วิธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด: การศึกษาสถานการณ์จำลองแบบมอนติคาร์โลและข้อมูลจริง

นายสิวะโชติ ศรีสุทธิยากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต
สาขาวิชาวิธีวิทยาการวิจัยการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2555

BAYESIAN ESTIMATION FOR MULTI-LEVEL DATA ANALYSIS WITH MEASUREMENT ERROR IN VARIABLES: MONTE CARLO SIMULATION AND EMPIRICAL DATA STUDIES

Mr. Siwachoat Srisuttiyakorn

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Educational Research Methodology
Department of Educational Research and Psychology
Faculty of Education
Chulalongkorn University
Academic Year 2012

หัวข้อวิทยานิพนธ์	วิธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด: การศึกษาสถานการณ์จำลองแบบมอนติคาร์โลและข้อมูลจริง.
โดย	นายสิวะโชติ ศรีสุทธิยากร
สาขาวิชา	วิธีวิทยาการวิจัยการศึกษา
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ศาสตราจารย์ ดร. เออิจิ มูรากิ

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

………………………………………….. คณบดีคณะครุศาสตร์
(ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

คณะกรรมการสอบวิทยานิพนธ์

……………………………………………ประธานกรรมการ
(รองศาสตราจารย์ ดร. อวยพร เรืองตระกูล)

………………………………………….. อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี)

…………………………………………… อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ศาสตราจารย์ ดร. เออิจิ มูรากิ)

…………………………………………… กรรมการ
(รองศาสตราจารย์ ดร. สุชาดา บวรกิติวงศ์)

…………………………………………… กรรมการ
(รองศาสตราจารย์ ดร. วรรณี แกมเกตุ)

…………………………………………… กรรมการภายนอกมหาวิทยาลัย
(ศาสตราจารย์กิตติคุณ ดร. นงลักษณ์ วิรัชชัย)

สิวะโชติ ศรีสุทธิยากร: วิธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด: การศึกษาสถานการณ์จำลองแบบมอนติคาร์โลและข้อมูลจริง.(BAYESIAN ESTIMATION FOR MULTI-LEVEL DATA ANALYSIS WITH MEASUREMENT ERROR IN VARIABLES: MONTE CARLO SIMULATION AND EMPIRICAL DATA STUDIES) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ศ. ดร. ศิริชัย กาญจนวาสี, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม :ศ. ดร. เออิจิ มูรากิ, 262 หน้า.

วัตถุประสงค์ของการวิจัยนี้มี 3 ประการ (1) เพื่อพัฒนาวิธีการประมาณในโมเดลสมการโครงสร้างพหุระดับที่พารามิเตอร์ในโมเดลเป็นแบบสุ่มและโมเดลการวัดมีน้ำหนักองค์ประกอบและความคลาดเคลื่อนจากการวัดเป็นแบบสุ่ม และ (2) เพื่อตรวจสอบและเปรียบเทียบความสามารถของวิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้นกับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus ภายใต้สถานการณ์จำลองแบบมอนติคาร์โลทั้งนี้ในการศึกษาวิธีการประมาณค่าแบบเบส์จะประมาณค่าภายใต้โมเดลการวัดที่มีน้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่ม ส่วนวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จะประมาณค่าภายใต้โมเดลการวัดที่มีน้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดคงที่ โดยผู้วิจัยกำหนดสถานการณ์จำลองให้มีค่าเฉลี่ยความเที่ยงรวมเท่ากับ 0.3, 0.5, 0.7 และ 0.9 จำนวนตัวอย่างในระดับที่สองเท่ากับ 15, 30 และ 50 หน่วย และจำนวนตัวอย่างในระดับที่หนึ่งเท่ากับ 30 หน่วย เกณฑ์การพิจารณาเปรียบเทียบความสามารถระหว่างวิธีการประมาณทั้งสองได้แก่ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (mean square error: MSE) และ (3) เพื่อทดลองใช้วิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่ผู้วิจัยพัฒนาขึ้นวิเคราะห์ข้อมูลจริงและเปรียบเทียบกับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus

ผลการวิจัยสามารถสรุปได้ดังนี้

(1) อัลกอริทึมที่พัฒนาขึ้นใช้การสุ่มตัวอย่างแบบกิบส์ (Gibb sampling algorithm) ซึ่งประกอบไปด้วยขั้นตอนการประมาณทั้งหมด 10 ขั้นตอน ซึ่งเป็นการสุ่มตัวอย่างพารามิเตอร์จากการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ในโมเดลได้แก่ พารามิเตอร์จุดตัดแกน พารามิเตอร์น้ำหนักองค์ประกอบพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด พารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝง คะแนนองค์ประกอบของตัวแปรแฝง พารามิเตอร์อิทธิพลคงที่ พารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1 และพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 วิธีการประมาณค่าดังกล่าวเป็นวิธีการประมาณค่าแบบทวนซ้ำ ซึ่งเมื่อทำการสุ่มตัวอย่างจากการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขในข้างต้นด้วยจำนวนรอบที่มากเพียงพอจะสามารถใช้ตัวอย่างสุ่มที่ได้เพื่อประมาณการแจกแจงความน่าจะเป็นภายหลังร่วมของพารามิเตอร์ในโมเดลที่ต้องการได้

(2) การศึกษาด้วยข้อมูลจำลอง ผลการวิจัยให้ข้อค้นพบว่า วิธีประมาณค่าแบบเบส์จะให้ค่า MSE ต่ำกว่าวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted ที่ได้จากโปรแกรม Mplus ในทุกกรณีที่ทำการศึกษา ปัจจัยค่าเฉลี่ยความเที่ยงรวมหรือจำนวนหน่วยตัวอย่างในระดับที่ สองที่เพิ่มขึ้นจะทำให้ค่า MSE ของวิธีการทั้งสองมีค่าลดลง และเมื่อค่าเฉลี่ยความเที่ยงรวมเพิ่มขึ้นหรือจำนวนหน่วยตัวอย่างเพิ่มขึ้นจะทำให้ค่า MSE ของวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted มีค่าต่ำลงลู่เข้าใกล้ค่า MSE ที่ได้จากวิธีประมาณค่าแบบเบส์ นอกจากนี้ยังพบว่าวิธีประมาณค่าแบบเบส์เป็นวิธีที่สามารถประมาณค่าได้ดีถึงแม้ว่าขนาดตัวอย่างจะมีขนาดเล็กก็ตาม

(3) การศึกษาด้วยข้อมูลเชิงประจักษ์ผู้วิจัยใช้ข้อมูลทุติยภูมิจากวิทยานิพนธ์เรื่อง อิทธิพลของพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลและสุขภาวะของครูที่มีต่อสุขภาวะของนักเรียน : โมเดลการปรับและการส่งผ่านพหุระดับ (ณมรัตน์ ศิริภาพ, 2554) พบว่าวิธีการประมาณค่าพารามิเตอร์ทั้งสองวิธีให้ค่าประมาณค่าพารามิเตอร์ที่ใกล้เคียงกันและให้ผลสรุปไปในทิศทางเดียวกันส่วนใหญ่ ยกเว้นค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบ พารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด อิทธิพลคงที่ บางตัวที่มีขนาดของค่าประมาณและข้อสรุปที่แตกต่างกัน ซึ่งพบว่าผลการประมาณค่าพารามิเตอร์ที่ได้จากวิธีการประมาณค่าแบบเบส์มีความสมเหตุสมผลมากกว่าการใช้วิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุด

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต..

สาขาวิชา วิธีวิทยาการวิจัยการศึกษา ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก....................................

ปีการศึกษา 2555 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....................................

5184260127 : MAJOR EDUCATIONAL RESEARCH METHODOLOGY
KEYWORDS : MULTILEVEL STRUCTURAL EQUATION MODEL / RANDOM FACTOR LOADINGS MODEL / BAYESIAN ESTIMATION / MARKOV CHAIN MONTE CARLO / GIBBS SAMPLING ALGORITHM

SIWACHOAT SRISUTTIYAKORN: BAYESIAN ESTIMATION FOR MULTI-LEVEL DATA ANALYSIS WITH MEASUREMENT ERROR IN VARIABLES: MONTE CARLO SIMULATION AND EMPIRICAL DATA STUDIES. ADVISOR : PROF. SIRICHAI KANJANAWASEE, Ph.D., CO-ADVISOR : PROF. EIJI MURAKI, Ph.D., 262 pp.

The three purposes of this research were (1) to develop the Bayesian estimation method to estimate the parameters in multi-level structural equation model with random factor loadings and random measurement error variances, (2) to verify and compare the accuracy between developed Bayesian estimation method and maximum likelihood estimation method via Mplus program under various simulation conditions. Simulation design was used with the level of average composite reliability are equal to 0.3, 0.5, 0.7 and 0.9, level-1 sample size is 30 units, and level-1 sample size are 15, 30, and 50 units. Mean square error (MSE) was used to compare the accuracy and efficiency between the estimators, and (3) to test the developed Bayesian estimation method on empirical data and compare the results with the results from maximum likelihood estimation method.

Summarized results of the research were:

(1) Developed algorithm is 10 steps Gibbs sampling algorithm. This algorithm is a simulation based approach that involves sequentially samplings from conditional probability distribution associated with each block of parameters such as measurement intercept, factor loadings, measurement error variances, variance-covariance of independent latent variables, factor scores, fixed effects, leve-1 model variance , and level-2 – model variances. The algorithm will guarantee that if we run for large number of iteration, a sequence of sample draw from the conditional probability distributions will converges in distribution to the joint posterior distribution.

(2) In simulation study, When comparing MSE of the Bayesian method and the maximum likelihood method, the MSE from Bayesian estimation method will always have value lower than the MSE from maximum likelihood method. In considering the trend of MSE value as the average of composite reliability value or number of clusters are increases, the MSE value from both methods are likely to have lower value. Moreover the result shown that Bayesian estimator is robust for small sample size cases.

(3) In empirical study, researcher used secondary data from the study “EFFECTS OF INTERPERSONAL TEACHER BEHAVIOR AND TEACHERS’ WELL-BEING ON STUDENTS’ WELL-BEING: A MULTI-LEVEL STRUCTURAL EQUATION MODELING WITH MODERATION AND MEDIATION” (Thomrat Siriparp, 2011). The estimated values of the two estimation methods are mostly coherent except the estimated value of factor loadings, measurement error variances, and fixed effect parameter. Researcher found that Bayesian estimation methods are more reasonable than the maximum likelihood method.

Department : Educational Research and Psychology	Student’s Signature ______________
Field of Study : Educational Research Methodology	Advisor’s Signature ______________
Academic Year : 2012	Co-advisor’s Signature ______________

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้จะสำเร็จลุล่วงไม่ได้หากขาดความช่วยเหลือ ความเมตตาและเอาใจใส่อย่างดียิ่งจาก ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี อาจารย์ที่ปรึกษาหลัก และศาสตราจารย์ ดร. เออิจิ มูรากิ อาจารย์ที่ปรึกษาร่วม ขอกราบขอบพระคุณอาจารย์ที่ได้เสียสละเวลาอันมีค่าให้คำชี้แนะ ปรึกษา ข้อคิดเห็น และแก้ไขข้อบกพร่องต่างๆ ซึ่งเป็นประโยชน์ทั้งในการทำวิทยานิพนธ์ และการทำงานในภายภาคหน้า

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. อวยพร เรืองตระกูล ประธานกรรมการสอบวิทยานิพนธ์ ศาสตราจารย์กิตติคุณ นงลักษณ์ วิรัชชัย รองศาสตราจารย์ ดร. สุชาดา บวรกิติวงษ์ และรองศาสตราจารย์ วรรณี แกมเกตุ กรรมการสอบวิทยานิพนธ์ ที่ได้ให้ความกรุณาตรวจสอบและให้ข้อเสนอแนะอันเป็นประโยชน์อย่างมากต่อการปรับปรุงแก้ไขวิทยานิพนธ์ให้มีความสมบูรณ์ ขอกราบขอบพระคุณ ศาสตราจารย์กิตติคุณ นงลักษณ์ วิรัชชัย ที่ได้ให้คำแนะนำและความรู้พื้นฐานที่สำคัญสำหรับการทำวิทยานิพนธ์ ให้คำปรึกษาแก่ผู้วิจัยตั้งแต่เริ่มพัฒนาหัวข้อ ชี้แนะแนวทางการแก้ปัญหาจวบจนวิทยานิพนธ์เสร็จสิ้นสมบูรณ์ ขอกราบขอบพระคุณ ศาสตราจารย์ ดร. สุวิมล ว่องวาณิช ที่ให้คำแนะนำในการเลือกหัวข้อ รวมทั้งดูแลผู้วิจัยเป็นอย่างดีตลอดระยะเวลาการศึกษา และขอกราบขอบพระคุณคณาจารย์ภาควิชาวิจัยและจิตวิทยาการศึกษาทุกท่านที่ได้ให้ทั้งความรู้อันมีค่า สั่งสอนและให้ความช่วยเหลือเอื้ออาทรเสมอมา

ขอขอบคุณ อ. ดร. ฌมรัตน์ ศิริภาพ ที่ได้ให้ความอนุเคราะห์ข้อมูลตัวอย่างสุขภาวะนักเรียนที่ใช้ในการศึกษา ให้ความช่วยเหลือ คำปรึกษา และดูแลผู้วิจัยเป็นอย่างดีตลอดระยะเวลาที่ศึกษา ขอขอบคุณ พันโทหญิง อภิญญา อินทรรัตน์ อ. ดร. เด่นดาว ชลวิทย์ และ อ. เสาวรส ยิ่งวรรณะ ตลอดจนเพื่อนพี่น้องนิสิตภาควิชาวิจัยและจิตวิทยาการศึกษา ที่ได้ร่วมแลกเปลี่ยนประสบการณ์การเรียนรู้ในวิชาสัมมนา รวมทั้งให้การช่วยเหลือและเป็นกำลังใจที่ดีตลอดมา

ขอกราบขอบพระคุณ คุณป้า รองศาสตราจารย์ พญ. ศรีริน สินธุภัค คุณลุง รองศาสตราจารย์ อัครเดช สินธุภัค คุณป้า ดร. รัตนา สินธุภัค และคุณน้า รองศาสตราจารย์ พญ. วัณณศรี สินธุภัค ที่ได้ให้การดูแล ให้คำแนะนำ ให้ความรัก และให้การสนับสนุนโอกาสทางการศึกษาอันที่ค่าอย่างยิ่ง ขอขอบคุณ คุณพี่ ผ.ศ. ดร. อภิสักส์ สินธุภัค ที่เป็นกำลังใจ และให้คำปรึกษาในด้านภาษาอังกฤษ และให้คำแนะนำอื่นๆที่เป็นประโยชน์อย่างมากตลอดระยะเวลาที่ผู้วิจัยศึกษา

ท้ายสุดนี้ขอกราบขอบพระคุณ คุณย่า เอ็ง ศรีสุทธิยากร ที่คอยเป็นกำลังใจและให้คำอวยพรจำนวนมากแก่ผู้วิจัยเสมอมา ขอกราบขอบพระคุณ คุณพ่อ สมศักดิ์ ศรีสุทธิยากร และคุณแม่ ศุภศรี ศรีสุทธิยากร และคุณพี่ ร้อยเอก พญ. ชุติกา ศรีสุทธิยากร ที่ได้เลี้ยงดู ให้ความรักเอาใจใส่และดูแลผู้วิจัยเป็นอย่างดี เป็นแรงผลักดันสำคัญ และให้โอกาสทางการศึกษาแก่ผู้วิจัย

