วารสารอิเล็กทรอนิกส์
ทางการศึกษา

OJED, Vol. 12, No. 2, 2017, pp. 1 - 17

OJED
An Online Journal
of Education
http://www.edu.chula.ac.th/ojed

ความแม่นยำของการประมาณค่าพารามิเตอร์ในการวิเคราะห์องค์ประกอบเชิงยืนยัน ภายใต้โมเดลที่มีการระบุผิดพลาด

ACCURACY OF PARAMETER ESTIMATIONS IN CONFIRMATORY FACTOR ANALYSIS UNDER MISSPECIFIED MODEL

นายกันตินันท์ นันทนาดิศัย*

Kantinan Nantanadisai

ดร.สิวะโชติ ศรีสุทธิยากร**

Siwachoat Srisuttiyakorn, Ph.D.

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์คือเปรียบเทียบความแม่นยำของการประมาณค่าพารามิเตอร์ และวิเคราะห์ผลกระทบของปัจจัย ในการวิเคราะห์องค์ประกอบเชิงยืนยัน ด้วยวิธีการ 4 วิธี ได้แก่ วิธีภาวะความควรจะเป็นสูงสุด (ML) วิธีภาวะความควรจะเป็นสูงสุดที่มีความแกร่ง (RML) วิธีกำลังสองน้อยสุดทั่วไป (GLS) และวิธีกำลังสองน้อยสุดถ่วงน้ำหนัก (WLS) ภายใต้สถานการณ์ที่มีการระบุพารามิเตอร์ความแปรปรวนร่วมของความคลาดเคลื่อนจากการวัดผิดพลาด และตัวแปรสังเกตได้มีการแจกแจงไม่ปกติ การวิจัยใช้ระเบียบวิจัยเชิงทดลองโดยใช้ข้อมูลจำลองด้วยเทคนิคมอนติคาร์โล ภายใต้ 20 เงื่อนไขการจำลองจาก 3 ปัจจัย ประกอบด้วย (1) ระดับการระบุโมเดลผิดพลาด 3 ระดับ คือ ต่ำ (RMSEA=.02) ปานกลาง (RMSEA=.04) และสูง (RMSEA=.06) (2) การแจกแจงความน่าจะเป็นของตัวแปรสังเกตมี 3 ระดับ คือ แบนราบกว่าปกติระดับน้อย (ku=-1) โด่งกว่าปกติระดับน้อย (ku=1) และโด่งกว่าปกติระดับมาก (ku=2) (3) ขนาดตัวอย่างเท่ากับ 200, 400, 600 และ 800 หน่วยตัวอย่าง โดยแต่ละเงื่อนไขการจำลองจะทำซ้ำ 1000 รอบ และใช้ค่าความเอนเอียงสัมพัทธ์ (RB) เป็นเกณฑ์การเปรียบเทียบความแม่นยำของการประมาณค่าพารามิเตอร์

ผลการวิจัยพบว่า (1) วิธี ML และ RML มีความเอนเอียงสัมพัทธ์ไม่แตกต่างกัน และมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี GLS และ WLS (2) ปัจจัยที่ส่งผลกระทบต่อความเอนเอียงสัมพัทธ์ของค่าประมาณพารามิเตอร์อย่างมีนัยสำคัญทางสถิติ (p-value= .000) เรียงตามขนาดอิทธิพลได้แก่ ประเภทของพารามิเตอร์ วิธีการประมาณค่าพารามิเตอร์ ระดับของการระบุโมเดลผิดพลาด และขนาดตัวอย่าง ขณะที่ระดับความโด่งของการแจกแจงความน่าจะเป็นของตัวแปรสังเกตมีไม่มีผลกระทบต่อค่า RB ของค่าประมาณพารามิเตอร์อย่างมีนัยสำคัญทางสถิติ (p-value= .237)

* นิสิตมหาบัณฑิตสาขาสถิติการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
E-mail Address: kantinan.yen@outlook.co.th
**อาจารย์ประจำสาขาวิชาสถิติการศึกษา ภาควิชาวิจัยและจิตวิทยา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
E-mail Address: choat.cu@gmail.com
ISSN1905-4491

Abstract

This research aimed to compare the accuracy of parameter estimation methods and effects of factors in confirmatory factor analysis with four parameter estimation methods including maximum likelihood (ML), robust maximum likelihood (RML), generalized least squares (GLS), and weighted least squares (WLS) when measurement error covariances are misspecified and observed variables are abnormally distributed. Data was generated by using the Monte Carlo technique under 20 simulation conditions from three factors consisting of (1) misspecification level included low (RMSEA=.02), moderate (RMSEA=.04), and high (RMSEA=.06) (2) kurtosis level of the observed variables distribution included low platykurtic (ku=-1), moderate leptokurtic (ku=1) and high leptokurtic (ku=2) and (3) sample sizes included 200, 400, 600 and 800 units. Each condition was repeated 1000 times. The relative bias (RB) was used to compare the accuracy of the estimation methods.

The results were: 1) ML and RML methods were equa tol RB and less RB than GLS and WLS methods; and 2) the factors that significantly affected RB of the estimated parameters (p-value= .000) were sorted by their effect size by the type of parameters, estimation methods, misspecification level, and sample size, while the kurtosis of the observed variable distribution did not significantly affect RB of tje estimated parameters (p-value= .237).

คำสำคัญ: การประมาณค่าพารามิเตอร์/ การวิเคราะห์องค์ประกอบเชิงยืนยัน/ โมเดลระบุผิดพลาด

KEYWORDS: PARAMETERS ESTIMATION/ CONFIRMATORY FACTOR ANALYSIS/ MODEL MISSPECIFICATION

บทนำ

โมเดลสมการโครงสร้างเป็นสถิติวิเคราะห์ที่มีประสิทธิภาพสูงสามารถใช้ทดสอบเพื่อยืนยันความตรงของโมเดลวิจัย ประมาณค่าขนาดอิทธิพลของตัวแปรอิสระที่มีต่อตัวแปรตาม ทั้งยังเป็นสถิติที่มีการขจัดปัญหาอันเกิดจากความคลาดเคลื่อนจากการวัดในตัวแปรสังเกตได้ และยอมให้ความคลาดเคลื่อนต่าง ๆ ในโมเดลมีความสัมพันธ์กันได้ จึงทำให้ผลการวิเคราะห์ที่ได้มีความแม่นยำและถูกต้องมากกว่าการใช้สถิติวิเคราะห์ความสัมพันธ์เชิงสาเหตุแบบดั้งเดิม (Jöreskog & Sörbom, 1996; Muthén & Muthén, 2010; นงลักษณ์ วิรัชชัย, 2542)

ในงานวิจัยที่ทำการศึกษาความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรด้วยโมเดลสมการโครงสร้างนั้น ผู้วิจัยจะต้องเริ่มด้วยการพัฒนาโมเดลวิจัย (research model) จากการสังเคราะห์แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง เนื่องจากโมเดลการวิจัยก็คือสมมติฐานการวิจัยดังนั้นนักวิจัยจึงมีความจำเป็นที่จะต้องตรวจสอบเพื่อยืนยันว่าโมเดลการวิจัยสามารถใช้เป็นตัวแทนความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรในประชากรได้ การตรวจสอบทำได้โดยการพิจารณาความแตกต่างระหว่างโมเดลวิจัยกับข้อมูลเชิงประจักษ์ที่สุ่มจากประชากร

เป้าหมาย และเรียกคุณสมบัติของโมเดลวิจัยที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์ว่าคุณสมบัติความสอดคล้องเชิงประจักษ์ (empirical fit) (Olsson, Foss, Troye, & Howell, 2000)

การประเมินความสอดคล้องเชิงประจักษ์ของโมเดลวิจัยด้วยการวิเคราะห์โมเดลสมการโครงสร้างทำได้โดยการเปรียบเทียบความแตกต่างระหว่างเมทริกซ์ความแปรปรวนร่วมสองเมทริกซ์ได้แก่ เมทริกซ์ความแปรปรวนร่วมระหว่างตัวแปรที่ศึกษาภายใต้โมเดลวิจัย (model-implied covariance matrix $\Sigma(\theta)$) กับเมทริกซ์ความแปรปรวนร่วมของตัวอย่าง (sample covariance matrix: S) เมื่อ θ คือค่าพารามิเตอร์ในโมเดลวิจัย การพิจารณาความแตกต่างระหว่างเมทริกซ์ทั้งสองจะใช้ฟังก์ชันความแตกต่าง (discrepancy function) เป็นเกณฑ์ในการพิจารณา โดยค่าของฟังก์ชันความแตกต่างที่ได้จากการคำนวณนั้นจะถูกนำไปใช้ประโยชน์ทั้งการหาค่าประมาณพารามิเตอร์ในโมเดล และการตรวจสอบความสอดคล้องเชิงประจักษ์ของโมเดลวิจัย (Jöreskog & Sörbom, 1996; Muthén & Muthén, 2010)

