เชิงเปรียบเทียบด้วยพื้นที่ใต้โค้ง ROC และกำลังการทดสอบระหว่างตัวแบบผสมเชิงเส้นนัยทั่วไปแบบลอจิตสะสมและตัวแบบผสมเชิงเส้นนัยทั่วไปแบบโพรบิตสะสม พบว่าเมื่อจำนวนกลุ่ม (N) และสหสัมพันธ์ภายในกลุ่ม (ICC) มีค่ามากขึ้น ตัวแบบทั้งสองมีประสิทธิภาพมากขึ้น โดยเฉพาะตัวแบบผสมเชิงเส้นนัยทั่วไปแบบลอจิตสะสมมีกำลังการทดสอบดีกว่าของตัวแบบผสมเชิงเส้นนัยทั่วไปแบบโพรบิตสะสมอย่างชัดเจน แต่เมื่อขนาดของกลุ่ม (n) เพิ่มมากขึ้น ประสิทธิภาพมีค่าลดลงบ้าง แต่ทั้งสองตัวแบบยังมีภาวะสารูปดีอย่างมีนัยสำคัญทางสถิติที่ 0.05 และมีความไวและประสิทธิภาพค่อนข้างสูงและใกล้เคียงกัน จึงอาจนำไปประยุกต์เพื่อการวิเคราะห์ข้อมูลจริงต่อไป โดยสามารถใช้โปรแกรมสำเร็จรูปทางสถิติช่วยในการประมวลผลโดยตรง

## เอกสารอ้างอิง


วีรานันท์ พงศาภักดี. (2555). การวิเคราะห์ข้อมูลจำแนกประเภท: ทฤษฎีและการประยุกต์ ด้วย GLIM, SPSS, SAS และ MTB พิมพ์ครั้งที่ 3 นครปฐม: โครงการตำราคณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร โรงพิมพ์มหาวิทยาลัยศิลปากร. หน้า. 165–196, 369-396, 419-454.

Agresti, A. (2002). Categorical Data Analysis. 2$^{nd}$ ed. New York: John Wiley & Sons. pp. 267–313.

Agresti, A (2007). Introduction to Categorical Data Analysis. New York: John Willey &Sons. pp. 173-179.

Agresti, A. (2013) Categorical Data Analysis. 3$^{nd}$ ed. New York: John Wiley & Sons. pp. 341-352, 455-525.

Breslow N. E. & Clayton, D. G. (1993). Approximate inference in generalized linear mixed model. *Journal of the American Statistical Association*, 88(421), 9-25; DOI: 10.2307/2290687.

Green, P.J. (1987). Penalized likelihood for general semi-parametric regression models. *Int. Statist Rev.;* 55(3), 245-260; DOI: 10.2307/1403404.

Goldstein, H. (1991). Nonlinear multilevel models with an application to discrete response data. *Biometrika*, 78(1), 45-51, DOI:10.2307/2336894.

Goldstein, H. (1995). *Multilevel Statistical Models.* 2nd Ed. Edward Arnold, London.

Hosmer, D, W. and Lemeshow, S. (1989). Applied Logistic Regression. New York: John Wiley & Sons. pp. 142-143.

Liang, K.Y. and Zeger, S.L. (1986) Longitudianl data analysis using generalized linear models. *Biometrika*; 73(1), 13-22.

Lipsitz, S.R., KIM, K. and Zhao, L. (1994). Analysis of repeated categorical data using generalized estimating equations. *Statist. Medic*. 13(11), 1149-1163; DOI: 10.1002/sim.4780131106.

Liu, Q. and Pierce, D.A. (1994). A note on Guass-Hermite quadrature. *Biometrika*; 81(3), 624-629; DOI: 10.2307/2337136,

McCullagh, P. (1983). Quasi-likelihood functions. Ann. Statist. 11, pp. 59-67.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*. 2nd Edn., London: Chapman and Hall. pp. 149-174.

McCulloch, C.E. and Searle, S. (2001). *Generalized, Linear, and Mixed Models*., New York: John Wiley and Sons. pp. 135-177.

Nelder, J.A. and Wedderburn, R.W.M. (1972). Generalized Linear Models. J. Roy. Statist. Soc. Ser A135(3): 370-384.

Pinheiro, J.C. and Bate, D.M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects models. *Journal of Computational and Graphical Statistics*, 4(1), 12-35; DOI: 10.2307/1390625,.

Pongsapukdee, V. and Sukgumphaphan, S. (2008). Three cumulative link models for ordinal responses: A review of methods and power comparisons. *Songklanakarin J. of Sci. Technol*; 30(6), 805-811.

Pongsapukdee, V. and Phonork, C. (2011). A Diagnosis and Prognosis for Generalized Linear Mixed Models with Correlated Ordinal Responses: Power and Area under ROC Curve. Modern Applied Science, Vol. 5, No. 5: 29-38; DOI: 10.5539/mas.v5n5p29.

Spiess, M. and Hamerle, A., A comparison of different methods for the estimation of regression models with correlated binary responses. *Computational Statistics and Data Analysis*. 33(4), 439-455, 2000; DOI: 10.1016/S0167-9473(99)00065-1. .

Tang, T.C. and Chi, L.C. (2005). Predicting multilateral trade credit risks: Comparisons of logit and fuzzy logic models using ROC curve analysis, *Expert Systems with Applications*, 28(3), 547-556; DOI: 10.1016/j.eswa.12.016.

❑❑❑❑❑