สารบัญ

สารบัญตาราง

สารบัญรูป

รูปที่ หน้า

รูปที่		หน้า

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในการวิจัยทางสังคมศาสตร์หรือทางการศึกษานั้น ผู้วิจัยมักพบว่าข้อมูลมีธรรมชาติที่หน่วยตัวอย่างในระดับที่ต่ำกว่า (micro-level unit) ซ้อน (nested) อยู่ในหน่วยตัวอย่างในระดับที่สูงกว่า (macro-level unit) เป็นลำดับไป ข้อมูลที่มีลักษณะดังกล่าวเรียกว่า ข้อมูลพหุระดับ (multi-level data) หรือ ข้อมูลระดับลดหลั่น (hierarchical data) ตัวอย่างของข้อมูลดังกล่าว เช่น ในงานวิจัยที่ทำการศึกษาระยะยาว (longitudinal study) จะพบว่าค่าสังเกตที่เก็บได้ในแต่ละช่วงเวลานั้นซ้อนอยู่ในหน่วยตัวอย่างหน่วยเดียวกัน ในการวิจัยที่ศึกษาทั่วประเทศจะพบว่าหน่วยตัวอย่างที่ใช้ในการศึกษาซ้อนอยู่ในเขตพื้นที่การปกครองในงานวิจัยที่ข้อมูลได้มาจากการสัมภาษณ์โดยใช้ผู้สัมภาษณ์หลายคน จะพบว่าหน่วยตัวอย่างแต่ละหน่วยซ้อนอยู่ในผู้สัมภาษณ์ ในงานวิจัยทางการศึกษาที่เก็บข้อมูลเกี่ยวกับนักเรียนจะพบว่านักเรียนที่เป็นหน่วยตัวอย่างจะซ้อนอยู่ในแบบสอบ หรือห้องเรียน หรือโรงเรียน หรือเขตพื้นที่การศึกษา เป็นต้น จากที่ได้กล่าวมาจะเห็นว่าธรรมชาติดังกล่าวก่อให้เกิดประเด็นปัญหาทั้งในด้านการวิจัยและด้านการวิเคราะห์ข้อมูล กล่าวคือ โครงสร้างของข้อมูลแบบพหุระดับทำให้การวัดตัวแปรสามารถวัดได้จากหลายระดับและการวิเคราะห์ข้อมูลเพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรนั้นก็สามารถกระทำได้หลายระดับตามลักษณะของข้อมูลเช่นเดียวกัน (ศิริชัย กาญจนวาสี, 2550; นงลักษณ์ วิรัชชัย และ สมหวัง พิธิยานุวัฒน์, 2543)

จากธรรมชาติของข้อมูลในข้างต้นเมื่อพิจารณาประเด็นปัญหาในด้านการวิเคราะห์ข้อมูลพบว่า แนวทางการวิเคราะห์ข้อมูลแบบพหุระดับนั้นอาจสามารถกระทำได้ 2 แนวทาง แนวทางแรกคือการวิเคราะห์แบบระดับเดียว (single-level analysis) เป็นแนวทางที่ดำเนินการวิเคราะห์ข้อมูลโดยไม่ได้มีการคำนึงถึงโครงสร้างของข้อมูลในขณะที่ข้อมูลมีความเป็นพหุระดับ ดังนั้นผู้วิจัยจึงมีความจำเป็นที่จะต้องจัดกระทำข้อมูลเพื่อให้ตัวแปรต่างๆอยู่ในระดับเดียวกันแล้วจึงทำการวิเคราะห์ ซึ่งสามารถกระทำได้ในสองลักษณะคือ การกระจายข้อมูลของตัวแปร

(disaggregation) และการรวมข้อมูลของตัวแปร (aggregation) แนวทางการวิเคราะห์ดังกล่าวอาจทำให้เกิดปัญหาในการวิเคราะห์ข้อมูลหลายประการได้แก่ ปัญหาความไม่เป็นอิสระกันของความคลาดเคลื่อนสุ่มอันเนื่องมาจากโครงสร้างของข้อมูลแบบพหุระดับซึ่งจะส่งผลให้การประมาณค่าคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์ในโมเดลมีความผิดพลาด (invalid estimated standard errors) ปัญหาความลำเอียงในการสรุปผลข้ามระดับ (aggregation biased) ปัญหาการปิดบังความผันแปรของสัมประสิทธิ์การถดถอยระหว่างกลุ่ม (heterogeneity of regression coefficients) นอกจากนี้ยังไม่สามารถศึกษาปฏิสัมพันธ์ของตัวแปรข้ามระดับได้ (cross-level interaction) (ศิริชัย กาญจนวาสี, 2550, สิวะโชติ ศรีสุทธิยากร, 2550; Terhorst, 2007; Mass & Hox, 2002) แนวทางที่สองคือการวิเคราะห์พหุระดับ (multi-level analysis) เป็นวิธีการวิเคราะห์โดยการใช้โมเดลทางสถิติที่มีการคำนึงถึงโครงสร้างความเป็นพหุระดับของข้อมูล เรียกโมเดลดังกล่าวว่า โมเดลพหุระดับ (multi-level model) ซึ่งทำให้ผู้วิจัยสามารถวิเคราะห์ความสัมพันธ์ความสัมพันธ์ของตัวแปรในแต่ละระดับและปฏิสัมพันธ์ของตัวแปรระหว่างระดับได้ ผลการวิเคราะห์ที่ได้จะมีความครอบคลุม ลึกซึ้งและน่าเชื่อถือมากกว่าการใช้การวิเคราะห์แบบระดับเดียว (ศิริชัย กาญจนวาสี, 2550; นงลักษณ์ วิรัชชัย และ สมหวัง พิธิยานุวัฒน์, 2543; Goldstein, 1995; Raudenbush and Bryk, 2002; Snijder and Bosker, 1999)

นอกจากนี้ธรรมชาติของศาสตรทางสังคมศาตร์และทางการศึกษาพบว่าสามารถแบ่งประเภทของตัวแปรออกได้เป็นสองประเภทได้แก่ ตัวแปรสังเกตได้ (manifest or observed variables) และตัวแปรแฝง (latent variables) ตัวแปรสังเกตได้เป็นตัวแปรที่ผู้วิจัยสามารถสังเกตค่าได้โดยตรง เช่น คะแนนผลสัมฤทธิ์ทางการเรียน รายได้ อายุ เป็นต้น ส่วนตัวแปรแฝงเป็นตัวแปรที่ผู้วิจัยไม่สามารถสังเกตค่าได้โดยตรงหรือไม่สามารถสังเกตค่าได้โดยอาศัยตัวแปรสังเกตได้เพียงตัวเดียว เช่น ทัศนคติ สุขภาวะ ความพึงพอใจ เป็นต้น ในทางปฏิบัติผู้วิจัยสามารถวัดค่าตัวแปรแฝงโดยผ่านตัวแปรสังเกตจำนวนหนึ่ง ซึ่งแต่ละตัวจะเป็นตัวแทนของมิติต่างๆของตัวแปรแฝงที่ต้องการ กระบวนการดังกล่าวจึงมักทำให้เกิดความคลาดเคลื่อนจากการวัดเสมอ ในทางสถิติผู้วิจัยสามารถวัดความน่าเชื่อถือของการวัดได้โดยใช้ตัวสถิติที่เรียกว่า ค่าความเที่ยง (reliability) ซึ่งใช้วัดระดับของความน่าเชื่อถือของการวัด (ศิริชัย กาญจนวาสี, 2550) โดยปกติแล้วความ

คลาดเคลื่อนจากการวัดสามารถเกิดได้ทั้งในตัวแปรตามและตัวแปรอิสระ ผลกระทบของความคลาดเคลื่อนดังกล่าวจะส่งผลต่อผลการวิเคราะห์ซึ่งอาจทำให้ตัวประมาณค่าแบบจุดไม่มีคุณสมบัติความแม่นยำหรือไม่มีประสิทธิภาพ การประมาณค่าแบบช่วงหรือการทดสอบสมมติฐานทางสถิติที่ใช้มีความไม่น่าเชื่อถือ (ไพฑูรย์ ไกรพรศักดิ์, 2548; Raudenbush and Bryk, 2002; Browne, Goldstein, Woodhouse and Yang, 2002; Fox, 2001) จากที่กล่าวมาจึงเป็นเรื่องสำคัญที่ผู้วิจัยจำเป็นต้องคำนึงถึงโมเดลที่จะใช้วัดค่าของตัวแปรแฝงที่ต้องการและวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรแฝงที่สนใจ วิธีการหนึ่งที่เป็นที่นิยมในปัจจุบันคือการใช้โมเดลสมการโครงสร้าง (structural equation model: SEM) โดยทั่วไปโมเดลสมการโครงสร้างจะประกอบไปด้วยโมเดลย่อยสองโมเดลได้แก่ โมเดลการวัด (measurement model) ซึ่งเป็นโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor alaysis model) ใช้ในการอธิบายความสัมพันธ์เชิงสถิติระหว่างตัวแปรสังเกตได้และตัวแปรแฝง และโมเดลสมการโครงสร้าง (structural model) ซึ่งใช้ในการอธิบายความสัมพันธ์เชิงสถิติระหว่างตัวแปรตามแฝง (dependent or endogeneous variables) กับตัวแปรอิสระแฝง (independent or exogeneous variables) (Jöreskog and Sörbom, 1996)

อย่างไรก็ตามโมเดล SEM ยังมีข้อจำกัดในการวิเคราะห์ข้อมูลเพราะไม่สามารถวิเคราะห์ข้อมูลที่มีโครงสร้างแบบพหุระดับได้ ที่ผ่านมาได้มีนักวิจัยหลายท่านพัฒนาโมเดลและวิธีการประมาณเพื่อรองรับปัญหาดังกล่าว โดยพบว่าอาจแบ่งได้เป็น 2 แนวทางตามลักษณะของโมเดลการวัดได้แก่ แนวทางการแก้ปัญหาที่ใช้โมเดลการวัดแบบดั้งเดิม (classical measurement model) (สิวะโชติ ศรีสุทธิยากร และ ศิริชัย กาญจนวาสี, 2551; Muthen and Asparouhov, 2010; Muthen and Muthen, 2010; Goldstein, Kounail, and Robinson, 2008; Browne, Goldstein, Woodhouse, and Yang, 2001; Woodhouse, Yang, Goldstein, and Rasbash, 1996; Woodhouse, 1996) และแนวทางการแก้ปัญหาโดยใช้โมเดลการวัดแบบโมเดลตอบสนองข้อสอบ (item response model) (Fox. J.-P., 2005; Fox. J. –P., and Glas, 2003; Fox. J. –P., and Glas, 2001) เมื่อพิจารณาแนวทางการแก้ไขปัญหาโดยใช้โมเดลการวัดแบบดั้งเดิมพบว่ามีผู้ที่พัฒนาวิธีการแก้ไขปัญหาไว้หลายวิธีการ เช่น วิธีการประมาณค่าแบบโมเมนต์ (Woodhouse &

M. Yang, 1996) และวิธี adjusted IGLS (Srisuttiyakorn and Kanjanawasee, 2009) เป็นต้น วิธีการในข้างต้นสามารถแก้ไขปัญหาได้ในระดับหนึ่งแต่มีข้อจำกัดที่ไม่สามารถขยายวิธีการดังกล่าวให้ใช้ในกรณีทั่วไปได้ วิธีการที่สามารถแก้ไขปัญหาดังกล่าวได้คือ วิธี adjusted MCMC (Goldstein, Kounail, and Robinson, 2008; Browne, Goldstein, Woodhouse, and Yang, 2001; Browne and Draper, 2000; Browne, 1998) วิธี adjusted MCMC แก้ไขปัญหาความคลาดเคลื่อนจากการวัดโดยการเพิ่มโมเดลการวัดแบบดั้งเดิมเป็นโมเดลย่อยอีกหนึ่งโมเดล เพื่อลดทอนความคลาดเคลื่อนจากการวัด และทำให้ประมาณค่าตัวแปรแฝงได้ ทำให้ผลการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับมีความถูกต้องมากยิ่งขึ้น อย่างไรก็ตามตามวิธี adjusted MCMC ยังมีข้อด้อยกล่าวคือเป็นวิธีการที่มีข้อสมมติให้น้ำหนักความสำคัญของตัวแปรสังเกตได้ทุกตัวที่ใช้วัดตัวแปรแฝงนั้นมีค่าเท่าเทียมกันซึ่งเป็นไปได้ยากในทางปฏิบัติ วิธีการที่มีประสิทธิภาพมากกว่าที่เป็นที่นิยมใช้ในปัจจุบันคือการใช้โมเดลสมการโครงสร้างพหุระดับ (multi-level structural equation model: MSEM) ซึ่งเป็นโมเดลที่ออกแบบมาเพื่อใช้ในการวิเคราะห์โมเดลสมการโครงสร้าง (structural equation model: SEM) สำหรับข้อมูลแบบพหุระดับ กล่าวคือโมเดลดังกล่าวพัฒนาต่อยอดโดยบูรณาการโมเดลเชิงเส้นพหุระดับ (multi-level linear model) เข้ากับโมเดล SEM ในทางทฤษฎี Muthen และ Asparouhov (2010) เสนอว่า สำหรับข้อมูลพหุระดับแบบสองระดับ โมเดลการวิเคราะห์ประกอบไปด้วยสองส่วนได้แก่ โมเดลการวัด (measurement model) และโมเดลสมการโครงสร้าง (structural model) เมื่อ $i = 1,2, \dots, n_j$ และ $j = 1,2, \dots, J$ จะได้ว่า

<u>โมเดลระดับที่ 1</u>: *โมเดลการวัด (measurement model)*

$$Y_{ij} = \nu_j + \Lambda_j \eta_{ij} + K_j X_{ij} + \varepsilon_{ij} \qquad (1.1)$$

โดยที่ Y_{ij} คือเวกเตอร์ของตัวแปรสังเกตได้ของหน่วยตัวอย่างที่ i ภายในกลุ่มที่ j, ν_j คือพารามิเตอร์จุดตัดแกน (measurement intercept) ของกลุ่มที่ j, Λ_j คือเมทริกซ์ของพารามิเตอร์น้ำหนักองค์ประกอบ (factor loading matrix), η_{ij} คือเวกเตอร์ของตัวแปรแฝงของหน่วยตัวอย่างที่ i ภายในกลุ่มที่ j, K_j คือเมทริกซ์ของสัมประสิทธิ์ความถดถอยของตัวแปรภูมิหลัง (regression coefficient matrix of background variables), X_{ij} คือเมทริกซ์ของตัวแปรภูมิหลัง (background variables matrix) ของหน่วยตัวอย่างที่ i ภายในกลุ่มที่ j และ ε_{ij} เวกเตอร์ของความคลาดเคลื่อนจากการวัด (measurement error vector) ของหน่วยตัวอย่างที่ i ภายในกลุ่มที่ j

<u>โมเดลระดับที่ 1</u>: *โมเดลสมการโครงสร้าง (structural model)*

$$\eta_{ij} = \alpha_j + \mathrm{B_j}\eta_{\mathrm{ij}} + \Gamma_{\mathrm{j}}X_{\mathrm{ij}} + \zeta_{\mathrm{ij}} \tag{1.2}$$

โดยที่ α_j เวกเตอร์ของพารามิเตอร์จุดตัดแกนของกลุ่มที่ j , $\mathrm{B_j}$ คือเมทริกซ์ของสัมประสิทธิ์ความถดถอยระหว่างตัวแปรแฝงของกลุ่มที่ j, Γ_{j} คือเมทริกว์ของสัมประสิทธิ์ความถดถอยระหว่างตัวแปรตามแฝงกับตัวแปรภูมิหลัง และ ζ_{ij} คือความคลาดเคลื่อนของโมเดลในระดับที่หนึ่ง (level-1 model errors)

<u>โมเดลระดับที่ 2</u>: *โมเดลสมการโครงสร้าง (structural model)*

$$\eta_j = \mu + \beta\eta_{\mathrm{j}} + \gamma X_{\mathrm{j}} + \zeta_{\mathrm{j}} \tag{1.3}$$

โดยที่ η_j คือเวกเตอร์ของตัวแปรแฝงในกลุ่มที่ j, μ เวกเตอร์ของพารามิเตอร์จุดตัดแกนในระดับที่สอง, β คือเมทริกซ์ของสัมประสิทธิ์ความถดถอยในระดับที่สองระหว่างตัวแปรแฝง, γ คือเมทริกซ์ของสัมประสิทธิ์ความถดถอยในระดับที่สองระหว่างตัวแปรตามแฝงและตัวแปรภูมิหลัง และ ζ_{j} คือเวกเตอร์ของความคลาดเคลื่อนของโมเดลในระดับที่สอง

จากโมเดลในข้างต้นจะเห็นว่าโมเดลสมการโครงสร้างพหุระดับมีความสามารถในการแก้ไขปัญหาความคลาดเคลื่อนจากการวัดที่เกิดขึ้นในข้อมูลพหุระดับได้อย่างเหมาะสม เนื่องจากได้มีการคำนึงถึงความผันแปรที่เกิดขึ้นจากความแตกต่างของหน่วย (Aspatouhov และ Muthen, 2011) การวิเคราะห์โมเดลในข้างต้นสามารถวิเคราะห์ได้โดยอาศัยโปรแกรม Mplus โปรแกรมดังกล่าวเป็นโปรแกรมสำเร็จรูปสำหรับพัฒนาโมเดลทางสถิติที่ได้รับความนิยมจากนักวิชาการในปัจจุบัน เนื่องด้วยเป็นโปรแกรมที่ถูกออกแบบให้สามารถใช้งานได้โดยง่ายสำหรับนักวิจัยโดยทั่วไป และมีความสามารถที่ครอบคลุมการวิเคราะห์รูปแบบของโมเดลทางสถิติขั้นสูงหลายโมเดลรวมถึงโมเดลสมการโครงสร้างพหุระดับ นอกจากนี้ Mplus ยังมีคุณสมบัติพิเศษอื่นๆ เช่น การประมาณค่าข้อมูลขาดหาย และการจำลองข้อมูลโดยใช้เทคนิค Monte Carlo เป็นต้น (Muthén and Muthén, 2010; Muthén and Asparouhov, 2008; Muthén and Muthén, 2003)