ฟังก์ชันความแตกต่างที่นิยมใช้ในปัจจุบันมีหลายฟังก์ชัน เช่น ฟังก์ชันภาวะความควรจะเป็น (likelihood function) ฟังก์ชันกำลังสองน้อยสุดทั่วไป (generalized least square) และฟังก์ชันกำลังสองน้อยสุดถ่วงน้ำหนัก (weighted least square) เป็นต้น ฟังก์ชันแต่ละฟังก์ชันยังมีประสิทธิภาพแตกต่างกัน ฟังก์ชันความแตกต่างที่นักวิจัยนิยมใช้คือฟังก์ชันภาวะความควรจะเป็น และเรียกวิธีการประมาณที่ใช้ฟังก์ชันนี้ว่า วิธีภาวะความควรจะเป็นสูงสุด (maximum likelihood function) ในทางทฤษฎีถ้าข้อมูลมีคุณสมบัติเป็นไปตามข้อกำหนดเบื้องต้นแล้วค่าประมาณที่ได้จากวิธีภาวะความควรจะเป็นสูงสุดจะมีคุณสมบัติที่ดีเช่น คุณสมบัติความไม่ลำเอียงเมื่อใกล้อนันต์ (asymptotical unbiased) และความมีประสิทธิภาพเมื่อใกล้อนันต์ (asymptotical efficient) และมีความคงเส้นคงวา (consistent) เป็นต้น แต่หากข้อมูลไม่เป็นไปตามข้อกำหนดเบื้องต้นแล้วประสิทธิภาพของการประมาณค่าจะเปลี่ยนแปลงไป (สิวะโชติ ศรีสุทธิยากร, 2555; Jöreskog & Sörbom, 1996; Muthén & Muthén, 2010)

ปัจจัยที่ส่งผลต่อประสิทธิภาพของการประมาณค่ามีอยู่หลายปัจจัย เช่น รูปแบบและความซับซ้อนของโมเดลวิจัย ขนาดตัวอย่าง การแจกแจงความน่าจะเป็นของตัวแปรสังเกตได้ รวมไปถึงรูปแบบและระดับของการระบุโมเดลผิดพลาด (Xiatao Fan & Lin Wang, 1998; Xitao Fan & Stephen A. Sivo, 2007, Pornprasertmanit, 2014) โดยการตรวจสอบระดับของการระบุโมเดลผิดพลาดในการวิเคราะห์โมเดลสมการโครงสร้างสามารถทำได้ผ่านการตรวจสอบความสอดคล้องเชิงประจักษ์ของโมเดลวิจัย ซึ่งมีหลายวิธีการ เช่น ดัชนีไคกำลังสองสัมพัทธ์ (relative chi-square) ดัชนีรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณค่า (root mean square error of approximation: RMSEA) เป็นต้น

ที่ผ่านมามีงานวิจัยที่ได้ศึกษาประสิทธิภาพของการประมาณค่าพารามิเตอร์ และดัชนีวัดความสอดคล้องเชิงประจักษ์ภายใต้สถานการณ์ที่โมเดลวิจัยมีการระบุผิดพลาด และตัวแปรสังเกตได้มีการแจกแจงไม่ปกติ พบว่าวิธีประมาณค่าพารามิเตอร์ส่วนใหญ่ที่นิยมศึกษามี 4 วิธีคือ วิธีภาวะความควรจะเป็นสูงสุด (maximum likelihood: ML) วิธีกำลังสองน้อยสุดทั่วไป (generalized least squares: GLS) วิธีกำลังสองน้อยสุดถ่วงน้ำหนัก (weighted least squares: WLS) และวิธีภาวะความควรจะเป็นสูงสุดที่มีความแกร่ง

(robust maximum likelihood: RML) ผลจากการศึกษาพบว่า (1) ภายใต้สถานการณ์ที่ตัวแปรสังเกตได้มีการแจกแจงไม่ปกติ วิธีกำลังสองน้อยสุดถ่วงน้ำหนักในกรณีที่มีขนาดตัวอย่างใหญ่ และวิธีภาวะความควรจะเป็นสูงสุดที่มีความแกร่ง มีประสิทธิภาพมากกว่าวิธีภาวะความควรจะเป็นสูงสุด และวิธีกำลังสองน้อยสุดทั่วไป แต่ยังไม่มีการเปรียบเทียบประสิทธิภาพระหว่างวิธีกำลังสองน้อยสุดถ่วงน้ำหนัก และวิธีภาวะความควรจะเป็นสูงสุดที่มีความแกร่งในกรณีที่มีขนาดตัวอย่างใหญ่ (Curran, 1994; Jeri Benson & John A. Fleishman, 1994; Curran, West, & Finch, 1996; Olsson, Foss, Troye, & Howell, 2000; Anne Boomsma & Jeffrey J. Hoogland, 2001; Ming Lei & Richard G. Lomax, 2005; Youngkyoung Min, 2008) (2) ภายใต้สถานการณ์ที่โมเดลวิจัยระบุผิดพลาดและตัวแปรสังเกตได้มีการแจกแจงไม่ปกติ พบว่าวิธีกำลังสองน้อยสุดถ่วงน้ำหนัก และวิธีกำลังสองน้อยวิธีภาวะความควรจะเป็นสูงสุดที่มีความแกร่งจะมีค่าน้อยลงกว่าค่าที่ควรจะเป็น (underestimated) แต่งานวิจัยส่วนใหญ่จะศึกษาระดับการระบุโมเดลผิดพลาดโดยใช้จำนวนพารามิเตอร์ที่ขาดหรือเกิน ซึ่งไม่มีเกณฑ์ที่แน่นอน ทำให้ในการนำผลวิจัยไปใช้จริงเป็นไปได้ยาก

จากประเด็นที่กล่าวมาข้างต้นพบว่าประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ภายใต้สถานการณ์ที่โมเดลระบุผิดพลาดและตัวแปรสังเกตได้มีการแจกแจงไม่ปกติยังไม่มีการศึกษาที่ให้ข้อสรุปที่ชัดเจน ดังนั้นการวิจัยในครั้งนี้จึงสนใจขยายการศึกษาประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์ รวมถึงปัจจัยที่ส่งผลกระทบต่อความเอนเอียงในการประมาณค่าพารามิเตอร์ เพื่อที่ผู้วิจัยจะได้สามารถเลือกใช้วิธีประมาณค่าที่มีประสิทธิภาพมากที่สุดในแต่ละสถานการณ์ โดยขอบเขตการวิจัยครั้งนี้จะสนใจศึกษาในโมเดลวิเคราะห์องค์ประกอบเชิงยืนยัน เนื่องจากการวิจัยทางการศึกษามักจะศึกษาเกี่ยวกับการสร้างเครื่องมือวัดความตรงของตัวแปรที่ใช้ในการวิจัยโดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยันในการวัดความตรงเชิงโครงสร้างของเครื่องมือ

วัตถุประสงค์

1. เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันภายใต้สถานการณ์ที่ไม่มีการระบุโมเดลผิดพลาด และมีการระบุโมเดลผิดพลาดโดยที่ตัวแปรสังเกตได้มีการแจกแจงไม่ปกติ
2. เพื่อวิเคราะห์ผลกระทบของปัจจัยระดับการระบุโมเดลผิดพลาด วิธีประมาณค่าพารามิเตอร์ รูปทรงการแจกแจงของข้อมูล ขนาดตัวอย่าง และประเภทของพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบที่ส่งผลต่อความเอนเอียงสัมพัทธ์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง (experimental research) โดยใช้แผนแบบการทดลองแบบแฟกทอเรียล (432 factorial design) ข้อมูลที่ใช้ในการศึกษาได้จากการจำลองด้วยวิธีการ มอนติคาร์โล (Monte Carlo simulation) โดยผู้วิจัยได้บางวิธีดำเนินการวิจัยออกเป็น 3 ส่วน ดังนี้