โมเดลการวัดซึ่งเป็นโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันในสมการที่ (1.1) จะเห็นว่าเป็นโมเดลที่ยังไม่สามารถระบุได้ (unidentified model) เนื่องจากมีจำนวนพารามิเตอร์มากเกินไป การแก้ปัญหาดังกล่าวสามารถกระทำได้โดยการกำหนดเงื่อนไข (restriction) ให้กับพารามิเตอร์บางตัวในโมเดล เมื่อพิจารณาสมการที่ (1.1) จะพบว่าสามารถกำหนดเงื่อนไขให้กับพารามิเตอร์ในโมเดลได้หลากหลายรูปแบบ ซึ่งในปัจจุบันโปรแกรม Mplus version 6.0 ถูกออกแบบมาเพื่อสามารถวิเคราะห์สมการโครงสร้างพหุระดับที่มีโมเดลการวัดแบบพหุระดับในลักษณะที่อาจเรียกว่า โมเดลการวัดแบบแบบสัมประสิทธิ์จุดตัดแกนแบบสุ่ม (random intercept measurement model) ซึ่งมีสมการดังต่อไปนี้

<u>โมเดลการวัดระดับที่หนึ่ง</u>

$$\underline{y}_{ij} = \underline{\mu}_j + \Lambda\underline{\omega}_{ij} + \underline{\epsilon}_{ij} \tag{1.4}$$

<u>โมเดลการวัดระดับที่สอง</u>

$$\underline{\mu}_j = \underline{\kappa} + \Lambda_b\underline{\varphi}_j + \underline{e}_j \tag{1.5}$$

ข้อสมมติเบื้องต้นของโมเดลกำหนดให้ $\underline{\epsilon}_{ij} \sim N\left(0, \Psi_{\epsilon j}\right)$, $\underline{\omega}_{ij} \sim N\left(\underline{0}, \Sigma_{\omega j}\right)$ โดยที่

$$\Sigma_{\omega j} = \begin{bmatrix} \underline{\beta}_j\Phi_j\underline{\beta}_j^T + \sigma_\delta^2 & \underline{\beta}_j^T\Phi_j \\ \underline{\beta}_j\Phi_j & \Phi_j \end{bmatrix} \tag{1.6}$$

$\underline{e}_j \sim N\left(\underline{0}, \Psi_b\right)$ โดยที่ $\Psi_{b\epsilon}$ เป็นเมทริกซ์ความแปรปรวนร่วมทแยงมุมของความคลาดเคลื่อนจากการวัดในระดับที่สอง และ $\underline{\varphi}_j \sim N\left(\underline{0}, \Phi_b\right)$ จากข้อสมมติดังกล่าวจะได้ว่า $\underline{\mu}_j \sim N\left(\underline{\alpha}, \Sigma_b\right)$ โดยที่ $\Sigma_b = \Lambda_b\Phi_b\Lambda_b^T + \Psi_b$

จากสมการที่ (1) จะสามารถเขียนสมการรวมได้ดังนี้

$$\underline{y}_{ij} = \underline{\kappa} + \Lambda_b\underline{\varphi}_j + \Lambda\underline{\omega}_{ij} + \underline{e}_j + \underline{\epsilon}_{ij} \tag{1.7}$$

จากข้อสมมติในข้างต้นจะได้ว่า

$$E\left(\underline{y}_{ij}\right) = \underline{\kappa} \tag{1.8}$$

$$Cov\left(\underline{y}_{ij} | \Phi_j, \Psi_{\epsilon j}\right) = \Lambda_b\Phi_b\Lambda_b^T + \Lambda\Sigma_{\omega j}\Lambda^T + \Psi_b + \Psi_{\epsilon j} \tag{1.9}$$

อย่างไรก็ตามเมื่อพิจารณางานวิจัยของ Ansari และ Jedidi (2002) พบว่างานวิจัยดังกล่าวได้เสนอรูปแบบโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันแบบพหุระดับ ซึ่งสามารถแบ่งประเภทออก

ได้เป็นสองประเภทใหญ่ได้แก่ 1) โมเดลการวิเคราะห์องค์ประกอบแบบพหุระดับที่มีความผันแปรในโครงสร้างของค่าเฉลี่ย (heterogeneous in mean structure factor analysis model) และ 2) โมเดลการวิเคราะห์องค์ประกอบที่มีความผันแปรทั้งในโครงสร้างของค่าเฉลี่ย ความแปรปรวนและความแปรปรวนร่วม (heterogeneous in mean and covariance structure factor analysis model) กรณีหลังนี้พบว่าสามารถเกิดขึ้นได้ในทางปฏิบัติกล่าวคือ ในกรณีที่ผู้วิจัยมีวัตถุประสงค์ที่ต้องการวัดและศึกษาความสัมพันธ์ในรูปแบบต่างๆของตัวแปรแฝงทั้งในระดับภายในกลุ่ม (within level) และระดับระหว่างกลุ่ม (between level) โมเดลการวัดแบบโมเดลการวิเคราะห์องค์ประกอบพหุระดับในโครงสร้างของค่าเฉลี่ยจะเป็นโมเดลที่สามารถตอบโจทย์การวิจัยนี้ได้ แต่ในกรณีที่จำนวนหน่วยมีจำนวนมากจะพบว่าคุณสมบัติความไม่แปรเปลี่ยนในการวัด (measurement invariance) มักถูกละเมิด กล่าวคือพารามิเตอร์ในโมเดลการวัดไม่ควรถูกกำหนดให้มีค่าคงที่ระหว่างหน่วย โมเดลการวัดที่เหมาะสมกับสถานการณ์ดังกล่าวอาจเป็นโมเดลการวัดที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่ม (random factor loadings and random measurement errors variance model) กรณีดังกล่าวจะเห็นว่าค่าเฉลี่ย ความแปรปรวนและความแปรปรวนร่วมของตัวแปรสังเกตได้ในโมเดลการวัดจะมีความผันแปรตามหน่วยตัวอย่างในระดับที่ 2 ซึ่งทำให้จัดประเภทของโมเดลการวัดในกรณีนี้อยู่ในจำพวกโมเดลการวัดที่มีความผันแปรทั้งในโครงสร้างของค่าเฉลี่ย ความแปรปรวนและความแปรปรวนร่วมดังที่กล่าวไว้ในข้างต้น กรณีนี้พบว่าหากผู้วิจัยในวิธีการประมาณค่าพารามิเตอร์แบบดั้งเดิม เช่น วิธีประมาณค่าแบบภาวะความควรจะเป็นสูงสุด หรือ วิธีกำลังสองน้อยสุดแบบถ่วงน้ำหนัก นั้นจะไม่สามารถประมาณค่าพารามิเตอร์ในโมเดลดังกล่าวได้ สาเหตุเนื่องด้วยโมเดลดังกล่าวมีจำนวนพารามิเตอร์จำนวนมาก ทำให้จำนวนมิติของการอินทิเกรตเพื่อหาผลเฉลยของค่าประมาณพารามิเตอร์มีจำนวนมากเกินไปจนไม่สามารถคำนวณค่าได้ (Asparouhov และ Muthen, 2012) ปัญหาดังกล่าวพบในโปรแกรม Mplus version 6.0 ทำให้ผู้วิเคราะห์ไม่สามารถวิเคราะห์โมเดลดังกล่าวได้ การแก้ไขปัญหาในข้างต้นทำได้โดยเปลี่ยนแนวคิดของวิธีการประมาณค่าพารามิเตอร์ ซึ่งวิธีการที่มีความสามารถแก้ไขปัญหาดังกล่าวคือวิธีการประมาณค่าแบบเบส์ (bayesian parameter estimation) ซึ่งยังไม่สามารถกระทำได้ในปัจจุบันด้วยโปรแกรม Mplus version 6.0

วิธีการประมาณค่าแบบเบส์มีข้อดีเหนือวิธีการแบบดั้งเดิมหลายประการ และมีหลักคิดในการแก้ปัญหาที่แตกต่างออกไป โดยมีข้อสมมติว่าพารามิเตอร์ในโมเดลที่สนใจเป็นตัวแปรสุ่ม (random variable) ที่อธิบายพฤติกรรมได้โดยใช้การแจกแจงความน่าจะเป็น (probability distribution) ส่วนข้อมูลค่าสังเกตจะถือว่าเป็นค่าคงที่ ดังนั้นในการประมาณค่าพารามิเตอร์ในโมเดลตามแนวทางสถิติแบบเบส์แท้จริงแล้วคือการประมาณการแจกแจงความน่าจะเป็นของพารามิเตอร์ดังกล่าวนั้นเอง การแจกแจงความน่าจะเป็นของพารามิเตอร์ที่ใช้เป็นเครื่องมือในการอนุมานเชิงสถิติในสถิติแบบเบส์จะเรียกว่า การแจกแจงความน่าจะเป็นภายหลัง (posterior distribution) การแจกแจงความน่าจะเป็นภายหลังนี้คำนวณได้จากการใช้ส่วนประกอบสองส่วนได้แก่ การแจกแจงความน่าจะเป็นก่อนหน้า (prior distribution) ซึ่งบรรจุสารสนเทศความรู้ก่อนหน้า (prior knowledge) ที่มีเกี่ยวกับพารามิเตอร์ที่สนใจและฟังก์ชันภาวะความควรจะเป็น (likelihood function) ซึ่งเป็นส่วนที่บรรจุสารสนเทศจากข้อมูลตัวอย่างและกำหนดรูปแบบของโมเดลที่ใช้ในการอธิบายปรากฏการณ์ที่สนใจ เครื่องมือที่ใช้เชื่อมสารสนเทศทั้งสองส่วนนี้เข้าด้วยกันจนกลายเป็นการแจกแจงความน่าจะเป็นภายหลังคือทฤษฎีของเบส์ (Bayes' theorem) ในทฤษฎีความน่าจะเป็น (probability theory) ดังสมการที่ (1.4)

$$p\left(\underline{\theta}|\underline{y}\right) = \frac{p\left(\underline{\theta},\underline{y}\right)}{p\left(\underline{y}\right)} = \frac{p\left(\underline{y}|\underline{\theta}\right)p\left(\underline{\theta}\right)}{p\left(\underline{y}\right)} \qquad (1.10)$$

เมื่อ $p(\underline{y}) = E\left(p(\underline{y};\underline{\theta})\right)$ ซึ่งมีค่าเท่ากับ $\iint_{\Theta}\cdots\int p(\underline{y}\,|\,\underline{\theta})p(\underline{\theta})d\underline{\theta} = \iint_{\Theta}\cdots\int p(\underline{\theta},\underline{y})d\underline{\theta}$ ในกรณีที่เป็นตัวแปรสุ่มต่อเนื่อง หรือเท่ากับ $\sum\sum_{\Theta}\cdots\sum p(\underline{y}\,|\,\underline{\theta})p(\underline{\theta}) = \sum\sum_{\Theta}\cdots\sum p(\underline{\theta},\underline{y})$ ในกรณีที่เป็นตัวแปรสุ่มไม่ต่อเนื่อง หลังจากทราบการแจกแจงความน่าจะเป็นภายหลัง ($p\left(\underline{\theta}|\underline{y}\right)$) ของพารามิเตอร์ที่ต้องการแล้ว ผู้วิจัยสามารถสรุปสารสนเทศจากการแจกแจงความน่าจะเป็นภายหลังดังกล่าวเพื่อใช้เป็น ค่าประมาณพารามิเตอร์ หรือเพื่อใช้อนุมานเกี่ยวกับค่าพารามิเตอร์ที่ต้องการได้โดยตรง อย่างไรก็ตามในทางปฏิบัติการคำนวณการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ในโมเดลส่วนใหญ่นั้นไม่สามารถคำนวณได้โดยตรงด้วยวิธีเชิงคณิตศาสตร์ สาเหตุเนื่องจากเมื่อโมเดลมีความซับซ้อนขึ้นจะทำให้พารามิเตอร์ในโมเดลจะมีจำนวนมากขึ้นจนส่งผลให้พจน์ค่าคงที่ $p\left(\underline{y}\right)$ มีจำนวนมิติที่ต้องอินทิเกรตมากเกินไป วิธีการที่นิยมใช้เพื่อแก้ไขปัญหาดัง

กล่าวคือการใช้วิธีการเชิงจำลอง (simulation base method) ซึ่งเทคนิคหนึ่งที่เป็นที่นิยมใช้อย่างกว้างขวางในปัจจุบันเรียกว่าเทคนิคลูกโซ่มาร์คอฟมอนติคาร์โล (Markov Chain Monte Carlo: MCMC) (Albert, 2009; Raudenbush & Bryk, 2002; Browne, 1998; Gelman, Carlin, Stern, and Rubin, 1995) ซึ่งจะกล่าวรายละเอียดของวิธีการนี้ต่อไปในวิทยานิพนธ์ฉบับนี้

ในปัจจุบันสถิติแบบเบส์เป็นแนวทางที่ได้รับความนิยมมากยิ่งขึ้นเนื่องด้วยมีข้อดีหลายประการดังนี้ (Albert, 2009; Bolstad, 2004; Browne, Draper, Goldstein, and Rasbash, 2002; Song, and Lee, 2001; Sc'neines, Hoijtink, and Boomsma, 1999; Browne, 1998; Gelman, Carlin, Stern, and Rubin, 1995; Iversen, 1989) ประการแรก สถิติแบบเบส์อาจกล่าวได้ว่าเป็นสถิติที่มีกระบวนทัศน์ที่สมบูรณ์ (complete paradigm) เนื่องจากเป็นแนวทางที่ใช้สารสนเทศจาก 2 แหล่งร่วมกันในกระบวนการวิเคราะห์ สารสนเทศทั้ง 2 แหล่ง คือ สารสนเทศจากความรู้ก่อนหน้า (prior knowledge) และสารสนเทศจากข้อมูลเชิงประจักษ์ (empirical knowledge) สารสนเทศจากข้อมูลก่อนหน้าจะมีความเป็นอัตนัย (subjective) เนื่องจากถูกกำหนดโดยผู้วิเคราะห์ก่อนการพิจารณาข้อมูลเชิงประจักษ์ ในขณะที่สารสนเทศจากข้อมูลเชิงประจักษ์จะมีความเป็นปรนัย (objective) การรวมสารสนเทศทั้งสองแหล่งเข้าด้วยกันจึงทำให้ผลการวิเคราะห์ที่ได้มีความยืดหยุ่นมากกว่าการวิเคราะห์แบบดั้งเดิม ในกรณีที่นักวิจัยมีความรู้ก่อนหน้าน้อยหรือยังไม่มีความชัดเจนเกี่ยวกับพารามิเตอร์ที่สนใจจะศึกษาอย่างมากเพียงพอนักวิจัยอาจเลือกใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศ (non-informative prior distribution) ซึ่งส่งอิทธิพลต่อผลการวิเคราะห์น้อยมากหรือไม่มีเลย กรณีนี้ผลการวิเคราะห์ที่ได้มีความใกล้เคียงกับผลการวิเคราะห์โดยใช้สถิติแบบดั้งเดิม แต่ในกรณีที่นักวิจัยมีความรู้ก่อนหน้าที่น่าเชื่อถือและชัดเจนนักวิจัยอาจเลือกใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบให้สารสนเทศ (informative prior distribution) ผลการวิเคราะห์ที่ได้นั้นจะได้รับอิทธิพลจากทั้งความรู้ก่อนหน้าที่ใส่เข้าไปและจากข้อมูลเชิงประจักษ์ ทำให้ผลการวิเคราะห์ที่ได้มีความแตกต่างจากสถิติแบบดั้งเดิม หากการแจกแจงความน่าจะเป็นก่อนหน้าที่ผู้วิเคราะห์กำหนดมีความเหมาะสมแล้วผลการวิเคราะห์ที่ได้ย่อมดีกว่าผลการวิเคราะห์ด้วยสถิติแบบดั้งเดิม