1. เงื่อนไขการจำลองข้อมูล

โมเดลประชากรที่ใช้ในการศึกษา คือโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis: CFA) โดยมีสมการและพารามิเตอร์ที่ใช้ในการศึกษาเป็นดังนี้

$$y_i = \Lambda \P_i + \¢_i \quad (1)$$

โดยที่ y_i คือเวกเตอร์ของตัวแปรสังเกตได้ (observed variable) ขนาด p x 1, Λ คือเมทริกซ์ของน้ำหนักองค์ประกอบ (factor loadings matrix) ขนาด p x q, $\P_i$ คือเวกเตอร์ของตัวแปรแฝง (latent variable) ขนาด q x 1 และ $\¢_i$ คือเวกเตอร์ของความคลาดเคลื่อนจากการวัด (measurement errors) ขนาด p x 1 รายละเอียดของการกำหนดค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันเป็นดังนี้

1. กำหนดให้มีตัวแปรแฝงในการศึกษาจำนวน 3 ตัวแปร ได้แก่ $\P_1, \P_2, \P_3$ (q=3)
2. กำหนดให้มีตัวแปรสังเกตได้ในการศึกษาจำนวน 12 ตัวแปร ได้แก่ $x_1, x_2, ..., x_{12}$ (p=12) โดยแต่ละตัวแปรแฝงจะประกอบด้วยตัวแปรสังเกตได้จำนวน 4 ตัวแปร
3. เมทริกซ์ของน้ำหนักองค์ประกอบขนาด 12 x 3 กำหนดให้มีค่าเท่ากับ

$$\Lambda_x(transposed) = \begin{bmatrix} 0.7 & 0.7 & 0.7 & 0.7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.7 & 0.7 & 0.7 & 0.7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7 & 0.7 & 0.7 & 0.7 \end{bmatrix}$$

4. เมทริกซ์ความแปรปรวนร่วมของตัวแปรแฝงขนาด 3 x 3 กำหนดให้มีค่าเท่ากับ

$$¤ = \begin{pmatrix} 1 & 0.3 & 0.3 \\ 0.3 & 1 & 0.3 \\ 0.3 & 0.3 & 1 \end{pmatrix}$$

5. เมทริกซ์ความแปรปรวนร่วมของความคลาดเคลื่อนจากการวัดขนาด 12 x 12 กำหนดให้มีค่าเท่ากับ

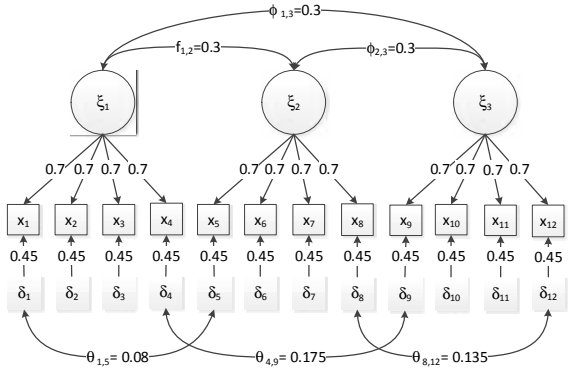
$$\theta_\delta = \begin{pmatrix} 0.45 & & & & & & & & & & & \\ 0 & 0.45 & & & & & & & & & & \\ 0 & 0 & 0.45 & & & & & & & & & \\ 0 & 0 & 0 & 0.45 & & & & & & & & \\ 0.08 & 0 & 0 & 0 & 0.45 & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0.45 & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.45 & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.45 & & & & \\ 0 & 0 & 0 & 0.175 & 0 & 0 & 0 & 0 & 0.45 & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.45 & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.135 & 0 & 0 & 0 & 0.45 & \end{pmatrix}$$

6. ในการศึกษาครั้งนี้จะศึกษาข้อมูลจากการจำลอง ภายใต้สถานการณ์ที่ตัวแปรสังเกตได้ไม่มีการแจกแจงปกติ ซึ่งระดับของการแจกแจงจะศึกษาถึงอิทธิพลของสัมประสิทธิ์ความโด่งเพียงอย่างเดียว เนื่องจากในการวิเคราะห์องค์ประกอบเชิงยืนยันจะใช้เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมในการวิเคราะห์ ซึ่งความโด่งมีอิทธิพลต่อการทดสอบนัยสำคัญทางสถิติที่ใช้ความแปรปรวน-ความแปรปรวนร่วมในการวิเคราะห์มากกว่าความเบ้ เนื่องจากความเบ้เป็นปัจจัยที่ส่งผลต่อค่าเฉลี่ย (Mardia, 1974; Olsson, Foss, Troye, & Howell, 2000) ดังนั้นจึงกำหนดให้ค่าสัมประสิทธิ์ความเบ้คงที่ โดยมีค่าเท่ากับ 0 และกำหนดค่าสัมประสิทธิ์ความโด่งแตกต่างกัน 3 ระดับ คือ ku=-1, ku=1, และ ku=2 ตามลำดับ

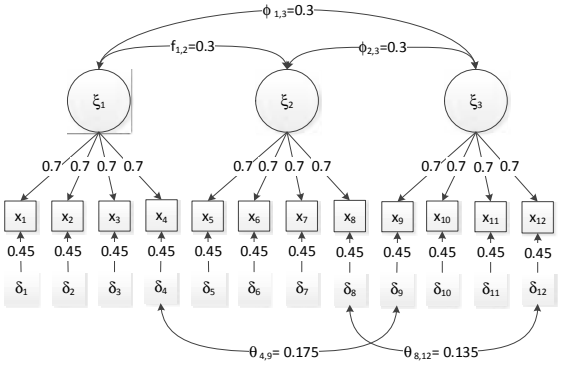
2. สถานการณ์ที่ใช้ในการศึกษา

แบบแผนการจำลองในครั้งนี้ประกอบไปด้วยตัวแปรต้น 5 ตัวแปร ประกอบด้วย (1) ประเภทของพารามิเตอร์ 3 ประเภท ได้แก่ ได้แก่ พารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อนจากการวัด (measurement error covariances) พารามิเตอร์น้ำหนักองค์ประกอบ (factor loading) และพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝง (latent variable covariances) (2) รูปทรงการแจกแจงของตัวแปร ในการศึกษาครั้งนี้จะศึกษาถึงอิทธิพลของสัมประสิทธิ์ความโด่งเพียงอย่างเดียว เนื่องจากในการวิเคราะห์องค์ประกอบเชิงยืนยันจะใช้เมทริกซ์ความแปรปรวนร่วมในการวิเคราะห์ ซึ่งความโด่งมีอิทธิพลต่อการทดสอบนัยสำคัญทางสถิติที่ใช้ความแปรปรวนในการวิเคราะห์ มากกว่าความเบ้ซึ่งเป็นปัจจัยที่ส่งผลต่อค่าเฉลี่ย (Mardia, 1974; Olsson et al., 2000) ดังนั้นจึงกำหนดให้ค่าสัมประสิทธิ์ความเบ้คงที่ โดยมีค่าเท่ากับ 0 และจากการศึกษางานวิจัยที่เกี่ยวข้องทางสังคมศาสตร์ที่ใช้การวิเคราะห์โมเดลสมการโครงสร้างและการวิเคราะห์องค์ประกอบเชิงยืนยัน จำนวน 50 งานวิจัยซึ่งมีตัวแปรสังเกตได้ทั้งหมด 955 ตัวแปร มีค่าสัมประสิทธิ์ความโด่งโดยเฉลี่ยเท่ากับ 0.157 โดยมีการกระจายตัวใกล้เคียงกับช่วง -1 ถึง 2 ดังนั้นจึงกำหนดให้ค่าสัมประสิทธิ์ความโด่งเท่ากับ -1, 1 และ 2 (3) ขนาดตัวอย่าง ที่ใช้ในการศึกษาครั้งนี้จะกำหนดจากวิธี The Rules of Thumb ซึ่งเป็นวิธีการกำหนดขนาดตัวอย่างจากสัดส่วนของพารามิเตอร์ที่ใช้ในการศึกษา โดยมีงานวิจัยที่พบว่าข้อมูลที่มีความโด่งผิดปกติควรจะใช้สัดส่วนของขนาดตัวอย่างต่อพารมมิเตอร์อย่างน้อยเป็น 10:1 (Hoogland and Boomsma, 1998) ผู้วิจัยจึงใช้ขนาดตัวอย่างเท่ากับ 10,15,20,25 เท่าของจำนวนพารามิเตอร์ ซึ่งในโมเดลที่ใช้ในการศึกษาครั้งนี้มีจำนวนพารามิเตอร์เท่ากับ 30 พารามิเตอร์ จะได้ขนาดตัวอย่างเท่ากับ 200, 300, 600, 750 จากนั้น ผู้วิจัยได้ทำการปรับให้ขนาดตัวอย่างในแต่ละช่วงมีช่วงห่างเท่าๆกันเพื่อความสะดวกในการพิจารณาแนวโน้มของผลการวิจัยและการนำไปใช้ ดังนั้น จะได้ขนาดตัวอย่างที่ใช้ในการศึกษาเท่ากับ 200, 400, 600, 800 (4) วิธีการประมาณค่าพารามิเตอร์ 4 วิธี ได้แก่ วิธีภาวะความควรจะเป็นสูงสุด (maximum likelihood: ML) วิธีภาวะความควรจะเป็นสูงสุดที่มีความแกร่ง (robust maximum likelihood: RML) วิธีกำลังสองน้อยสุดทั่วไป (generalized least squares: GLS) และวิธีกำลังสองน้อยสุดถ่วงน้ำหนัก (weighted least squares: WLS) (5) ระดับของการระบุโมเดลผิดพลาด โมเดลประชากรที่ใช้ในการศึกษาครั้งนี้ประกอบด้วยตัวแปรแฝง 3 ตัวแปร ตัวแปรสังเกตได้ 12 ตัวแปร และโมเดลที่มีการระบุผิดพลาด 4 ระดับ ได้แก่โมเดลระบุไม่ผิดพลาด (no misspecification) มีค่า RMSEA เท่ากับ 0.00 โมเดลระบุผิดพลาดระดับน้อย (small misspecification) มีค่า RMSEA เท่ากับ 0.02 โมเดลระบุผิดพลาดระดับปานกลาง (moderate misspecification) มีค่า RMSEA เท่ากับ 0.04 และโมเดลระบุผิดพลาดระดับมาก (severe misspecification) มีค่า RMSEA เท่ากับ 0.06 โดยจะกำหนดพารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อนที่แตกต่างจากโมเดลประชากรเป็นดังนี้ โมเดลระบุผิดพลาดระดับน้อยจะกำหนดให้พารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน $\theta_{1,5} = 0$ ดังรูปที่ 1 ข. โมเดลระบุผิดพลาดระดับปานกลางจะกำหนดให้พารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน $\theta_{1,5} = 0, \theta_{8,12} = 0$ ดังรูปที่

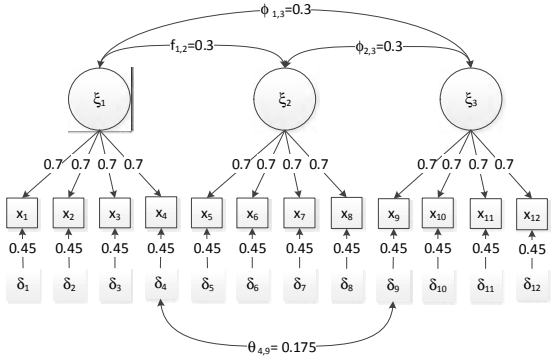
1 ค. และโมเดลระบุผิดพลาดระดับมากจะกำหนดให้พารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน $\theta_{1,5} = 0, \theta_{4,9} = 0, \theta_{8,12} = 0$ ดังรูป 1 ง.



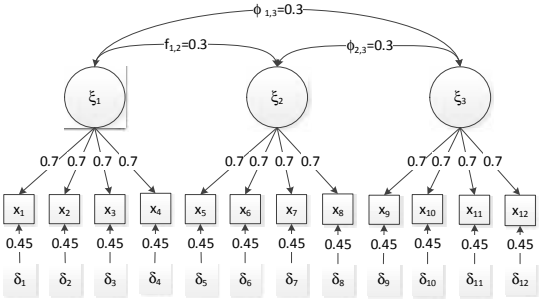
ก. โมเดลประชากร



ข. โมเดลระบุผิดพลาดระดับน้อย



ค. โมเดลระบุผิดพลาดระดับปานกลาง



ง. โมเดลระบุผิดพลาดระดับมาก

รูป 1: โมเดลประชากรและโมเดลที่มีการระบุผิดพลาดระดับต่าง ๆ ที่ใช้ในการศึกษา

3. เกณฑ์การพิจารณา

เกณฑ์การพิจารณาเพื่อเปรียบเทียบความแม่นยำของของวิธีประมาณค่าพารามิเตอร์ในการวิจัยครั้งนี้จะพิจารณาจากความเอนเอียงสัมพัทธ์ (relative bias : RB) เป็นเกณฑ์ที่ใช้วัดความเอนเอียงของพารามิเตอร์ $\bar{\hat{\theta}}_j$ ในการประมาณค่าพารามิเตอร์ประชากร $\bar{\theta}_j$ โดยมีสูตรการคำนวณเป็นดังนี้

$$RB\left(\hat{\theta}_j\right) = \frac{\left|\bar{\hat{\theta}}_j - \bar{\theta}_j\right|}{\bar{\theta}_j} \times 100, \; j = 1,2,3,\ldots,n$$

เมื่อ $\bar{\hat{\theta}}_j$ คือ ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ได้จากการจำลอง n รอบ

$\bar{\theta}_j$ คือ ค่าประมาณพารามิเตอร์รอบที่ j

ผลการวิจัย

ผลการวิจัยครั้งนี้แบ่งออกได้เป็น 2 ส่วน ได้แก่ 1) เปรียบเทียบความแม่นยำของการประมาณค่าพารามิเตอร์ 2) วิเคราะห์ผลกระทบของปัจจัยระดับการระบุโมเดลผิดพลาด วิธีประมาณค่าพารามิเตอร์ รูปทรงการแจกแจงของข้อมูล ขนาดตัวอย่าง และประเภทของพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบที่มีต่อของความแม่นยำการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันวิเคราะห์อิทธิพลของตัวแปร ระดับการระบุโมเดลผิดพลาด วิธีประมาณค่าพารามิเตอร์ รูปทรงการแจกแจงของข้อมูล ขนาดตัวอย่าง และประเภทของพารามิเตอร์ ที่ส่งผลต่อความเอนเอียงสัมพัทธ์

1. การเปรียบเทียบความแม่นยำของการประมาณค่าพารามิเตอร์

การเปรียบเทียบความแม่นยำของการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบจะพิจารณาจากค่าความเอนเอียงสัมพัทธ์ โดยจำแนกตามประเภทของพารามิเตอร์ทั้งสามในโมเดล ได้แก่ พารามิเตอร์น้ำหนักองค์ประกอบ ความแปรปรวนร่วมระหว่างตัวแปรแฝง และความแปรปรวนร่วมของความคลาดเคลื่อนจากการวัด มีรายละเอียดดังนี้

ความแม่นยำของการประมาณค่าพารามิเตอร์น้ำหนักองค์ประกอบ

จากตารางที่ 1 ความเอนเอียงสัมพัทธ์จากการประมาณค่าพารามิเตอร์น้ำหนักองค์ประกอบพบว่า 1) วิธีประมาณค่า ML และ RML มีความเอนเอียงสัมพัทธ์ไม่แตกต่างกัน และมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี GLS และ WLS ทุกขนาดตัวอย่าง ระดับการระบุโมเดลผิดพลาด และรูปทรงการแจกแจงของข้อมูล 2) วิธี GLS มีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์ลดลงเมื่อมีขนาดตัวอย่างเพิ่มขึ้น และมีความเอนเอียงสัมพัทธ์ใกล้เคียงกันทุกระดับการระบุโมเดลผิดพลาด โดยที่จะมีความเอนเอียงสัมพัทธ์สูงกว่าวิธี ML แลพ RML แต่น้อยกว่าวิธี WLS ยกเว้นในกรณีที่มีสัมประสิทธิ์ความโด่งเท่ากับ -1 และเดลระบุผิดพลาดระดับน้อยและปานกลาง วิธี GLS มีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์สูงกว่าวิธี WLS 3) วิธี WLS มีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์เพิ่มขึ้นเมื่อมีระดับการระบุโมเดลผิดพลาดมากขึ้น โดยมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์สูงกว่าวิธีอื่น ๆ

ตาราง 1: ความเอนเอียงสัมพัทธ์ของการประมาณค่าพารามิเตอร์น้ำหนักองค์ประกอบ จำแนกตามระดับการระบุโมเดลผิดพลาด วิธีประมาณค่า รูปทรงการแจกแจงของข้อมูล และขนาดตัวอย่าง