ประการที่สอง ในการอนุมานเกี่ยวกับพารามิเตอร์ที่สนใจนั้น สถิติแบบเบส์จะทำการอนุมานเชิงสถิติโดยใช้การแจกแจงความน่าจะเป็นภายหลัง (posterior distribution) เป็นเครื่องมือ การแจกแจงความน่าจะเป็นดังกล่าวเป็นการแจกแจงความน่าจะเป็นของพารามิเตอร์หลังจากที่ได้ปรับความเชื่อเกี่ยวกับความรู้ก่อนหน้าของพารามิเตอร์ด้วยข้อมูลเชิงประจักษ์แล้ว จะเห็นว่าหลักการอนุมานเชิงสถิติแบบเบส์มีแตกต่างจากสถิติแบบดั้งเดิมที่อาศัยการแจกแจงความน่าจะเป็นของตัวอย่างสุ่ม (sampling distribution) ซึ่งเป็นการแจกแจงความน่าจะเป็นของตัวสถิติที่ได้จากตัวอย่างสุ่มเป็นเครื่องมือในการอนุมานเชิงสถิติ อาจกล่าวได้ว่าสถิติแบบดั้งเดิมมีแนวคิดในการอนุมานโดยใช้สารสนเทศจากปริภูมิตัวอย่าง (sample space) เพื่อไปอนุมานพารามิเตอร์ในปริภูมิพารามิเตอร์ (parameter space) ในขณะที่สถิติแบบเบส์นั้นมีแนวคิดในการอนุมานโดยใช้สารสนเทศจากปริภูมิพารามิเตอร์ เพื่ออนุมานพารามิเตอร์ที่อยู่ในปริภูมิพารามิเตอร์เดียวกัน จากคำกล่าวในข้างต้นจะเห็นได้ชัดเจนกว่าการแปลผลจากการใช้สถิติแบบเบส์จะมีความตรงไปตรงมากและเข้าใจง่ายมากกว่าสถิติแบบดั้งเดิม เพราะเครื่องมือในการอนุมานเชิงสถิตินั้นอยู่ในปริภูมิเดียวกับพารามิเตอร์ที่สนใจ

ประการที่สาม หลักการของสถิติแบบเบส์นั้นสามารถนำไปใช้แก้ปัญหาทางสถิติที่ซับซ้อนได้ง่ายกว่าและมีความสะดวกกว่าสถิติแบบดั้งเดิม ทำให้สามารถแก้ไขปัญหาทางสถิติที่สถิติแบบดั้งเดิมไม่สามารถกระทำได้ (Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995) ยกตัวอย่างเช่น การแก้ปัญหาในโมเดลพหุระดับ (Browne, and Draper, 2000; Browne, 1998) การแก้ไขปัญหาในโมเดลพหุระดับที่มีความแปรปรวนซับซ้อนในเทอมความคลาดเคลื่อนของโมเดลระดับที่ 1 (multilevel model with complex variation in level-1 errors) (Browne, Draper, Goldstein, and Rasbash, 2002) การแก้ไขปัญหาความคลาดเคลื่อนจากการวัด (Goldstein, Kounali, and Robinson, 2008; Browne, Goldstein, Woodhouse, and Yang, 2001; Nounou et al., 2001) การแก้ปัญหาในโมเดลการวิเคราะห์องค์ประกอบ (factor analysis) (Ansari and Jedidi, 2002; Song, and Lee, 2001; Lee, 1981) การแก้ปัญหาในโมเดลสมการโครงสร้าง (Dunson, Palomo, and Bollen, 2005; Sc’neines, Hoijtink, and Boomsma, 1999) การแก้ปัญหาในโมเดลการตอบสนองข้อสอบ (Chaimongkol, Huffer, and Kamata, 2006; Fox, 2005; Fox and Glas, 2001) เป็นต้น

ประการสุดท้าย ในกรณีที่ต้องการทดสอบสมมติฐานเพื่อเลือกโมเดลที่เหมาะสมนั้น (model selection) หรืออาจเรียกว่าการทดสอบสมมติฐานเพื่อเปรียบเทียบโมเดลนั้น (model comparison) จะพบว่าการดำเนินการทดสอบสมมติฐานตามแนวทางของสถิติแบบดั้งเดิมนั้นเกิดปัญหาขึ้นหลายประการ ได้แก่ 1) การทดสอบสมมติฐานเพื่อทดสอบความสอดคล้องของโมเดลตามแนวทางสถิติแบบดั้งเดิมนั้น ตัวสถิติทดสอบจะมีโอกาสที่จะปฏิเสธสมมติฐานหลักมากขึ้นเมื่อขนาดตัวอย่างมีขนาดใหญ่ ถึงแม้ว่าในความเป็นจริงแล้วโมเดลดังกล่าวจะมีความสอดคล้องกับข้อมูลเชิงประจักษ์ก็ตาม 2) สถิติที่ใช้ในการทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์นั้นในความเป็นจริงแล้วตรรกะของการทดสอบดังกล่าวเป็นการทดสอบเพื่อมุ่งหวังที่จะปฏิเสธสมมติฐานหลัก ดังนั้นการที่ไม่สามารถปฏิเสธสมมติฐานหลักนั้นไม่ได้หมายความว่าโมเดลการวิจัยที่เสนอนั้นมีความสอดคล้องกับข้อมูลเชิงประจักษ์หรือโมเดลตามสมมติฐานหลักนั้นเป็นโมเดลที่เหมาะสม เพียงแต่ไม่มีหลักฐานที่เพียงพอที่จะสามารถกล่าวได้อย่างมีนัยสำคัญว่าโมเดลดังกล่าวไม่สอดคล้องกับข้อมูลเชิงประจักษ์เท่านั้น ดังนั้นการไม่สามารถปฏิเสธสมมติฐานหลักดังกล่าวจึงไม่อาจใช้เป็นหลักฐานที่หนักแน่นเพื่อแสดงความสอดคล้องของโมเดลการวิจัยกับข้อมูลเชิงประจักษ์ได้ 3) การทดสอบเพื่อเปรียบเทียบโมเดลตามแนวทางของสถิติแบบดั้งเดิมนั้นไม่สามารถที่จะใช้ทดสอบเพื่อเปรียบเทียบระหว่างโมเดลที่ไม่ซ้อนกันได้ (non-nested models) ปัญหาดังกล่าวเกิดขึ้นเพราะในการทดสอบสมมติฐานด้วยสถิติแบบดั้งเดิมนั้นใช้หลักการของการสร้างเกณฑ์การทดสอบด้วยวิธีอัตราส่วนภาวะความควรจะเป็น (likelihood ratio) ซึ่งจำเป็นที่จะต้องตั้งข้อสมมติไว้เบื้องต้นก่อนว่าสมมติฐานหลักนั้นเป็นจริงจากนั้นจึงสร้างเกณฑ์ที่ใช้ในการปฏิเสธสมมติฐานหลักดังกล่าว ในขณะที่สถิติแบบเบส์ไม่จำเป็นต้องมีข้อสมมติดังกล่าวเพราะการทดสอบสมมติฐานด้วยสถิติแบบเบส์สามารถดำเนินการได้โดยใช้ตัวสถิติ Bayes factor ซึ่งใช้หลักการเปรียบเทียบความเหมาะสมระหว่างโมเดลสองโมเดลที่ต้องการเปรียบเทียบกันซึ่งแตกต่างจากสถิติแบบดั้งเดิม ทำให้ตัวสถิติ Baye factor สามารถใช้ในการเปรียบเทียบโมเดลที่ไม่ซ้อนกันได้ (Berger, 1985; Kass and Raftery, 1995)

เมื่อพิจารณางานวิจัยที่เกี่ยวกับโมเดลการวิเคราะห์องค์ประกอบและโมเดลสมการโครงสร้างที่ประมาณค่าด้วยวิธีการประมาณค่าแบบเบส์พบว่ามีนักสถิติได้พัฒนาอัลกอริทึมไว้

จำนวนหนึ่ง เช่น อัลกอริทึมของ Lee และ Song (2004, 2002, 2001) และ Lee (1981) สำหรับการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันและโมเดลสมการโครงสร้างซึ่งสามารถใช้ได้กับทั้งกรณีที่ตัวแปรสังเกตได้เป็นตัวแปรต่อเนื่องและไม่ต่อเนื่อง อัลกอริทึมของ Dunson และคณะ (Dunson, Palomo, & Bollen, 2005) เป็นต้น แนวคิดของวิธีการดังกล่าวผู้วิจัยสามารถนำมาต่อยอดเพื่อแก้ไขปัญหาการประมาณค่าพารามิเตอร์ในโมเดลสมการโครงสร้างพหุระดับที่โมเดลการวัดมีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่มได้ ซึ่งผู้วิจัยเชื่อว่าโมเดลการวิเคราะห์ดังกล่าวจะมีความสอดคล้องกับสภาพความเป็นจริงมากขึ้น ทำให้ผลการวิเคราะห์ที่ได้มีความถูกต้องและน่าเชื่อถือมากยิ่งขึ้น อย่างไรก็ตามยังไม่พบว่ามีงานวิจัยที่ได้ทำการพัฒนาวิธีการประมาณค่าแบบเบส์ในโมเดลสมการโครงสร้างพหุระดับที่โมเดลการวัดมีพารามิเตอร์แบบสุ่ม ในงานวิจัยดังกล่าวจึงมีวัตถุประสงค์หลักเพื่อพัฒนาวิธีการประมาณค่าพารามิเตอร์ในโมเดลสมการโครงสร้างพหุระดับที่โมเดลการวัดมีพารามิเตอร์แบบสุ่ม โดยเลือกกรณีที่พารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่มมาศึกษา ซึ่งเป็นกรณีที่สามารถเกิดขึ้นจริงในทางปฏิบัติ และโปรแกรม Mplus ไม่สามารถวิเคราะห์ได้ ซึ่งผู้วิจัยเชื่อว่าสำหรับงานวิจัยที่ไม่ได้มีวัตถุประสงค์ที่ต้องการวัดตัวแปรแฝงซึ่งอยู่ในระดับที่สอง แต่ต้องการศึกษาความสัมพันธ์ระหว่างตัวแปรแฝงในโมเดลสมการโครงสร้างโดยมีการคำนึงถึงความเป็นพหุระดับในโมเดลการวัด และโมเดลสมการโครงสร้าง วิธีการประมาณค่าพารามิเตอร์ภายใต้โมเดลการวิเคราะห์ดังกล่าวจะมีความสอดคล้องกับสภาพความเป็นจริงมากขึ้น ทำให้ผลการวิเคราะห์ที่ได้มีความถูกต้องและน่าเชื่อถือมากยิ่งขึ้น

คำถามวิจัย

1. การพัฒนาอัลกอริทึมการประมาณค่าในโมเดลสมการโครงสร้างพหุระดับที่พารามิเตอร์ในโมเดลเป็นแบบสุ่มและโมเดลการวัดมีน้ำหนักองค์ประกอบและความคลาดเคลื่อนจากการวัดเป็นแบบสุ่มจะสามารถกระทำได้อย่างไร

2. ความถูกต้องของวิธีการประมาณค่าแบบเบส์และความสามารถของวิธีการประมาณค่าพารามิเตอร์แบบเบส์กับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted จะมีความเหมือนความแตกต่างกันอย่างไรภายใต้สถานการณ์จำลองต่างๆ
3. เมื่อใช้วิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้นกับข้อมูลจริง ผลการวิเคราะห์และข้อสรุปที่ได้จะมีความแตกต่างจากการใช้วิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted หรือไม่ อย่างไร

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการประมาณในโมเดลสมการโครงสร้างพหุระดับที่พารามิเตอร์ในโมเดลเป็นแบบสุ่มและโมเดลการวัดมีน้ำหนักองค์ประกอบและความคลาดเคลื่อนจากการวัดเป็นแบบสุ่ม
2. เพื่อตรวจสอบความถูกต้องของวิธีการประมาณค่าแบบเบส์และเปรียบเทียบความสามารถของวิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้นกับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted ภายใต้สถานการณ์จำลองแบบมอนติคาร์โล
3. เพื่อทดลองใช้วิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่ผู้วิจัยพัฒนาขึ้นวิเคราะห์ข้อมูลจริงและเปรียบเทียบกับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted

ขอบเขตของการวิจัย

1) ขอบเขตของการวิจัยในส่วนของการพัฒนาวิธีการประมาณค่าพารามิเตอร์แบบเบส์

1.1 โมเดลที่ใช้ในการศึกษาจะใช้โมเดลสมการโครงสร้างแบบพหุระดับที่มีสองระดับโดยมีสมการดังนี้

<u>โมเดลการวัด (measurement model)</u>

$$\underline{y}_{ij} = \underline{\mu} + \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij} \qquad (1.11)$$

<u>โมเดลสมการโครงสร้างพหุระดับ (multi-level structural model)</u>

$$\eta_{ij} = \beta_j^T \underline{\Xi}_{ij} + \delta_{ij} = \underline{\gamma}^T \underline{\Xi}_{ij} + \underline{u}_j^T \underline{\Xi}_{ij} + \delta_{ij} \tag{1.12}$$

เมื่อ $\underline{\Xi}_{ij} = \left(\underline{1} \quad \underline{\xi}_{ij}\right)^T$ และ $\beta_j = \underline{\gamma} + \underline{u}_j$,

$\underline{\epsilon}_{ij} \sim N(\underline{0}, \Psi_{\epsilon j})$ โดยที่ $\Psi_{\epsilon j} = diag(\psi_{\epsilon kj})$,

$\underline{\omega}_{ij} = \left(\eta_{ij}, \underline{\xi}_{ij}\right)^T \sim N\left(\underline{\theta}_{\omega j}, \Sigma_{\omega j}\right)$

โดยที่ $\underline{\theta}_{\omega j} = \begin{pmatrix} \underline{\beta}_j^T \underline{\nu}_j \\ \underline{\nu}_j \end{pmatrix}$, $\Sigma_{\omega j} = \begin{bmatrix} \underline{\beta}_j \Phi \underline{\beta}_j^T + \sigma_\delta^2 & \underline{\beta}_j^T \Phi \\ \underline{\beta}_j \Phi & \Phi \end{bmatrix}$

และ $\delta_{ij} \sim N\left(\underline{0}, \sigma_\delta^2\right)$

1.2 วิธีการประมาณค่าที่ใช้จะใช้วิธีการประมาณค่าพารามิเตอร์แบบเบส์ โดยใช้เทคนิคลูกโซ่มาร์คอฟมอนติคาร์โล (Markov Chain Monte Carlo: MCMC) อัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ (Gibbs-sampling)

1.3 การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดล กำหนดให้ใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบวงศ์คู่สังยุค (conjugacy prior) และมีรายละเอียดดังนี้

1) $\underline{\mu} \sim Normal(\kappa, \Sigma_\mu)$
2) $\underline{\nu}_j \sim Normal(0, \Delta)$ โดยที่ Δ คือเมทริกซ์ความแปรปรวนร่วมก่อนหน้าขนาด $q_2 \times q_2$
3) $\underline{\Lambda}_{kj} \sim Normal\left(\underline{g}_{kj}, H_{kj}\right)$ โดยที่ $\underline{\Lambda}_{kj}$ คือเวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์น้ำหนักองค์ประกอบ Λ_j ที่เป็น free parameters
4) $\psi_{\epsilon kj}^{-1} \sim Gamma(\alpha_{kj}, \beta_{kj})$ โดยที่ $\psi_{\epsilon kj} \in \Psi_{\epsilon j}$
5) $\left(\sigma_\delta^2\right)^{-1} \sim Gamma(\alpha_\delta, \beta_\delta)$
6) $\Phi^{-1} \sim Wishart\left(\rho_\phi, R_\phi\right)$
7) $\Sigma_u^{-1} \sim Wishart(\rho_u, R_u)$
8) $\underline{\gamma} \sim Normal\left(\underline{\gamma}_0, \Sigma_{\gamma 0}\right)$

2) ขอบเขตของการวิจัยในส่วนของการศึกษาด้วยข้อมูลจำลอง

2.1 วิธีการประมาณค่าพารามิเตอร์ที่ใช้มี 2 วิธีการได้แก่ วิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่ผู้วิจัยพัฒนาขึ้น และวิธีการประมาณค่าพารามิเตอร์ภาวะความควรจะเป็นสูงสุดแบบ restricted