ระดับการระบุโมเดลผิดพลาด	วิธีประมาณค่า	แบนราบกว่าปกติน้อย (ku=-1)				โด่งกว่าปกติน้อย (ku=1)				โด่งกว่าปกติมาก (ku=2)			
		200	400	600	800	200	400	600	800	200	400	600	800
โมเดลระบุไม่ผิดพลาด (RMSEA=0)	ML	0.325	0.182	0.211	0.094	0.369	0.256	0.238	0.109	0.211	0.370	0.197	0.193
	GLS	2.318	1.153	0.855	0.507	2.286	1.216	0.869	0.550	2.261	1.271	0.819	0.659
	WLS	2.008	1.006	0.534	0.571	5.734	2.899	2.103	1.537	5.341	4.372	3.059	2.379
	RLM	0.325	0.182	0.211	0.094	0.369	0.256	0.238	0.109	0.211	0.370	0.197	0.193
โมเดลระบุผิดพลาดระดับน้อย (RMSEA=0.02)	ML	0.508	0.517	0.241	0.177	0.369	0.192	0.298	0.180	0.517	0.287	0.269	0.285
	GLS	2.476	2.504	0.904	0.566	2.286	1.150	0.870	0.559	2.504	1.011	0.863	0.712
	WLS	2.027	8.080	0.820	0.783	5.734	2.988	2.249	1.658	8.080	4.451	3.370	2.827
	RLM	0.508	0.517	0.241	0.177	0.369	0.192	0.298	0.180	0.517	0.287	0.269	0.285

ตาราง 1: ความเอนเอียงสัมพัทธ์ของการประมาณค่าพารามิเตอร์น้ำหนักองค์ประกอบ จำแนกตามระดับการระบุโมเดลผิดพลาด วิธีประมาณค่า รูปทรงการแจกแจงของข้อมูล และขนาดตัวอย่าง (ต่อ)

ระดับการระบุโมเดลผิดพลาด	วิธีประมาณค่า	แบนราบกว่าปกติน้อย (ku=-1)				โด่งกว่าปกติน้อย (ku=1)				โด่งกว่าปกติมาก (ku=2)			
		200	400	600	800	200	400	600	800	200	400	600	800
โมเดลระบุ	ML	0.364	0.361	0.260	0.241	0.285	0.361	0.260	0.366	0.380	0.299	0.241	0.345
ผิดพลาดระดับ	GLS	2.202	1.065	0.767	0.717	0.712	1.065	0.767	0.720	2.221	1.114	0.717	0.620
ปานกลาง	WLS	3.185	3.218	1.724	3.845	2.827	3.218	1.724	2.121	8.528	5.042	3.845	3.242
(RMSEA=0.04)	RLM	0.364	0.361	0.260	0.241	0.285	0.361	0.260	0.366	0.380	0.299	0.241	0.345
โมเดลระบุ	ML	0.505	0.396	0.409	0.446	0.431	0.465	0.405	0.400	0.431	0.439	0.349	0.398
ผิดพลาดระดับ	GLS	1.991	1.042	0.799	0.782	2.308	1.034	0.836	0.704	2.308	1.248	0.803	0.729
มาก	WLS	4.519	3.281	2.994	2.778	9.146	3.911	3.142	2.656	9.146	6.141	4.875	4.251
(RMSEA=0.06)	RLM	0.505	0.396	0.409	0.446	0.431	0.465	0.405	0.400	0.431	0.439	0.349	0.398

ความแม่นยำของการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝง

จากตาราง 2 ความเอนเอียงสัมพัทธ์จากการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝง พบว่า 1) วิธีประมาณค่า ML และ RML มีความเอนเอียงสัมพัทธ์ไม่แตกต่างกัน และมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี GLS และ WLS ทุกขนาดตัวอย่าง ระดับการระบุโมเดลผิดพลาด โดยมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์สูงขึ้น เมื่อมีระดับการระบุโมเดลผิดพลาดมากขึ้น ยกเว้นในกรณีที่มีสัมประสิทธิ์ความโด่งเท่ากับ 1 และ 2 วิธี WLS จะมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี ML และ RML เมื่อโมเดลระบุผิดพลาดระดับปานกลาง และมาก 2) วิธี GLS มีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์ลดลงเมื่อมีขนาดตัวอย่างเพิ่มขึ้น และมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์สูงขึ้นเมื่อมีระดับการระบุโมเดลผิดพลาดมากขึ้น โดยจะมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี WLS เมื่อมีสัมประสิทธิ์ความโด่งเท่ากับ -1 และมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์สูงกว่าวิธีอื่นๆเมื่อมีสัมประสิทธิ์ความโด่งเท่ากับ 1 และ 2 3) วิธี WLS มีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์มากกว่าวิธีอื่นๆเมื่อมีสัมประสิทธิ์ความโด่งเท่ากับ -1 และจะมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธีอื่นๆเมื่อมีระดับการระบุโมเดลผิดพลาดมากขึ้น และมีสัมประสิทธิ์ความโด่งเท่ากับ 1 และ 2

ตาราง 2: ความเอนเอียงสัมพัทธ์ของการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝง จำแนกตามระดับการระบุโมเดลผิดพลาด วิธีประมาณค่า รูปทรงการแจกแจงของข้อมูล และขนาดตัวอย่าง

ระดับการระบุโมเดลผิดพลาด	วิธีประมาณค่า	แบนราบกว่าปกติน้อย (ku=-1)				โด่งกว่าปกติน้อย (ku=1)				โด่งกว่าปกติมาก (ku=2)			
		200	400	600	800	200	400	600	800	200	400	600	800
โมเดลระบุไม่ผิดพลาด (RMSEA=0)	ML	0.388	0.718	0.336	0.126	0.932	0.312	0.634	0.857	0.611	0.417	0.634	0.186
	GLS	4.058	1.881	1.065	0.968	3.952	2.282	0.989	0.238	3.478	1.849	0.989	0.951
	WLS	10.832	4.766	2.823	2.193	1.561	0.429	0.651	1.333	0.726	2.546	0.651	1.122
	RLM	0.388	0.718	0.336	0.126	0.932	0.312	0.634	0.857	0.611	0.417	0.634	0.186
โมเดลระบุ	ML	1.229	1.355	1.038	1.145	0.932	1.885	1.316	1.772	1.355	1.088	1.236	1.671
ผิดพลาดระดับ	GLS	3.624	4.779	1.539	2.154	3.952	3.852	2.300	2.745	4.779	2.872	1.563	2.727
น้อย	WLS	10.628	3.288	3.799	3.923	1.561	1.136	1.473	1.023	3.288	1.936	1.903	0.980
(RMSEA=0.02)	RLM	1.229	1.355	1.038	1.145	0.932	1.885	1.316	1.772	1.355	1.088	1.236	1.671

ตาราง 2: ความเอนเอียงสัมพัทธ์ของการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝง จำแนกตามระดับการระบุโมเดลผิดพลาด วิธีประมาณค่า รูปทรงการแจกแจงของข้อมูล และขนาดตัวอย่าง (ต่อ)

ระดับการระบุโมเดลผิดพลาด	วิธีประมาณค่า	แบนราบกว่าปกติน้อย (ku=-1)				โด่งกว่าปกติน้อย (ku=1)				โด่งกว่าปกติมาก (ku=2)			
		200	400	600	800	200	400	600	800	200	400	600	800
โมเดลระบุ	ML	3.408	3.100	3.176	3.138	1.671	3.100	3.176	3.427	4.592	3.386	3.138	3.524
ผิดพลาดระดับ	GLS	5.873	4.028	4.098	4.269	2.727	4.028	4.098	4.176	6.583	5.087	4.269	4.184
ปานกลาง	WLS	14.118	2.402	7.546	1.544	0.980	2.402	7.546	2.669	2.920	1.340	1.544	1.833
(RMSEA=0.04)	RLM	3.408	3.100	3.176	3.138	1.671	3.100	3.176	3.427	4.592	3.386	3.138	3.524
โมเดลระบุ	ML	5.278	5.956	5.686	5.679	5.491	6.000	5.374	5.543	5.491	5.040	5.782	5.443
ผิดพลาดระดับ	GLS	8.547	7.961	7.298	7.065	8.589	8.235	7.064	6.907	8.589	7.186	7.344	6.871
มาก	WLS	19.225	14.595	12.594	12.100	0.824	4.161	3.695	3.492	0.824	1.313	1.446	1.245
(RMSEA=0.06)	RLM	5.278	5.956	5.686	5.679	5.491	6.000	5.374	5.543	5.491	5.040	5.782	5.443

ความแม่นยำของการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน

จากตาราง 3 ความเอนเอียงสัมพัทธ์จากการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน พบว่า 1) วิธีประมาณค่า ML และ RML มีความเอนเอียงสัมพัทธ์ไม่แตกต่างกัน และมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี GLS และ WLS ทุกขนาดตัวอย่าง ระดับการระบุโมเดลผิดพลาด และรูปทรงการแจกแจงของข้อมูล โดยมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์ต่ำลง เมื่อมีระดับการระบุโมเดลผิดพลาดมากขึ้น 2) วิธี GLS มีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์สูงขึ้นเมื่อมีระดับการระบุโมเดลผิดพลาดมากขึ้น โดยมีความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี WLS เมื่อมีสัมประสิทธิ์ความโด่งเท่ากับ 2 และมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์ใกล้เคียงกับวิธี WLS เมื่อมีขนาดตัวอย่างมากขึ้น เมื่อมีสัมประสิทธิ์ความโด่งเท่ากับ -1 และ 1 3) วิธี WLS มีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์เพิ่มขึ้นเมื่อมีระดับการระบุโมเดลผิดพลาดมากขึ้น โดยมีแนวโน้มที่จะมีความเอนเอียงสัมพัทธ์สูงกว่าวิธีอื่น ๆ

ตาราง 3: ความเอนเอียงสัมพัทธ์ของการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อนจำแนกตามระดับการระบุโมเดลผิดพลาด วิธีประมาณค่า รูปทรงการแจกแจงของข้อมูล และขนาดตัวอย่าง

ระดับการระบุโมเดลผิดพลาด	วิธีประมาณค่า	แบนราบกว่าปกติน้อย (ku=-1)				โด่งกว่าปกติน้อย (ku=1)				โด่งกว่าปกติมาก (ku=2)			
		200	400	600	800	200	400	600	800	200	400	600	800
โมเดลระบุไม่ผิดพลาด (RMSEA=0)	ML	4.494	3.995	3.821	3.784	7.696	4.218	4.119	3.729	4.676	3.973	3.831	3.875
	GLS	10.798	7.409	6.058	5.410	15.993	7.609	6.356	5.393	11.023	7.447	5.967	5.600
	WLS	10.312	6.964	5.767	5.231	18.331	8.682	7.188	6.059	12.579	9.848	7.674	6.934
	RLM	4.494	3.995	3.821	3.784	7.696	4.218	4.119	3.729	4.676	3.973	3.831	3.875
โมเดลระบุผิดพลาดระดับน้อย (RMSEA=0.02)	ML	4.552	5.123	3.950	4.000	1.103	4.384	4.111	4.118	5.123	4.401	4.269	4.073
	GLS	11.779	12.378	6.894	6.407	9.992	8.512	7.060	6.472	12.378	8.418	7.236	6.432
	WLS	11.399	16.135	6.615	6.223	12.497	9.681	7.929	7.200	16.135	10.912	9.034	8.051
	RLM	4.552	5.123	3.950	4.000	1.103	4.384	4.111	4.118	5.123	4.401	4.269	4.073

ตาราง 3: ความเอนเอียงสัมพัทธ์ของการประมาณค่าพารามิเตอร์ความแปรปรวนร่วมระหว่างความคลาดเคลื่อนจำแนกตามระดับการระบุโมเดลผิดพลาด วิธีประมาณค่า รูปทรงการแจกแจงของข้อมูล และขนาดตัวอย่าง (ต่อ)

ระดับการระบุโมเดลผิดพลาด	วิธีประมาณค่า	แบนราบกว่าปกตินิ้อย (ku=-1)				โด่งกว่าปกตินิ้อย (ku=1)				โด่งกว่าปกติมาก (ku=2)			
		200	400	600	800	200	400	600	800	200	400	600	800
โมเดลระบุ	ML	3.208	2.882	2.896	2.934	2.631	2.882	2.896	2.795	3.365	3.124	2.934	2.868
ผิดพลาดระดับ	GLS	12.609	9.042	7.766	7.884	5.033	9.042	7.766	7.194	12.829	9.325	7.884	7.138
ปานกลาง	WLS	12.231	10.420	7.461	10.252	6.666	10.420	7.461	8.207	17.120	12.197	10.252	9.220
(RMSEA=0.04)	RLM	3.208	2.882	2.896	2.934	2.631	2.882	2.896	2.795	3.365	3.124	2.934	2.868
โมเดลระบุ	ML	1.277	0.781	0.808	0.769	1.269	0.833	0.831	0.688	1.269	0.943	0.801	0.829
ผิดพลาดระดับ	GLS	15.606	11.297	9.668	9.008	15.849	11.334	9.815	9.032	15.849	11.480	10.016	9.142
มาก	WLS	15.276	10.828	9.263	8.658	21.591	13.344	11.532	10.591	21.591	15.628	13.618	12.455
(RMSEA=0.06)	RLM	1.277	0.781	0.808	0.769	1.269	0.833	0.831	0.688	1.269	0.943	0.801	0.829

2. วิเคราะห์ผลกระทบของปัจจัยระดับการระบุโมเดลผิดพลาด วิธีประมาณค่าพารามิเตอร์ รูปทรงการแจกแจงของข้อมูล ขนาดตัวอย่าง และประเภทของพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบ ที่ส่งผลต่อความเอนเอียงสัมพัทธ์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน

การวิเคราะห์ผลกระทบของปัจจัยระดับการระบุโมเดลผิดพลาด วิธีประมาณค่าพารามิเตอร์ รูปทรงการแจกแจงของข้อมูล ขนาดตัวอย่าง และประเภทของพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบ ที่ส่งผลต่อความเอนเอียงสัมพัทธ์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน จะวิเคราะห์ว่าปัจจัยใดที่ส่งผลกระทบต่อความเอนเอียงสัมพัทธ์ และส่งอย่างไร ในระดับไหน ด้วยการวิเคราะห์ความแปรปรวน ผลการวิเคราะห์แสดงในตาราง 4 รายละเอียดเป็นดังนี้

จากผลการวิเคราะห์ความแปรปรวนในตาราง 4 พบว่ามีปัจจัยทั้งสิ้น 4 ปัจจัยที่ส่งผลกระทบอย่างมีนัยสำคัญทางสถิติต่อความเอนเอียงสัมพัทธ์ของค่าประมาณพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน ซึ่งสามารถเรียงตามลำดับจากมากไปน้อยด้วยค่าขนาดอิทธิพล Partial Eta Squared (η^2) เป็นดังนี้ ปัจจัยประเภทของพารามิเตอร์ ส่งผลต่อความเอนเอียงสัมพัทธ์ของค่าประมาณพารามิเตอร์มากที่สุด (η^2=.427, p-value=.000) รองลงมาคือปัจจัย วิธีประมาณค่าพารามิเตอร์ (η^2=.346, p-value=.000) ระดับการระบุโมเดลผิดพลาด(η^2=.081, p-value=.000) และขนาดตัวอย่าง (η^2=.079, p-value=.000) ตามลำดับ ส่วนปัจจัยรูปทรงการแจกแจงด้านความโด่ง ของข้อมูล พบว่ามีผลกระทบต่อความเอนเอียงสัมพัทธ์ในระดับต่ำมาก และไม่มีนัยสำคัญทางสถิติ (η^2=.005, p-value=.237) โดยที่ 1) ปัจจัยประเภทของพารามิเตอร์ การประมาณค่าพารามิเตอร์น้ำหนักองค์ประกอบ พารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝง และความแปรปรวนร่วมระหว่างความคลาดเคลื่อน มีความเอนเอียงในการประมาณค่าแตกต่างกันทุกพารามิเตอร์ อย่างมีนัยสำคัญ 2) ปัจจัยวิธีประมาณค่าพารามิเตอร์ การประมาณค่าพารามิเตอร์ด้วยวิธี ML จะมีความเอนเอียงสัมพัทธ์แตกต่างกับวิธี GLS และ WLS การประมาณค่าพารามิเตอร์ด้วยวิธี GLS จะมีความ