2.2 การกำหนดพารามิเตอร์สำหรับการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดล เนื่องจากผู้วิจัยต้องการเปรียบเทียบความสามารถของวิธีการประมาณค่าแบบเบส์กับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted ซึ่งวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted เป็นวิธีที่ไม่มีการผนวกการแจกแจงความน่าจะเป็นก่อนหน้าเข้าไว้ในการวิเคราะห์เช่นเดียวกับวิธีการประมาณค่าแบบเบส์ ดังนั้นเพื่อไม่ให้มีอิทธิพลของการแจกแจงความน่าจะเป็นก่อนหน้าเป็นตัวแปรรบกวนผลการวิจัยผู้วิจัยจึงเลือกการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดลให้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศ (non-informative prior distribution) ทั้งหมด

2.3 เนื่องด้วยข้อจำกัดทางด้านเวลา ในการศึกษาผู้วิจัยจึงกำหนดให้จำนวนตัวแปรตามแฝงเท่ากับ 1 ตัวแปร และจำนวนตัวแปรอิสระแฝงเท่ากับ 1 ตัวแปร

2.4 จากข้อสมมติของโมเดลจึงจำลองข้อมูลตัวแปรอิสระแฝงให้มีการแจกแจงแบบปกติ

2.5 จากข้อสมมติของโมเดลจึงจำลองข้อมูลความคลาดเคลื่อนสุ่มในโมเดลสมการโครงสร้างพหุระดับ และในโมเดลการวัดให้มีการแจกแจงความน่าจะเป็นแบบปกติ

2.6 จากการทบทวนวรรณกรรมพบว่าการกำหนดขนาดตัวอย่างสำหรับงานวิจัยที่มีการวิเคราะห์ข้อมูลแบบพหุระดับนั้น จำนวนตัวอย่างในระดับที่ 2 จะมีความสำคัญมากกว่าตัวอย่างในระดับที่ 1 (Hox and Mass, 2001; Muthen, 1991) Muthen (1991) เสนอว่าขนาดตัวอย่างที่เหมาะสมควรมีจำนวนตัวอย่างในระดับที่สอง อย่างต่ำตั้งแต่ 50-100 กลุ่ม ส่วน Preacher, Zhang, and Zyphur (2011) เสนอว่าจำนวนตัวอย่างในระดับที่ 2 ควรมีอย่างน้อย 20 กลุ่ม ประกอบกับคุณสมบัติของสถิติแบบเบส์ที่จะมีความสามารถเหนือกว่าสถิติแบบดั้งเดิมเมื่อขนาดตัวอย่างมีจำนวนน้อย เพราะไม่ได้ขึ้นกับทฤษฎีลู่เข้าเชิงกำกับ (asymptotic theoty) เช่นเดียวกับสถิติแบบดั้งเดิม ผู้วิจัยจึงกำหนดขนาด

ตัวอย่างทั้งสองระดับ โดยเลือกกำหนดให้จำนวนหน่วยตัวอย่างในระดับที่ 1 คงที่คือ 30 หน่วย และจำนวนตัวอย่างในระดับที่ 2 คือ 15, 30 และ 50 หน่วย

2.7 เนื่องจากโมเดลที่ใช้ในการศึกษาเป็นโมเดลที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นแบบสุ่ม จึงทำให้ค่าความเที่ยงที่ได้เป็นค่าความเที่ยงแบบสุ่มด้วย ในการสร้างข้อมูลจำลองของแต่ละสถานการณ์ต่างที่มีความเที่ยงแตกต่างกัน ผู้วิจัยจึงเลือกกำหนดระดับของความเที่ยงโดยใช้ค่าเฉลี่ย และจากธรรมชาติของค่าสถิติความเที่ยงที่มีขอบเขตอยู่ในช่วง [0,1] จึงกำหนดขอบเขตระดับของค่าเฉลี่ยความเที่ยงรวมที่ใช้ในการศึกษาให้มีค่าเท่ากับ 0.3, 0.5, 0.7 และ 0.9 โดยต้องการให้มีการกระจายของระดับความเที่ยงรวมที่ใช้ในการศึกษาให้ทั่วถึงทั้งขอบเขตที่เป็นไปได้ของความเที่ยง

2.8 การกำหนดพารามิเตอร์องค์ประกอบแบบสุ่มจะกำหนดโดยสุ่มตัวอย่างจากการแจกแจงความน่าจะเป็นแบบปกติ

$$\Lambda_{kj} \sim N(E[\Lambda_k], cov[\Lambda_k])$$

โดยที่ Λ_{kj} คือเวกเตอร์ของสมาชิกใน $E[\Lambda_k] = (1,0.8,0.6)^T$ และ

$$cov[\Lambda_k] = \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0.0064 & 0 \\ 0 & 0 & 0.0036 \end{pmatrix}$$

2.9 สร้างค่าพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่ม โดยกำหนดให้พารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดของตัวแปรสังเกตได้ในหน่วยที่ j มีค่าเท่ากันกล่าวคือ $\psi_{\epsilon kj} = \psi_{\epsilon k'j}\ \forall k \neq k'$ และกำหนดค่าโดยการคำนวณกลับจากสูตรของความเที่ยงรวมสำหรับกลุ่มที่ j

$$\rho_{mj}^2 = \frac{\left(\sum_k \lambda_{kmj}\right)^2}{\left(\sum_k \lambda_{kmj}\right)^2 + \sum_k \psi_{\epsilon kj}}$$

โดยที่ ρ_{mj}^2 คือค่าความเที่ยงของตัวแปรแฝงที่ m ของกลุ่มที่ j, $\lambda_{kmj} \in \Lambda_{kj}$ คือพารามิเตอร์น้ำหนักองค์ประกอบของตัวแปรแฝงที่ m ที่ส่งอิทธิพลต่อตัวแปรสังเกตได้ที่ k ของกลุ่มที่ j และ $\psi_{\epsilon kj} \in \Psi_{\epsilon j}$ คือความแปรปรวนของความคลาดเคลื่อนจากการวัดของตัวแปรสังเกตได้ที่ k

2.10 ข้อมูลที่ใช้ในการวิจัยส่วนนี้สร้างขึ้นโดยอาศัยเทคนิคมอนติคาร์โล ซึ่งจะกระทำซ้ำ 100 ครั้งต่อสถานการณ์ โปรแกรมที่ใช้ในการจำลองข้อมูลได้แก่ R และโปรแกรมที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ R, Mplus และ Microsoft Excel

2.11 พิจารณาความแม่นยำและประสิทธิภาพโดยใช้เกณฑ์ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (MSE) และค่าอัตราส่วนระหว่างค่าคลาดเคลื่อนกำลังสองของวิธีประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted ต่อวิธีประมาณค่าแบบเบส์ (RAMSE) โดยการพิจารณาแยกตามประเภทของพารามิเตอร์ในโมเดลซึ่งแบ่งออกได้เป็น 6 ประเภท ได้แก่ พารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่ม พารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่ม พารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝง พารามิเตอร์ของอิทธิพลคงที่ พารามิเตอร์ของความแปรปรวนของโมเดลระดับที่สอง และพารามิเตอร์ของความแปรปรวนของโมเดลระดับที่หนึ่ง เกณฑ์ MSE ใช้สำหรับพิจารณาความแม่นยำและประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์ที่สนใจ และเกณฑ์ RAMSE ใช้พิจารณาเปรียบเทียบขนาดของค่า MSE ว่าวิธีประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted มีค่า MSE มากกว่าวิธีประมาณค่าแบบเบส์กี่เท่า

รายละเอียดเงื่อนไขที่ใช้สำหรับสร้างข้อมูลจำลองข้างต้น สามารถแสดงสรุปไว้ในรูปแบบตารางในตารางที่ 1.1 ซึ่งจะเห็นว่ามีสถานการณ์จำลองที่ทำการศึกษา 12 สถานการณ์จำลอง

<u>ตารางที่ 1.1</u> รายละเอียดเงื่อนไขที่ใช้สำหรับการสร้างข้อมูลจำลอง

ค่าเฉลี่ยความเที่ยงรวม	จำนวนตัวแปรตามแฝง/จำนวนตัวแปรอิสระแฝง	จำนวนหน่วยตัวอย่างในระดับที่ 1	จำนวนหน่วยตัวอย่างในระดับที่ 2	จำนวนหน่วยตัวอย่างรวม
0.3	1/1	30	15	450
			30	900
			50	1500
0.5	1/1	30	15	450
			30	900
			50	1500
0.7	1/1	30	15	450

			30	900
			50	1500
0.9	1/1	30	15	450
			30	900
			50	1500

3) ขอบเขตของการวิจัยในส่วนของการศึกษาด้วยข้อมูลทุติยภูมิ

3.1ข้อมูลที่ใช้ในการศึกษาในส่วนนี้เป็นข้อมูลทุติยภูมิจากสถานการณ์จริงซึ่งจะใช้ข้อมูลจากวิทยานิพนธ์ระดับดุษฎีบัณฑิต สาขาวิชาวิธีวิทยาวิจัยทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เรื่อง “อิทธิพลของพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลและสุขภาวะครูที่มีต่อสุขภาวะของนักเรียน : โมเดลการปรับและการส่งผ่านพหุระดับ” (ถมรัตน์ ศิริภาพ, 2554)

3.2 เนื่องจากวิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้นนี้พัฒนาขึ้นภายใต้โมเดลแบบสัมประสิทธิ์ความถดถอยแบบสุ่ม (random coefficients model) ดังนั้นเพื่อให้กรอบแนวคิดการวิจัยในส่วนนี้สอดคล้องกับโมเดลของวิธีการประมาณค่าพารามิเตอร์ที่พัฒนาขึ้น ผู้วิจัยจึงตัวแปรที่เลือกมาเพื่อใช้ในการศึกษาจำนวน 3 ตัวแปรได้แก่

3.2.1 ตัวแปรสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) วัดได้จากตัวแปรสังเกตได้จำนวน 6 ตัวแปรได้แก่ เจตคติและอารมณ์เชิงบวก (AFF) ความเพลิดเพลิน (ENJ) อัตมโนทัศน์เชิงวิชาการ (SEL) ปัญหาทางสังคม (SOC) ความวิตกกังวล (WOR) และปัญหาสุขภาพกาย (PHY)

3.2.2 ตัวแปรพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) วัดได้จากตัวแปรสังเกตได้จำนวน 4 ตัวแปรได้แก่ การใช้อำนาจ-การร่วมมือ (DC) การร่วมมือ-การคล้อยตาม (CS) การคล้อยตาม-การต่อต้าน (SO) และการต่อต้าน-การใช้อำนาจ (OD)

3.2.3 ตัวแปรผลสัมฤทธิ์ทางการเรียนในระดับนักเรียน (ACH_w)พิจารณาจากเกรดเฉลี่ยสะสมของนักเรียน

3.3 โมเดลที่ใช้ในการวิเคราะห์แสดงได้เป็นสมการได้ดังนี้

<u>โมเดลการวัด</u>

$$\underline{y}_{ij} = \underline{\mu} + \Lambda_j \underline{\omega}_{ij} + \underline{\varepsilon}_{ij} \quad (1.13)$$

เมื่อ $\underline{y}_{ij} = (AFF, ENJ, SEL, SOC, WOR, PHY, DC, CS, SO, OD)^T$,

$\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_{10})^T$,

$\Lambda_j = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & \lambda_{14} & 1^* & \lambda_{16} & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 0^* & 0^* & \lambda_{21} & \lambda_{22} & \lambda_{23} & \lambda_{24} \end{bmatrix}_j^T$,

$\underline{\omega}_{ij} = (SWB_w, SIB_w)_{ij}^T$,

$\underline{\varepsilon}_{ij} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{10})_{ij}^T$ โดยที่ $\varepsilon_{kij} \sim N(\underline{0}, \psi_{\varepsilon k(j)})$

และ ε_{kij} กับ $\varepsilon_{ki'j'}$ เป็นอิสระซึ่งกันและกัน

<u>โมเดลระดับที่ 1 (level-1 model)</u>

$$SWB_{wij} = \Xi_{ij}\underline{\beta}_j + \delta_{ij} \quad (1.14)$$

เมื่อ $\Xi_{ij} = (ACH_w \quad SIB_w)_{ij}$, $\underline{\beta}_j = (\beta_1, \beta_2)_j^T$ และ $\delta_{ij} \sim N(0, \sigma_\delta^2)$

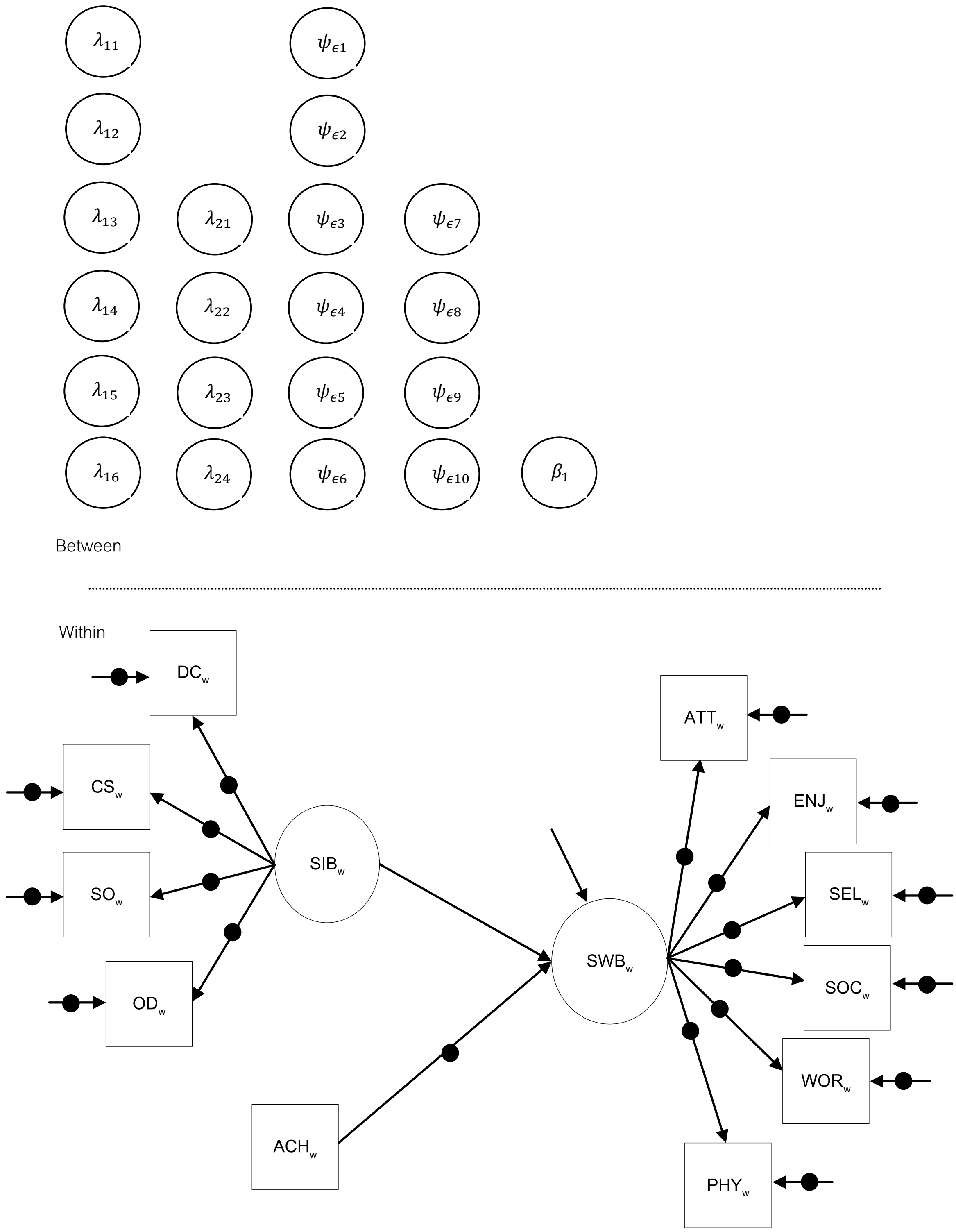
<u>โมเดลระดับที่ 2 (level-2 model)</u>

$$\beta_j = \underline{\gamma} + \underline{u}_j \quad (1.15)$$

เมื่อ $\underline{\gamma} = (\gamma_{01}, \gamma_{02})^T$, $\underline{u}_j = (u_{1j}, 0)^T \sim MVN(\underline{0}, \Sigma_u)$ โดยที่ $\Sigma_u = \begin{bmatrix} \tau_{11} & 0 \\ 0 & 0 \end{bmatrix}$

และ $\underline{u}_j$ กับ $\underline{u}_{j'}$ เป็นอิสระซึ่งกันและกัน

จากโมเดลในสมการข้างต้นสามารถเขียนเป็นรูปภาพได้ดังรูปที่ 1.1



λ_{11}
λ_{12}
λ_{13}
λ_{14}
λ_{15}
λ_{16}
λ_{21}
λ_{22}
λ_{23}
λ_{24}
$\psi_{\epsilon1}$
$\psi_{\epsilon2}$
$\psi_{\epsilon3}$
$\psi_{\epsilon4}$
$\psi_{\epsilon5}$
$\psi_{\epsilon6}$
$\psi_{\epsilon7}$
$\psi_{\epsilon8}$
$\psi_{\epsilon9}$
$\psi_{\epsilon10}$
β_1
Between
Within
DC_w
CS_w
SO_w
OD_w
SIB_w
ACH_w
SWB_w
ATT_w
ENJ_w
SEL_w
SOC_w
WOR_w
PHY_w

รูปที่ 1.1 โมเดลที่ใช้ในการวิเคราะห์ส่วนการศึกษาด้วยข้อมูลจริง

นิยามศัพท์เฉพาะที่ใช้ในการวิจัย

1. ข้อมูลพหุระดับ (multi-level data) คือ ข้อมูลที่มีลักษณะโครงสร้างที่หน่วยตัวอย่างในระดับที่ต่ำกว่าจะซ้อน (nested) อยู่ในหน่วยตัวอย่างในระดับที่สูงกว่าเป็นลำดับไป (hierarchies)

2. โมเดลเชิงเส้นพหุระดับ (multi-level linear model) คือ โมเดลเชิงเส้นตรงทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลพหุระดับ พารามิเตอร์ในโมเดลสามารถมีความผันแปรตามหน่วยตัวอย่างในระดับที่สูงกว่าซึ่งความผันแปรดังกล่าวอาจเป็นความผันแปรอย่างคงที่ (deterministic) หรืออย่างสุ่ม (random)

3. ความคลาดเคลื่อนจากการวัด (measurement error) คือ ส่วนประกอบในพจน์ที่เป็นความคลาดเคลื่อนของโมเดลการวัดแบบดั้งเดิม (classical test theory) โดยในงานวิจัยนี้จะหมายถึงส่วนความคลาดเคลื่อนจากการวัดแบบสุ่ม (random measurement error) เท่านั้น

4. สถิติแบบเบส์ (Bayesian statistics) คือ วิธีการทางสถิติที่อนุมานเกี่ยวกับค่าพารามิเตอร์ที่สนใจจากการแจกแจงความน่าจะเป็นภายหลัง (posterior distribution) ซึ่งสร้างโดยอาศัยการรวมสารสนเทศจากความรู้ก่อนหน้า (prior knowledge) ที่เก็บรวบรวมอยู่ในการแจกแจงความน่าจะเป็นก่อนหน้า (prior distribution) และสารสนเทศจากข้อมูลเชิงประจักษ์ที่เก็บรวบรวมอยู่ในฟังก์ชันภาวะความควรจะเป็น (likelihood function) โดยใช้หลักการจากทฤษฎีของเบส์ (Bayes's theorem) ในทฤษฎีความน่าจะเป็น (probability theory)

5. ฟังก์ชันภาวะความควรจะเป็น (likelihood function) เขียนแทนด้วยสัญลักษณ์ $f(x|\theta)$ เมื่อ θ เป็นพารามิเตอร์ในโมเดล และ $x=\{x_1, x_2,..., x_n\}$ เป็นข้อมูลค่าสังเกต คือรูปแบบความน่าจะเป็นร่วมของการเกิดขึ้นของข้อมูลค่าสังเกต x ซึ่งเป็นการแสดงโอกาสที่จะเกิดขึ้นของ

เหตุการณ์ที่สนใจทั้งหมดใน n มิติไปพร้อมกัน หรืออาจกล่าวได้ว่าเป็นรูปแบบความน่าจะเป็นที่อธิบายการเกิดขึ้นของชุดข้อมูลตัวอย่าง

6. วิธีภาวะความควรจะเป็นสูงสุด คือ วิธีหาค่าประมาณพารามิเตอร์ที่ทำให้ฟังก์ชันภาวะความควรจะเป็นมีค่าสูงสุด หรืออาจกล่าวได้ว่าเป็นวิธีหาค่าประมาณพารามิเตอร์ที่สอดคล้องกับข้อมูลเชิงประจักษ์มากที่สุด

7. วิธีภาวะความควรจะเป็นสูงสุดแบบ restricted คือ วิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดโดยที่การหาค่าประมาณค่าพารามิเตอร์ไม่ได้ขึ้นกับฟังก์ชันภาวะความควรจะเป็นแบบมีสารสนเทศเต็มรูป (full information likelihood function) แต่จะใช้ฟังก์ชันภาวะความควรของค่าสังเกตที่ผ่านการกระบวนแปลงทำให้พารามิเตอร์รบกวน (nuisance parameters) ไม่มีผลกระทบต่อฟังก์ชันภาวะความควรจะเป็น ยกตัวอย่างเช่น ในการประมาณค่าส่วนประกอบความแปรปรวน ข้อมูลสังเกตได้ที่ใช้ในการประมาณค่าส่วนประกอบความแปรปรวนจะใช้เศษเหลือ ซึ่งเศษเหลือดังกล่าวจะพบว่าไม่มีความสัมพันธ์กับพารามิเตอร์อิทธิพลคงที่ในโมเดล ดังนั้นฟังก์ชันภาวะความควรจะเป็นของเศษเหลือดังกล่าวจึงไม่ได้รับอิทธิพลรบกวนจากอิทธิพลคงที่ซึ่งเป็นตัวแปรรบกวน เรียกฟังก์ชันภาวะความควรจะเป็นดังกล่าวนี้ว่า ฟังก์ชันภาวะความควรจะเป็นแบบ restricted

8. การแจกแจงความน่าจะเป็นก่อนหน้า (prior distribution) เขียนแทนด้วยสัญลักษณ์ $p(\theta)$ เมื่อ θ เป็นพารามิเตอร์ในโมเดล คือ การแจกแจงความน่าจะเป็นที่ใช้ในการอธิบายสารสนเทศของข้อมูลก่อนหน้า (prior knowledge) ซึ่งจะนำไปรวมเข้ากับกระบวนการวิเคราะห์พารามิเตอร์ของการแจกแจงความน่าจะเป็นก่อนหน้าจะเรียกว่า ไฮเปอร์พารามิเตอร์ (hyperparameters) เพื่อให้มีความแตกต่างจากพารามิเตอร์ในโมเดล

9. การแจกแจงความน่าจะเป็นภายหลัง (posterior distribution) เขียนแทนด้วยสัญลักษณ์ $p(\theta|x)$ เมื่อ θ เป็นพารามิเตอร์ในโมเดล และ x ข้อมูลค่าสังเกตคือ การแจกแจงความน่าจะเป็นแบบมีเงื่อนไข (conditional probability distribution) ของพารามิเตอร์เมื่อกำหนดข้อมูลค่าสังเกต ซึ่งคำนวณได้จากทฤษฎีของเบส์ ดังนี้

$$p(\theta|x)\,\alpha\, p(\theta)\times f(x|\theta)$$

10. การแจกแจงความน่าจะเป็นก่อนหน้าแบบวงศ์คู่สังยุค (conjugacy prior) คือ การแจกแจงความน่าจะเป็นก่อนหน้าที่มีการแจกแจงความน่าจะเป็นอยู่ในวงศ์ (family) ของการแจกแจง D ที่มีพารามิเตอร์ α จะเป็นการแจกแจงที่เป็นวงศ์คู่สังยุคกับการแจกแจง $f(\underline{y}\,|\,\underline{\theta})$ ถ้าการแจกแจงความน่าจะเป็นภายหลัง $f(\underline{\theta}\,|\,\underline{y})$ เป็นการแจกแจงที่เป็นสมาชิกในวงศ์ของ D เช่นเดียวกัน กล่าวคือ ถ้า $\underline{\theta} \sim D(\alpha)$ แล้ว $\underline{\theta}\,|\,\underline{y} \sim D(\widetilde{\alpha})$

11. ลูกโซ่มาร์คอฟ (Markov Chain) คือกระบวนการเฟ้นสุ่ม (stochastic process) ประเภทหนึ่งซึ่งใช้อธิบายความไม่แน่นอนในการเปลี่ยนแปลงสถานะ จากสถานะหนึ่งไปยังอีกสถานะหนึ่งโดยใช้ความน่าจะเป็นของการเปลี่ยนสถานะ (transition probability) และความน่าจะเป็นของการเปลี่ยนสถานะนั้นมีคุณสมบัติที่ขึ้นกับสถานะในอดีตที่ติดกับสถานการณ์ปัจจุบันล่าสุดเพียงสถานะเดียวเท่านั้น

12. ลูกโซ่มาร์คอฟมอนติคาร์โล (Markov Chain Monte Carlo: MCMC) คือวิธีการที่ใช้ในการสุ่มตัวอย่างจากการแจกแจงความน่าจะเป็นที่ต้องการ โดยที่อาศัยคุณสมบัติของระบบลูกโซ่มาร์คอฟ ซึ่งหากมีการออกแบบให้ระบบลูกโซ่มาร์คอฟที่จำลองมีคุณสมบัติที่ดีบางประการจะทำให้ ตัวอย่างจากการสุ่มตามอัลกอริทึม MCMC ลู่เข้าสู่การแจกแจงความน่าจะเป็นสมดุล (equilibrium distribution) หรือการแจกแจงความน่าจะเป็นคงที่ (stationary distribution) เมื่อจำนวนตัวอย่างที่สุ่มมีขนาดใหญ่เพียงพอ (หรือจำนวนการเปลี่ยนแปลงสถานะ (state) ของลูกโซ่มาร์คอฟมีเพียงพอ)

13. การสุ่มตัวอย่างแบบกิบบ์ (Gibbs sampling) คือวิธีการที่ใช้ในการประมาณการแจกแจงความน่าจะเป็นภายหลังในกรณีที่ไม่สามารถคำนวณได้โดยตรง โดยการสุ่มตัวอย่างจากการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ทีละตัว ในแต่ละขั้นตอนของวิธีการ ซึ่งหลังจากดำเนินกระบวนการดังกล่าวด้วยจำนวนรอบของการทวนซ้ำที่เพียงพอแล้วตัวอย่างสุ่มที่ได้หลังจากนั้นจะเป็นตัวอย่างสุ่มที่สุ่มได้จากการแจกแจงความน่าจะเป็นภายหลังที่ต้องการ

ประโยชน์ที่ได้รับ

จากการศึกษาครั้งนี้ ทำให้เกิดประโยชน์ในเชิงวิชาการและในเชิงปฏิบัติ ดังนี้

<u>ประโยชน์ในเชิงวิชาการ</u>

1. สร้างองค์ความรู้ในทางสถิติเกี่ยวกับการประมาณค่าพารามิเตอร์แบบเบส์ในการวิเคราะห์ข้อมูลภายใต้โมเดลสมการโครงสร้างพหุระดับที่โมเดลการวัดมีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่ม

2. ได้ข้อค้นพบทำให้ทราบถึงความสำคัญของโมเดลการวัดที่มีพารามิเตอร์แบบสุ่ม ซึ่งจากงานวิจัยจะเห็นว่าการประมาณค่าพารามิเตอร์ภายใต้โมเดลการวัดที่มีพารามิเตอร์แบบคงที่ จะทำให้เกิดความคลาดเคลื่อนของค่าประมาณมากกว่าโมเดลการวัดที่มีพารามิเตอร์แบบสุ่มอย่างเห็นได้ชัดเจน

3. ได้ข้อค้นพบเกี่ยวกับข้อเด่นของการใช้สถิติแบบเบส์ กล่าวคือ สถิติแบบเบส์สามารถใช้ในการวิเคราะห์ข้อมูลภายใต้โมเดลที่สถิติแบบดั้งเดิมกระทำได้ลำบากหรือไม่สามารถกระทำได้ และสถิติแบบเบส์ยังมีข้อเด่นเหนือสถิติแบบดั้งเดิมในเรื่องของจำนวนหน่วยตัวอย่างที่ต้องการ ทำให้ทราบว่าการใช้สถิติแบบเบส์จะสามารถลดจำนวนหน่วยตัวอย่างที่ต้องการลงได้มากกว่าการใช้สถิติแบบดั้งเดิม

4. แนวคิดการประมาณค่าพารามิเตอร์ที่ได้พัฒนาขึ้นจากงานวิจัยนี้ สามารถพัฒนาขยายแนวคิดในงานวิจัยดังกล่าวเพื่อใช้ในโมเดลการวิเคราะห์ที่มีความซับซ้อนมากยิ่งขึ้นซึ่งสามารถทำการวิจัยต่อได้ในอนาคต เช่น โมเดลที่มีตัวแปรตามแฝงหลายตัวแปร โมเดลที่มีตัวแปรส่งผ่าน (mediator model) โมเดลที่มีตัวแปรปรับ (moderator model) เป็นต้น

<u>ประโยชน์ในเชิงปฏิบัติ</u>

1. จากวิธีการประมาณค่าพารามิเตอร์ที่พัฒนาขึ้นทำให้ในการวิจัยเชิงประยุกต์ ผู้วิจัยสามารถขยายขอบเขตการวิเคราะห์ภายใต้โมเดลที่ Mplus ยังไม่สามารถกระทำได้ ซึ่งจะทำให้การวิเคราะห์ข้อมูลสามารถกระทำได้อย่างถูกต้อง และมีความลึกซึ้งมากยิ่งขึ้น

2. การใช้สถิติแบบเบส์ในการวิเคราะห์ข้อมูลทำให้ผู้วิจัยสามารถนำความรู้ก่อนหน้าผนวกไว้ในกระบวนการวิเคราะห์ซึ่งจะทำให้ผลการวิเคราะห์ที่ได้มีความยืดหยุ่นมากกว่าการใช้สารสนเทศจากข้อมูลเชิงประจักษ์แต่เพียงอย่างเดียว

3. จากข้อค้นพบการใช้สถิติแบบเบส์ทำให้ผู้วิจัยสามารถลดขนาดตัวอย่างที่ต้องการลงได้ ทำให้เป็นการประหยัดทรัพยากรที่ใช้ในการวิจัย ทั้งด้านค่าใช้จ่าย และเวลา

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

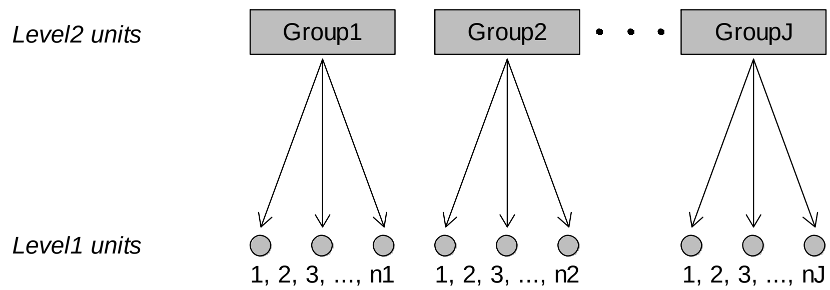
ในส่วนของแนวคิดทฤษฎีที่เกี่ยวข้องผู้วิจัยได้กำหนดกรอบของการนำเสนอซึ่งสามารถแบ่งออกได้เป็น 4 ตอนได้แก่ ตอนที่ 1 แนวคิดทฤษฎีที่เกี่ยวกับแนวคิดการวิเคราะห์ข้อมูลพหุระดับทางสังคมศาสตร์ ตอนที่ 2 แนวคิดทฤษฎีเกี่ยวกับสถิติแบบเบส์ ตอนที่ 3 แนวคิดทฤษฎีเกี่ยวกับการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับด้วยสถิติแบบเบส์ ตอนที่ 4 แนวคิดทฤษฎีที่เกี่ยวกับความคลาดเคลื่อนจากการวัด ตอนที่ 5 แนวคิดในการพัฒนาวิธีการประมาณค่าพารามิเตอร์ของผู้วิจัย และตอนที่ 6 กรอบแนวคิดการวิจัย รายละเอียดเป็นดังต่อไปนี้

ตอนที่ 1: แนวคิดและทฤษฎีเกี่ยวกับการวิเคราะห์ข้อมูลพหุระดับทางสังคมศาสตร์

ในส่วนนี้จะนำเสนอหลักทฤษฎีเกี่ยวกับลักษณะข้อมูลทางสังคมศาสตร์ โมเดลการวิเคราะห์ ข้อสมมติเบื้องต้นของโมเดลและวิธีการประมาณค่าพารามิเตอร์ในโมเดล โดยในส่วนของรายละเอียดของวิธีการประมาณค่าพารามิเตอร์นั้นจะกล่าวถึงเฉพาะวิธีการประมาณค่าพารามิเตอร์แบบดั้งเดิมเท่านั้น ส่วนวิธีการประมาณค่าพารามิเตอร์แบบเบส์จะนำเสนอภายหลังในตอนที่ 3 ซึ่งเป็นส่วนที่นำเสนอแนวคิดและทฤษฎีที่เกี่ยวกับสถิติแบบเบส์