เอนเอียงสัมพัทธ์แตกต่างกับวิธี WLS และ RML การประมาณค่าพารามิเตอร์ด้วยวิธี WLS จะมีความเอนเอียงสัมพัทธ์แตกต่างกับวิธี RML อย่างมีนัยสำคัญ 3) ปัจจัยระดับการระบุโมเดลผิดพลาด การประมาณค่าพารามิเตอร์ในโมเดลที่ระบุไม่ผิดพลาดจะมีความเอนเอียงสัมพัทธ์แตกต่างกับการประมาณค่าพารามิเตอร์ในโมเดลที่ระบุผิดพลาดระดับมาก และการประมาณค่าพารามิเตอร์ในโมเดลที่ระบุผิดพลาดระดับน้อยจะมีความเอนเอียงสัมพัทธ์แตกต่างกับการประมาณค่าพารามิเตอร์ในโมเดลที่มีการระบุผิดพลาดระดับมาก 4) ปัจจัยขนาดตัวอย่าง การประมาณค่าพารามิเตอร์โดยใช้ขนาดตัวอย่างเท่ากับ 200 จะมีความเอนเอียงสัมพัทธ์แตกต่างกับการใช้ขนาดตัวอย่างเท่ากับ 400 600 และ 800 อย่างมีนัยสำคัญ 5) ปัจจัยรูปทรงการแจกแจงของข้อมูล การประมาณค่าพารามิเตอร์จากข้อมูลที่มีลักษณะแบนราบกว่าปกติน้อย โด่งกว่าปกติน้อยและโด่งกว่าปกติมาก มีความเอนเอียงสัมพัทธ์ไม่แตกต่างกันอย่างมีนัยสำคัญ

ตาราง 4: ตารางวิเคราะห์ความแปรปรวนเพื่อเปรียบเทียบค่าเฉลี่ยของความเอนเอียงสัมพัทธ์ จำแนกตามปัจจัยประเภทของพารามิเตอร์ วิธีประมาณค่าพารามิเตอร์ ระดับของการระบุโมเดลผิดพลาด ขนาดตัวอย่าง และรูปทรงการแจกแจงของข้อมูล

ปัจจัย	Partial Eta Squared	ระดับปัจจัย (I)	ระดับปัจจัย (J)	Mean Difference (I-J)	p-value
ประเภทของพารามิเตอร์*	.427	ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน	น้ำหนักองค์ประกอบ	.0514*	.000
F = 209.280			ความแปรปรวนร่วมระหว่างตัวแปรแฝง	.0314*	.000
(p-value = .000)		ความแปรปรวนร่วมระหว่างตัวแปรแฝง	น้ำหนักองค์ประกอบ	.0200*	.000
วิธีประมาณค่าพารามิเตอร์*	.346	ML	RML	.000	1.000
		GLS	ML	.0287*	.000
			RML	.0287*	.000
		WLS	ML	.0406*	.000
F = 99.162			GLS	.0119*	.001
(p-value = .000)			RML	.0406*	.000
ระดับของการระบุโมเดลผิดพลาด*	.081	ระบุผิดพลาดระดับน้อย	ระบุไม่ผิดพลาด	.0051	.387
		ระบุผิดพลาดระดับปานกลาง	ระบุไม่ผิดพลาด	.0080	.060
			ระบุผิดพลาดระดับน้อย	.0029	.806
F = 16.481		ระบุผิดพลาดระดับมาก	ระบุไม่ผิดพลาด	.0198*	.000
			ระบุผิดพลาดระดับน้อย	.0147*	.000
(p-value = .000)			ระบุผิดพลาดระดับปานกลาง	.0118	.001

ตาราง 4: ตารางวิเคราะห์ความแปรปรวนเพื่อเปรียบเทียบค่าเฉลี่ยของความเอนเอียงสัมพัทธ์ จำแนกตามปัจจัยประเภทของพารามิเตอร์ วิธีประมาณค่าพารามิเตอร์ ระดับของการระบุโมเดลผิดพลาด ขนาดตัวอย่าง และรูปทรงการแจกแจงของข้อมูล (ต่อ)

ปัจจัย	Partial Eta Squared	ระดับปัจจัย (I)	ระดับปัจจัย (J)	Mean Difference (I-J)	p-value
ขนาดตัวอย่าง* F = 16.111 (p-value = .000)	.079	400	200	.0104*	.006
			600	.0062	.220
			800	.0081	.055
		600	200	.0165*	.000
			800	.0019	.933
		800	200	.0184*	.000
สัมประสิทธิ์ความโด่ง F = 1.445 (p-value = .237)	.005	แบนราบกว่าปกติน้อย (ku=-1)	โด่งกว่าปกติน้อย (ku=1)	.0041	.266
			โด่งกว่าปกติมาก (ku=2)	.0010	.924
		โด่งกว่าปกติมาก (ku=2)	โด่งกว่าปกติน้อย (ku=1)	.0031	.469

หมายเหตุ : * คือ มีนัยสำคัญทางสถิติที่ระดับ 0.05

อภิปรายผล

จากผลการวิจัยในข้างต้นผู้วิจัยมีประเด็นอภิปรายจำนวน 4 ประเด็น ดังนี้

(1) ระดับความเอนเอียงสัมพันธ์ของค่าประมาณพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝงมีค่าสูงมากที่สุดเมื่อเปรียบเทียบกับความเอนเอียงสัมพัทธ์ของค่าประมาณพารามิเตอร์ประเภทอื่น ๆ สาเหตุเนื่องมาจากตัวแปรแฝงในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันเป็นตัวแปรที่ไม่มีค่าสังเกตโดยตรง แต่ใช้การประมาณค่าเอาจริงข้อมูลของตัวแปรสังเกตได้ที่ผู้วิจัยใช้เป็นตัวบ่งชี้ของตัวแปรแฝงนั้น ๆ จึงย่อมก่อให้เกิดความคลาดเคลื่อนในการประมาณที่สูงกว่าการประมาณค่าพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนร่วมของความคลาดเคลื่อนจากการวัดที่ประมาณค่าผ่านค่าสังเกตของตัวแปรสังเกตได้โดยตรง (Jichuan Wang, Xiaoquin Wang ; 2012)

(2) การประมาณค่าพารามิเตอร์ด้วยวิธี ML กับวิธี RML มีค่าความเอนเอียงสัมพัทธ์ไม่แตกต่างกัน ทั้งนี้เป็นเพราะวิธีการทั้งสองใช้ฟังก์ชันวัตถุประสงค์ (objective function) เดียวกันในการหาค่าประมาณพารามิเตอร์ของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน สิ่งที่ทำให้ RML แตกต่างจาก ML คือ RML มีการปรับสูตรการประมาณค่าคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์ในโมเดลโดยใช้หลักการประมาณค่าแบบแซนวิช (sandwich estimator) เพื่อให้ค่าประมาณของค่าคลาดเคลื่อนมาตรฐานนั้นมีความแกร่งต่อการละเมิดข้อตกลงเบื้องต้นด้านการแจกแจงปกติของข้อมูลตัวแปรสังเกตได้ ซึ่งการปรับสูตรการประมาณค่า

คลาดเคลื่อนมาตรฐานนี้เป็นกระบวนการที่แยกออกมาจากการหาค่าประมาณค่าพารามิเตอร์ ดังนั้น ค่าประมาณของพารามิเตอร์ที่ได้จากวิธี ML และ RML จึงไม่แตกต่างกัน