1.1 <u>ข้อมูลพหุระดับและการวิเคราะห์ข้อมูลพหุระดับ</u>

ในการวิจัยทางสังคมศาสตร์ ทางการศึกษา หรือทางการแพทย์ ผู้วิจัยมักพบข้อมูลที่ลักษณะซ้อน (nested) หรือมีโครงสร้างในลักษณะลดหลั่นเป็นลำดับชั้น โครงสร้างของข้อมูลดังกล่าวจะมีลักษณะที่หน่วยตัวอย่างในระดับล่างจะซ้อนอยู่กับหน่วยตัวอย่างในระดับที่สูงกว่า เรียกข้อมูลที่มีลักษณะดังกล่าวว่า ข้อมูลพหุระดับ (multi-level data) ยกตัวอย่างเช่นในกรณีที่โครงสร้างของข้อมูลเป็น 2 ระดับ (two-level hierarchies) โครงสร้างของข้อมูลจะเป็นโครงสร้างของที่หน่วยตัวอย่างในระดับที่ 1 ซึ่งเป็นระดับล่างโดยอาจเรียกว่าระดับจุลภาค (micro- level) ซ้อนอยู่ในหน่วยตัวอย่างระดับที่ 2 ซึ่งอาจเรียกว่าในระดับมหภาค (macro-level) ลักษณะดังกล่าวแสดงไว้ในรูปที่ 2.1 ตัวอย่างของข้อมูลที่มีโครงสร้างแบบ 2 ระดับเช่น ข้อมูลของนักเรียนที่รวมกลุ่มอยู่ในแต่ละโรงเรียน ข้อมูลพนักงานรวมกลุ่มในแต่ละองค์กร ข้อมูลของครัวเรือนรวมกลุ่มในแต่ละจังหวัด หรือข้อมูลคนไข้ที่รวมกลุ่มในการได้รับการรักษาจากแพทย์แต่ละคน เป็นต้น โครงสร้างของข้อมูลอาจมีโครงสร้างมากกว่า 2 ระดับได้ เช่น หากเป็นงานวิจัยทางการศึกษาที่สนใจศึกษาเกี่ยวกับนักเรียน จะพบว่านักเรียนที่เป็นหน่วยตัวอย่างนั้นซ้อนอยู่ในห้องเรียน ห้องเรียนซ้อนอยู่ในโรงเรียน โรงเรียนซ้อนอยู่ในเขตพื้นที่การศึกษา เป็นต้น



<u>รูปที่ 2.1</u> ลักษณะโครงสร้างข้อมูลพหุระดับที่มี 2 ระดับ

หากข้อมูลในงานวิจัยมีโครงสร้างในลักษณะลดหลั่นเป็นลำดับชั้นดังที่กล่าวมาแล้วผู้วิจัยละเลยต่อโครงสร้างของข้อมูลในการวิเคราะห์ เช่น ใช้การวิเคราะห์การถดถอยแบบดั้งเดิมซึ่งเป็นการวิเคราะห์แบบระดับเดียว (single-level analysis) ผู้วิจัยจะไม่สามารถวิเคราะห์โดยใช้ตัวแปรที่อยู่ต่างระดับกันวิเคราะห์ร่วมกันได้ แต่จำเป็นที่จะต้องจัดกระทำ (manipulate) ตัวแปรต่างๆใหม่

เพื่อให้เป็นตัวแปรที่อยู่ในระดับเดียวกันและจึงทำการวิเคราะห์ การจัดกระทำตัวแปรดังกล่าวนั้นสามารถกระทำได้ 2 วิธีการคือ การกระจายตัวแปร (disaggregation) และการรวมตัวแปร (aggregation) (ศิริชัย กาญจนวาสี, 2550; Raudenbush & Bryk, 2002; Snider & Bosker, 1999)

วิธีการกระจายข้อมูลของตัวแปร (disaggregation) เป็นวิธีการที่ผู้วิจัยกระจายข้อมูลตัวแปรของหน่วยตัวอย่างในระดับที่สูงกว่าให้เป็นตัวแปรในระดับที่ต่ำกว่า กล่าวคือหน่วยตัวอย่างที่อยู่ในกลุ่มเดียวกันจะมีค่าของตัวแปรนั้นเท่ากันทั้งกลุ่ม จากนั้นจึงนำตัวแปรที่ได้ไปใช้ในการวิเคราะห์ที่ต้องการต่อไป วิธีการดังกล่าวสามารถกระทำได้หากหน่วยตัวอย่างในระดับที่ต่ำกว่านั้นไม่ได้มีความสัมพันธ์กันเองภายในกลุ่ม อย่างไรก็ตามโดยธรรมชาติของข้อมูลแบบพหุระดับแล้วมักพบว่าหน่วยตัวอย่างที่ได้รับการสุ่มจากกลุ่มย่อยต่างๆของประชากรนั้นมักมีความคล้ายคลึงกันภายในกลุ่มมากกว่าหน่วยตัวอย่างที่ได้รับการสุ่มจากประชากรเดียวที่ไม่ได้มีการแบ่งกลุ่มย่อย จากสถานการณ์ดังกล่าวหน่วยตัวอย่างที่อยู่ภายในกลุ่มเดียวกันนั้นจึงมักมีความสัมพันธ์กันเองซึ่งเป็นการละเมิดข้อกำหนดเบื้องต้นของโมเดลในการวิเคราะห์ข้อมูล ซึ่งจำทำให้ค่าความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ในโมเดลมีความผิดพลาดค่า (misestimated standard error) ซึ่งจะมีค่าที่ต่ำกว่าปกติเมื่อเปรียบเทียบกับสถานการณ์ที่หน่วยตัวอย่างไม่ได้มีความสัมพันธ์กัน ทำให้การประมาณค่าแบบช่วงนั้นได้ช่วงความเชื่อมั่นที่แคบกว่าปกติ และการทดสอบสมมติฐานเกี่ยวกับค่าพารามิเตอร์ที่สนใจจะมีโอกาสที่จะปฏิเสธสมมติฐานหลักมากกว่าปกติ ซึ่งจะทำให้ผลการวิเคราะห์ที่ได้มีความผิดพลาด นอกจากนี้การกระจายข้อมูลยังเป็นการไม่คำนึงถึงโครงสร้างของข้อมูลที่เป็นแบบพหุระดับซึ่งจะถือว่าค่าพารามิเตอร์ของโมเดลนั้นมีค่าเท่ากันระหว่างกลุ่ม จึงเป็นการบดบังความสัมพันธ์ของตัวแปรที่แตกต่างกันในแต่ละกลุ่ม เรียกปัญหาดังกล่าวว่า ปัญหาการปิดบังความผันแปรของสัมประสิทธิ์การถดถอยระหว่างกลุ่ม (heterogeneity of regression coefficients) ตัวสถิติตัวหนึ่งที่ใช้ในการวัดระดับของความสัมพันธ์ภายในกลุ่มดังกล่าวคือ สัมประสิทธิ์สหสัมพันธ์ภายในกลุ่ม (intraclass correlation: ICC) ซึ่งคำนวณจากสัดส่วนของความแปรปรวนระหว่างกลุ่ม (between group variance) กับความแปรปรวนรวม (total variance) โดยทั่วไปในงานวิจัยทางการศึกษามักพบว่าค่า ICC จะมีค่า

อยู่ระหว่าง 0.05 ถึง 0.3 (ศิริชัย กาญจนวาสี, 2550; สิวะโชติ ศรีสุทธิยากร, 2550; Terhorst, 2007; Mass & Hox, 2002)

วิธีการรวมข้อมูลตัวแปร (aggregation) เป็นวิธีที่ผู้วิจัยรวมข้อมูลของตัวแปรในระดับที่ต่ำกว่าขึ้นไปเป็นข้อมูลของตัวแปรในระดับที่สูงกว่า ซึ่งอาจใช้การเฉลี่ยหรือการรวมคะแนนของตัวแปรรายกลุ่ม เพื่อให้ตัวแปรที่สร้างใหม่มีความหมายเป็นตัวแปรของหน่วยตัวอย่างในระดับที่สูงกว่า วิธีการดังกล่าวมีข้อด้อยเพราะผู้วิจัยจะสูญเสียสารสนเทศของข้อมูลในแต่ละกลุ่มไปจำนวนมาก ซึ่งความสัมพันธ์ระหว่างตัวแปรของหน่วยตัวอย่างในระดับสูงที่สร้างขึ้นใหม่นั้นจะบดบังความสัมพันธ์ระหว่างตัวแปรของหน่วยตัวอย่างในระดับที่ต่ำกว่า ทำให้ความสัมพันธ์ของตัวแปรที่ทำการรวมข้อมูลเป็นตัวแปรระดับกลุ่มแล้วมีความแตกต่างจากความสัมพันธ์ระหว่างตัวแปรที่ยังไม่ได้กระทำการรวมตัวแปร ก่อให้เกิดความคลาดเคลื่อนในการสรุปผลการวิเคราะห์ เรียกความผิดพลาดดังกล่าวว่า ความเอนเอียงจากการสรุปผลข้ามระดับ (aggregation bias) (ศิริชัย กาญจนวาสี, 2550; Fox, 2001; Snijder and Bosker, 1999)

กล่าวโดยสรุปการวิเคราะห์ข้อมูลแบบพหุระดับในข้างต้นนั้นเป็นการวิเคราะห์โดยที่ผู้วิจัยไม่ได้คำนึงถึงโครงสร้างของข้อมูลตามธรรมชาติ ทำให้เกิดปัญหาตามมาหลายประการ ซึ่งไม่ว่าจะเลือกวิธีใดจะทำให้ผู้วิจัยสูญเสียสารสนเทศที่มี รวมทั้งทำให้เกิดความผิดพลาดในการสรุปผลที่ได้จากการวิเคราะห์ การวิเคราะห์โดยใช้การวิเคราะห์แบบระดับเดียวยังทำให้นักวิจัยไม่สามารถศึกษาปฏิสัมพันธ์ระหว่างตัวแปรข้ามระดับ (cross-level interaction) การตอบคำถามวิจัยจึงไม่มีความลึกซึ้ง ดังนั้นการวิเคราะห์ข้อมูลพหุระดับจึงควรมีการคำนึงถึงโครงสร้างตามธรรมชาติของข้อมูล ซึ่งจะทำให้ผลการวิเคราะห์มีความถูกต้องน่าเชื่อถือ อีกทั้งยังสามารถตอบคำถามวิจัยได้อย่างลึกซึ้งมากยิ่งขึ้น แนวทางการวิเคราะห์ข้อมูลที่เหมาะสมคือ การวิเคราะห์พหุระดับ (multi-level analysis) ซึ่งเป็นการวิเคราะห์ที่มีการคำนึงถึงความเป็นพหุระดับของข้อมูล ผลการวิเคราะห์ที่ได้นั้นมีความถูกต้องและน่าเชื่อถือ อีกทั้งยังไม่เสียสารสนเทศของข้อมูลในแต่ละระดับทำให้สามารถตอบคำถามวิจัยได้ครอบคลุมมากยิ่งขึ้น โดยจะนำเสนอรายละเอียดให้หัวข้อต่อไป (ศิริชัย กาญจนวาสี, 2550; Goldstein, 1995; Snijder and Bosker, 1999, Raudenbush and Bryk; 2002)

1.2 __โมเดลการวิเคราะห์พหุระดับ__ __(multi-level model)__

การวิเคราะห์พหุระดับเป็นเทคนิคการวิเคราะห์ข้อมูลโดยการพัฒนาโมเดลทางสถิติเพื่อให้มีการคำนึงถึงลักษณะโครงสร้างของข้อมูลที่เป็นระดับชั้น โมเดลที่พัฒนาขึ้นจะแบ่งแยกความผันแปรที่เกิดขึ้นจากข้อมูลในแต่ละระดับทำให้ผู้วิจัยสามารถวิเคราะห์ข้อมูลได้โดยที่ไม่เสียสารสนเทศของข้อมูล ซึ่งจะทำให้ผู้วิจัยสามารถศึกษาความสัมพันธ์ระหว่างตัวแปรทั้งในกลุ่มและระหว่างกลุ่ม นอกจากนี้ยังสามารถศึกษาปฏิสัมพันธ์ระหว่างตัวแปรข้ามระดับได้ ผลการวิเคราะห์ที่ได้จากวิธีการนี้จึงใช้สารสนเทศทั้งหมดที่ผู้วิจัยมีในการวิเคราะห์ ทำให้ผลการวิเคราะห์มีความน่าเชื่อถือ ครอบคลุม และตอบคำถามวิจัยได้อย่างลึกซึ้งกว่าการใช้สถิติวิเคราะห์แบบระดับเดียวดังที่กล่าวไว้ในข้างต้น โมเดลการวิเคราะห์ข้อมูลแบบพหุระดับโมเดลหนึ่งที่นิยมใช้กันในงานวิจัยทางสังคมศาสตร์อย่างแพร่หลายคือ โมเดลเชิงเส้นพหุระดับ (multi-level linear model) โมเดลดังกล่าวเป็นกรณีทั่วไปโมเดลหลายโมเดล เช่น โมเดลการวิเคราะห์ความแปรปรวนแบบอิทธิพลสุ่ม (random effect ANOVA) โมเดลการวิเคราะห์ความแปรปรวนร่วมแบบอิทธิพลสุ่ม (random effect ANCOVA) โมเดลสัมประสิทธิ์ความถดถอยแบบสุ่ม (random coefficients model) และโมเดลการวิเคราะห์พัฒนาการ (growth model) เป็นต้น (ศิริชัย กาญจนวาสี, 2550; Darandari, 2004; Raudenbush and Bryk; 2002; Snijder and Bosker, 1999; Goldstein, 1995)

ในกรณีที่โครงสร้างของข้อมูลเป็น 2 ระดับโมเดลเชิงเส้นพหุระดับจะมีรูปแบบทั่วไปดังนี้

โมเดลทั่วไประดับที่ 1 (level-1 general model)

$$\underline{Y}_j = X_j \underline{\beta}_j + \underline{\varepsilon}_j \tag{2.1}$$

โมเดลทั่วไประดับที่ 2 (level-2 general model)

$$\underline{\beta}_j = W_j \underline{\gamma} + \underline{u}_j \tag{2.2}$$

แทนสมการ (2) ลงใน (1) จะได้โมเดลจริงรวม (true combined model) ดังนี้

$$\underline{Y}_j = X_j W_j \underline{\gamma} + X_j \underline{u}_j + \underline{\varepsilon}_j \tag{2.3}$$

เมื่อ

$\underline{Y}_j$ เป็นเวกเตอร์ตัวแปรตามในระดับที่ 1 ของหน่วยตัวอย่างในกลุ่มที่ j ขนาด $n_j \times 1$

X_j เป็น design เมทริกซ์ของตัวแปรอิสระในระดับที่ 1 ของหน่วยตัวอย่างในกลุ่มที่ j ขนาด $n_j \times (p+1)$

$\underline{\beta}_j$ เป็นสัมประสิทธิ์ความถดถอยในระดับที่ 1 (level-1 regression coefficients) ของหน่วยตัวอย่างในกลุ่มที่ j ขนาด $(p+1) \times 1$

$\underline{\varepsilon}_j$ เป็นเวกเตอร์ของความคลาดเคลื่อนในโมเดลระดับที่ 1 ของหน่วยตัวอย่างในกลุ่มที่ j ขนาด $n_j \times 1$

W_j เป็น design เมทริกซ์ของตัวแปรอิสระในระดับที่ 2 ของหน่วยตัวอย่างในกลุ่มที่ j ขนาด $(p+1) \times (s+1)$

$\underline{\gamma}$ เป็นเวกเตอร์ของสัมประสิทธิ์ความถดถอยในระดับที่ 2 (level-2 regression coefficients) ขนาด $(s+1) \times 1$