(3) วิธี ML และ RML เป็นวิธีการประมาณค่าพารามิเตอร์ที่มีค่าความเอนเอียงสัมพัทธ์ต่ำกว่าวิธี GLS และ WLS ในทุกสถานการณ์ที่ศึกษา ทั้งนี้เนื่องจาก 1) วิธี GLS จะมีแนวโน้มให้ค่าประมาณพารามิเตอร์ที่มีความเอียงเมื่อมีการฝ่าฝืนข้อตกลงเบื้องต้น เช่นการแจกแจงของข้อมูลไม่เป็นแบบปกติ (Boomsma & Hoogland, 2001; Lei & Lomax, 2005; Min, 2008; Olsson et al., 2000; Shimizu & Kano, 2008) โดยผลการวิจัยครั้งนี้พบว่าวิธี GLS ยังคงคงมีความเอนเอียงสัมพัทธ์สูงกว่ากว่าวิธี ML และ RML ถึงแม้ว่าข้อมูลจะมีการแจกแจงแตกต่างไปจากการแจกแจงปกติไม่มาก 2) วิธี WLS เป็นวิธีที่ไม่มีข้อจำกัดเกี่ยวกับการแจกแจงของข้อมูล อย่างไรก็ตามงานวิจัยของ Boomsma & Hoogland (2001) พบว่าวิธี WLS ต้องการขนาดตัวอย่างจำนวนมากในการวิเคราะห์ จึงจะได้ค่าประมาณที่มีความแม่นยำ และมีความคงเส้นคงวา โดยในงานวิจัยใช้จำนวนขนาดตัวอย่างเท่ากับ 40 เท่าต่อจำนวนพารามิเตอร์ในโมเดล วิธี WLS มีความแม่นยำดีกว่าวิธี ML โดยในการวิจัยครั้งนี้ใช้จำนวนขนาดตัวอย่างเท่ากับ 20 เท่าต่อจำนวนพารามิเตอร์ในโมเดล พบว่าวิธี WLS มีความแม่นยำที่ต่ำกว่าวิธี ML และ RML แต่ความเอนเอียงสัมพัทธ์มีแนวโน้มลดลงเมื่อเพิ่มขนาดตัวอย่างมากขึ้น

(4) ผลจากการเปรียบเทียบความเอนเอียงสัมพัทธ์ของการประมาณค่าพารามิเตอร์ตามระดับของรูปทรงการแจกแจงของข้อมูล พบว่าปัจจัยรูปทรงการแจกแจงของข้อมูลไม่มีนัยสำคัญต่อความเอนเอียงการประมาณค่าพารามิเตอร์ ซึ่งไม่สอดคล้องกับงานวิจัยของ Benson & Fleishman (1994); Curran et al. (1996); Lei & Lomax (2005); Min (2008); Olsson et al. (2000) ที่มีผลการวิจัยว่ารูปแบบการแจกแจงของข้อมูลส่งผลต่อประสิทธิภาพของการประมาณค่าพารามิเตอร์ เนื่องจากในการวิจัยครั้งนี้ศึกษาเฉพาะสัมประสิทธิ์ความโด่งเพียง 3 ระดับ คือ ku=-1, ku=1 และ ku=2 ซึ่งมีความแตกต่างกันน้อยเกินไปจนไม่สามารถวิเคราะห์ความแตกต่างของผลการวิเคราะห์ได้ โดยในงานวิจัยของ Curran et al. (1996) พบว่าข้อมูลที่มีการแจกแจงไม่ปกติโดยมี สัมประสิทธิ์ความเบ้เท่ากับ 2 และ สัมประสิทธิ์ความโด่งเท่ากับ 7 จะส่งผลต่อประสิทธิภาพของการประมาณค่าพารามิเตอร์

ข้อเสนอแนะในการนำผลวิจัยไปใช้

ผู้วิจัยควรให้ความสำคัญกับการทบทวนวรรณกรรมเพื่อกำหนดพารามิเตอร์ในโมเดลให้มีความถูกต้องมากที่สุดโดยเฉพาะการกำหนดพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝง เนื่องจากผลการวิจัยชี้ให้เห็นว่าการระบุพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบส่งผลต่อความเอนเอียงสัมพัทธ์มากที่สุดโดยเฉพาะพารามิเตอร์ความแปรปรวนร่วมระหว่างตัวแปรแฝงที่มีแนวโน้มเกิดความเอนเอียงในการประมาณค่ามากกว่าพารามิเตอร์อื่น ๆ ในโมเดล รองลงมาคือการเลือกใช้วิธีประมาณค่าพารามิเตอร์ ที่ผลการวิจัยชี้ให้เห็นว่าวิธี ML และ RML มีความเอนเอียงในการประมาณค่าพารามิเตอร์ต่ำที่สุดทุก ๆ ประเภทของ

พารามิเตอร์ที่ประมาณค่า ระดับการระบุโมเดลผิดพลาด ขนาดตัวอย่าง และรูปทรงการแจกแจงของข้อมูล ที่ใช้ในการวิจัยครั้งนี้ ส่วนรูปทรงการแจกแจงของข้อมูล ผลการวิจัยชี้ให้เห็นว่าข้อมูลที่มีสัมประสิทธิ์ความโด่งไม่สูงนักไม่ส่งผลต่อความเอนเอียงสัมพัทธ์

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ประยุกต์วิธีดำเนินการวิจัยไปใช้กับโมเดลการวิเคราะห์อื่น ๆ เช่น การวิเคราะห์โมเดลสมการโครงสร้าง (Structural Equation Modeling: SEM) เพื่อศึกษาอิทธิพลของพารามิเตอร์อื่น ๆ ในการวิเคราะห์เพิ่มเติม

2. ขยายกรอบแนวคิดการวิจัยให้ตัวแปรรูปทรงการแจกแจงของข้อมูล และขนาดตัวอย่างให้มีความหลากหลายมากขึ้น เพื่อที่จะให้ได้ผลการวิจัยที่มีความชัดเจนและครอบคลุม เนื่องจากผลการวิจัยครั้งนี้ใช้ขนาดตัวอย่างไม่เพียงพอต่อการศึกษาความเอนเอียงของวิธี WLS ได้อย่างชัดเจน และยังไม่สามารถวิเคราะห์อิทธิพลของรูปทรงการแจกแจงของข้อมูลต่อความเอนเอียงของการประมาณค่าพารามิเตอร์ได้อย่างมีนัยสำคัญ

3. ขยายการถึงศึกษาประสิทธิภาพของดัชนีวัดความสอดคล้องเชิงประจักษ์ ที่ได้จากการประมาณค่าพารามิเตอร์ รวมถึงผลกระทบของวิธีประมาณค่าพารามิเตอร์ที่มีต่อดัชนีวัดความสอดคล้องเชิงประจักษ์เมื่อมีปัจจัยที่ทำให้เกิดความเอนเอียงสัมพัทธ์

รายการอ้างอิง

ภาษาไทย

นงลักษณ์ วิรัชชัย. (2542). *โมเดลลิสเรล : สถิติวิเคราะห์สำหรับการวิจัย.* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สิวะโชติ ศรีสุทธิยากร. (2555). *วิธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด: การศึกษาสถานการณ์จำลองแบบมอนติคาร์โลและข้อมูลจริง* (วิทยานิพนธ์ปริญญามหาบัณฑิต). สาขาวิธีวิทยาการวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร.

ภาษาอังกฤษ

Benson, J., & Fleishman, J. A. (1994). The robustness of maximum likelihood and distribution-free estimators to non-normality in confirmatory factor analysis. Quality and Quantity, 28(2), 117-136.

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog, 139-168.

Curran, P. J. (1994). The robustness of confirmatory factor analysis to model misspecification and violations of normality. ProQuest Information & Learning.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychological methods, 1(1), 16.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. Multivariate Behavioral Research, 42(3), 509-529.

Fan, X., & Wang, L. (1998). Effects of potential confounding factors on fit indices and parameter estimates for true and misspecified SEM models. Educational and Psychological Measurement, 58(5), 701-735.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. Sociological Methods & Research, 26(3), 329-367.

Jöreskog, K. G., & Sörbom, D. (1996). LISREL 8: User's reference guide: Scientific Software International.

Lei, M., & Lomax, R. G. (2005). The effect of varying degrees of nonnormality in structural equation modeling. Structural equation modeling, 12(1), 1-27.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. Sankhyā: The Indian Journal of Statistics, Series B, 36(2), 115-128.

Marsh, H. W., & Hau, K.-T. (1999). Confirmatory factor analysis: Strategies for small sample sizes. Statistical strategies for small sample research, 1, 251-284.

Min, Y. (2008). Robustness in confirmatory factor analysis: The effect of sample size, degree of non-normality, model, and estimation method on accuracy of estimation for standard errors. Citeseer.

Muthén, L. K., & Muthén, B. O. (2010). Mplus User's Guide: Statistical Analysis with Latent Variables: User'ss Guide: Muthén & Muthén.

Olsson, U. H., Foss, T., Troye, S. V., & Howell, R. D. (2000). The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. Structural equation modeling, 7(4), 557-595.

Pornprasertmanit, S. (2014). The Unified Approach for Model Evaluation in Structural Equation Modeling.

Shimizu, S., & Kano, Y. (2008). Use of non-normality in structural equation modeling: Application to direction of causation. Journal of Statistical Planning and Inference, 138(11), 3483-3491.

Wang, J., & Wang, X. (2012). Structural equation modeling: Applications using Mplus: John Wiley & Sons.