$\underline{u}_j$ เป็นเวกเตอร์ของความคลาดเคลื่อนในโมเดลระดับที่ 2 ของหน่วยตัวอย่างในกลุ่มที่ j ขนาด $p \times 1$

n_j เป็นจำนวนหน่วยตัวอย่างในระดับที่ 1 ของกลุ่มที่ j

p เป็นจำนวนตัวแปรอิสระในระดับที่ 1

s เป็นจำนวนตัวแปรอิสระในระดับที่ 2

เมื่อเปรียบเทียบสมการ (2.3) กับโมเดลการวิเคราะห์ความถดถอยเชิงเส้นแบบปกติจะพบว่า หากผู้วิจัยเลือกใช้การวิเคราะห์ความถดถอยเชิงเส้นแบบปกติพจน์ของความคลาดเคลื่อนในสมการที่ (2.3) จะถูกรวมเข้าเป็นพจน์เดียว ดังนั้นการใช้โมเดลการวิเคราะห์ความถดถอยแบบปกติในการวิเคราะห์ความสัมพันธ์ของตัวแปรที่ข้อมูลมีโครงสร้างเป็นพหุระดับนั้นจะทำให้เกิดปัญหาการละเมิดข้อสมมติเบื้องต้นของการวิเคราะห์ความถดถอยแบบปกติได้แก่ ข้อสมมติของความเป็นอิสระของความคลาดเคลื่อนสุ่ม (independent) และข้อสมมติของความคงที่ของความแปรปรวนของความคลาดเคลื่อนสุ่ม (homoscedasticity) ซึ่งจะส่งผลให้ตัวประมาณกำลังสองน้อยสุด (least squares estimator) ไม่มีคุณสมบัติที่ดี กล่าวคือ ไม่เป็นตัวประมาณ BLUE (Best

Linear Unbiased Estimator) นอกจากนี้ยังส่งผลให้การประมาณค่าความคลาดเคลื่อนมาตรฐาน (standard error) ของตัวประมาณสัมประสิทธิ์ความถดถอยมีความผิดพลาด ดังนั้นการอนุมานเกี่ยวกับพารามิเตอร์ที่สนใจจึงมีความคลาดเคลื่อนไปด้วย ดังนั้นหากข้อมูลมีโครงสร้างของความเป็นพหุระดับการใช้โมเดลเชิงเส้นพหุระดับจึงมีความเหมาะสมมากกว่า

ข้อตกลงเบื้องต้นของโมเดลเชิงเส้นพหุระดับมีความคล้ายคลึงกับข้อตกลงเบื้องต้นของโมเดลเชิงเส้นแบบปกติทั่วไป แต่จะมีการกำหนดข้อตกลงเบื้องต้นเกี่ยวกับความคลาดเคลื่อนในระดับที่สองที่เพิ่มขึ้นมาในโมเดล รายละเอียดเป็นดังนี้

1. ความคลาดเคลื่อนสุ่มในระดับที่ 1 ($\underline{\varepsilon}_j$) เป็นอิสระและมีการแจกแจงเหมือนกัน (i.i.d) โดยมีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ย $\underline{0}$ และความแปรปรวน $\sigma_\varepsilon^2 \cdot I_{(n_j)}$
2. ค่าคลาดเคลื่อนสุ่มในระดับที่ 2 ($\underline{u}_j$) เป็นอิสระและมีการแจกแจงเหมือนกัน (i.i.d) มีการแจกแจงแบบปกติหลายตัวแปร ที่มีค่าเฉลี่ย $\underline{0}$ และเมทริกซ์ความแปรปรวนร่วม Σ โดยที่

$$\Sigma = \begin{bmatrix} \tau_{00} & \cdots & \tau_{0p} \\ \vdots & \ddots & \vdots \\ \tau_{p0} & \cdots & \tau_{pp} \end{bmatrix}$$

3. ค่าคลาดเคลื่อนสุ่มในทั้ง 2 ระดับเป็นอิสระต่อกัน
4. ตัวแปรอิสระในระดับที่ 1 (X_{ij}) เป็นอิสระกับค่าคลาดเคลื่อนทั้ง 2 ระดับ
5. ตัวแปรอิสระในทั้งสองระดับวัดได้โดยที่ไม่มีความคลาดเคลื่อนในการวัด

จากข้อสมมติเบื้องต้นและสมการ (2.3) ทำให้ได้ว่าค่าคาดคะเนแบบมีเงื่อนไข (conditional expectation) ของเวกเตอร์ $\underline{Y}_j$ เมื่อกำหนด X_j และ W_j มีค่าเท่ากับ

$$E(\underline{Y}_j \mid X_j, W_j) = X_j W_j \underline{\gamma} \tag{2.4}$$

และมีเมทริกซ์ความแปรปรวนร่วม (variance-covariance matrix) เป็น

$$V_j = Cov(\underline{Y}_j) = Cov(X_j W_j \underline{\gamma} + X_j \underline{u}_j + \underline{\varepsilon}_j) = X_j \cdot \Sigma \cdot X_j^T + \sigma_\varepsilon^2 \cdot I_{(n_j)} \tag{2.5}$$

<u>*การประมาณค่าและการอนุมานพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับโดยใช้สถิติแบบดั้งเดิม*</u>

การประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับโดยทั่วไปแล้วสามารถแบ่งออกได้เป็น 2 แนวทางได้แก่ การประมาณค่าพารามิเตอร์แบบดั้งเดิม (classical parameter estimation) และการประมาณค่าพารามิเตอร์แบบเบส์ (bayesian parameter estimation) โดยในหัวข้อนี้จะกล่าวถึงเฉพาะวิธีการประมาณค่าโดยใช้สถิติแบบดั้งเดิมเท่านั้น และจะกล่าวถึงการประมาณค่าโดยใช้สถิติแบบเบส์ในตอนที่ 2 ของรายงานส่วนนี้

วิธีประมาณค่าแบบภาวะความควรจะเป็นสูงสุด (maximum likelihood estimation)

วิธีการประมาณค่าพารามิเตอร์แบบดั้งเดิมมีหลักการคือคัดเลือกค่าประมาณพารามิเตอร์โดยพิจารณาจากฟังก์ชันวัตถุประสงค์ (objective function) วิธีการที่นิยมใช้คือวิธีภาวะความควรจะเป็นสูงสุด (maximum likelihood method: ML) กำหนดให้ $\underline{Y}$ เป็นเวกเตอร์ของค่าสังเกตตัวแปรตามในระดับที่ 1 และ $\underline{u}$ เป็นเวกเตอร์ของความคลาดเคลื่อนสุ่มในระดับที่ 2 และ $\underline{\omega} = \left(\underline{\gamma}, \Sigma, \sigma_{\varepsilon}^{2}\right)$ เป็นเวกเตอร์ของพารามิเตอร์ที่ต้องการประมาณในโมเดล ดังนั้นจะได้ว่าฟังก์ชันภาวะความควรจะเป็นของข้อมูลค่าสังเกต $\underline{Y}$ เมื่อกำหนดพารามิเตอร์ $\underline{\omega}$ คือ

$$L(\underline{Y} \mid \underline{\omega}) = \int f(\underline{Y} \mid \underline{u}, \underline{\omega}) \cdot p(\underline{u} \mid \underline{\omega}) d\underline{u} \tag{2.6}$$

โดยที่ $f(\underline{Y} \mid \underline{u}, \underline{\omega})$ คือฟังก์ชันความน่าจะเป็นของค่าสังเกตในระดับที่ 1 เมื่อกำหนดพารามิเตอร์ $\underline{\omega}$ และพารามิเตอร์ความคลาดเคลื่อน $\underline{u}$ วิธีการภาวะความควรจะเป็นสูงสุดนี้จะใช้ฟังก์ชันภาวะความควรจะเป็น(likelihood function) เป็นฟังก์ชันวัตถุประสงค์ โดยจะเลือกค่าประมาณพารามิเตอร์ $\underline{\omega}$ ที่ทำให้ฟังก์ชันภาวะความควรจะเป็นมีค่าสูงสุด (maximize likelihood)

ในกรณีทั่วไปจะไม่สามารถพิสูจน์สูตรของการประมาณค่าให้อยู่ในรูปปิดได้ ดังนั้นการเลือกค่าประมาณพารามิเตอร์เพื่อทำให้ฟังก์ชันความควรจะเป็นมีค่าสูงสุดนั้นจึงอาศัยอัลกอริทึมแบบทวนซ้ำ (iterative algorithm) เพื่อหาค่าประมาณที่ดีที่สุด อัลกอริทึมที่ใช้ในการประมาณค่าพารามิเตอร์นั้นได้มีผู้พัฒนาไว้หลายอัลกอริทึมดังนี้ วิธี iterative generalized least squares (IGLS) (Goldstein, 1986) วิธี restricted iterative generalized least squares (RIGLS) (Goldstein, 1989) วิธี expectation-maximization (EM) (Dempster, rubin, & Tsutakawa, 1981; Dempster, Laird, & Rubin, 1977) วิธี fisher scoring (Longford, 1987) อัลกอริทึมจะประมาณค่าพารามิเตอร์แบบทวนซ้ำ (iterative process) โดยประมาณแยกทีละส่วนซึ่งจะอาศัยความรู้เกี่ยวกับพารามิเตอร์ในอีกส่วนหนึ่งเพื่อประมาณค่าพารามิเตอร์อีกส่วนหนึ่ง กระบวนการประมาณดังกล่าวจะหยุดเมื่อค่าประมาณที่ได้จากอัลกอริทึมในข้างต้นลู่เข้าสู่ค่าประมาณที่เหมาะสมแล้ว ในกรณีที่ข้อสมมติเบื้องต้นเป็นจริง อัลกอริทึมต่างๆจะได้ค่าประมาณที่ไม่แตกต่างกัน ความแตกต่างจะอยู่ที่ความยากง่ายของการคำนวณและความรวดเร็วในการลู่เข้าของค่าประมาณในแต่ละอัลกอริทึม (สิวะโชติ ศรีสุทธิยากร, 2550; Mass & Hox, 2002) ในส่วนต่อไปจะพิจารณารายละเอียดของอัลกอริทึม IGLS ซึ่งเป็นอัลกอริทึมหนึ่งที่ใช้ในการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับตามแนวคิดของวิธีภาวะความควรจะเป็นสูงสุดดังต่อไปนี้

วิธีการประมาณ IGLS เป็นวิธีการที่มีหลักการคือประมาณค่าไปกลับระหว่างค่าประมาณของพารามิเตอร์ที่เป็นอิทธิพลคงที่ และค่าประมาณของพารามิเตอร์ที่เป็นอิทธิพลสุ่ม โดยจะมีการเปลี่ยนการถ่วงน้ำหนัก (reweighted) ด้วยค่าส่วนประกอบความแปรปรวน (V_j) ในทุกรอบการประมาณ จนกระทั่งค่าประมาณในทั้ง 2 ส่วนลู่เข้าสู่ค่าที่เหมาะสมจึงสิ้นสุดการประมาณ (Terhorst, 2007; Browne, 1998; Goldstein, 1995, 1986) รายละเอียดของวิธี IGLS เป็นดังนี้

สมมติว่าทราบค่าประมาณของอิทธิพลสุ่มหรือส่วนประกอบความแปรปรวน จะได้ว่าเราสามารถใช้ตัวประมาณกำลังสองน้อยสุดทั่วไปประมาณอิทธิพลคงที่ได้ดังนี้

$$\underline{\hat{\gamma}} = (\sum_j X_j^T V_j^{-1} X_j)^{-1} \sum_j X_j^T V_j^{-1} \underline{Y}_j \qquad (2.7)$$

และจะได้ว่าเมทริกซ์ความแปรปรวนร่วมของ $\underline{\hat{\gamma}}$ คือ

$$Cov(\underline{\hat{\gamma}}) = (\sum_j X_j^T V_j^{-1} X_j)^{-1} \sum_j X_j^T V_j^{-1} Cov(\underline{Y}_j) V_j^{-1} X_j (\sum_j X_j^T V_j^{-1} X_j)^{-1}$$

$$= (\sum_j X_j^T V_j^{-1} X_j)^{-1} \qquad (2.8)$$

จากค่าประมาณอิทธิพลคงที่ได้ใน (2.7) กำหนดให้

$$E_j^* = \underline{E}_j \cdot \underline{E}_j^T \qquad \text{เมื่อ} \quad \underline{E}_j = \underline{Y}_j - X_j \underline{\gamma} \qquad (2.9)$$

และ $E_j^{**} = vec(E_j^*)$ (2.10)

ซึ่งเป็นการเรียงเมทริกซ์ E_j^* ให้อยู่ในรูปเวกเตอร์โดยเรียงคอลัมม์ทุกคอลัมม์ในอยู่ในคอลัมม์แรกเพียงคอลัมม์เดียว

เราสามารถเขียนโมเดลเชิงเส้นของ E_j^{**} ได้ดังนี้

$$E_j^{**} = Z_j \underline{\theta} + \underline{R}_j \quad (2.11)$$

เมื่อ Z_j แทน design matrix ของพารามิเตอร์ผลกระทบสุ่ม $\underline{\theta}$

$\underline{\theta}$ แทนพารามิเตอร์ผลกระทบสุ่ม หรือส่วนประกอบความแปรปรวนทั้งในระดับที่ 1 และระดับที่ 2

ดังนั้นจากโมเดลในสมการที่ (2.11) จะสามารถประมาณส่วนประกอบความแปรปรวนได้โดยวิธีสองน้อยสุดนัยทั่วไปเช่นเดียวกัน ดังนี้

$$\hat{\underline{\theta}} = (\sum_j Z_j^T V_j^{*-1} Z_j)^{-1} \sum_j Z_j^T V_j^{*-1} E_j^{**} \quad \text{โดยที่ } V_j^* = V_j \otimes V_j \quad (2.12)$$

เมื่อ $\otimes$ คือ kronecker product หรือ direct product

และจะได้ว่าเมทริกซ์ความแปรปรวนร่วมของ $\hat{\underline{\theta}}$ คือ

$$Cov(\hat{\theta}) = 2(\sum_j Z_j^T V_j^{*-1} Z_j)^{-1} \quad (2.13)$$

วิธี IGLS เป็นอัลกอริทึมที่จะประมาณค่าทวนซ้ำไปมาระหว่างสมการที่ (2.7) และ (2.12) ตัวประมาณที่ได้มาในแต่ละรอบ จะลู่เข้าสู่ค่าที่เหมาะสม โดยในการเลือกค่าเริ่มต้นของการประมาณ (initial) ผู้วิเคราะห์อาจเลือกค่าเริ่มต้นโดยใช้ตัวประมาณกำลังสองน้อยสุดแบบสามัญ (ordinary least square estimator) โดยในรอบแรกที่ทำการประมาณจะกำหนดให้ $V_j = \sigma_\varepsilon^2 I_{n_j}$

วิธีการนี้ไม่ได้มีการคำนึงความผันแปรจากการสุ่มตัวอย่าง (sampling variation) กล่าวคือไม่ได้มีการพิจารณาความไม่แน่นอนของค่า $\underline{\gamma}$ ในการประมาณส่วนประกอบความแปรปรวน ดังนั้นในกรณีที่ขนาดตัวอย่างมีขนาดเล็กจะทำให้ค่าประมาณพารามิเตอร์ที่ได้จากวิธีการภาวะความควรจะเป็นสูงสุดนี้ในส่วนของค่าประมาณพารามิเตอร์ส่วนประกอบความ

แปรปรวน (variance component) เป็นตัวประมาณที่มีความเอนเอียง (biased) อย่างไรก็ตาม หากข้อสมมติของโมเดลเป็นจริงตัวประมาณที่ได้จากวิธีการนี้จะมีคุณสมบัติที่ดีได้แก่ คุณสมบัติคงเส้นคงวา (consistency) กล่าวคือตัวประมาณดังกล่าวจะมีความน่าจะเป็นที่จะเข้าใกล้ค่าพารามิเตอร์ที่แท้จริงมากขึ้น เมื่อขนาดตัวอย่างเพิ่มขึ้น และคุณสมบัติเป็นตัวประมาณที่มีประสิทธิภาพเมื่อขนาดตัวอย่างใหญ่ (asymptotically efficient) กล่าวคือ ตัวประมาณดังกล่าวจะเป็นตัวประมาณที่ไม่เอนเอียง (unbiased) และมีความแปรปรวนต่ำสุด (minimum variance) เมื่อขนาดตัวอย่างใหญ่เพียงพอ นอกจากนี้ตัวประมาณความควรจะเป็นสูงสุดยังมีคุณสมบัติไม่แปรเปลี่ยน (invariance) อีกด้วย (สิวะโชติ ศรีสุทธิยากร, 2550; Raudenbush & Bryk, 2002)

วิธีประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted (restricted maximum likelihood estimation)

เพื่อแก้ไขปัญหาดังกล่าวจึงได้มีการพัฒนาวิธีภาวะความควรจะเป็นสูงสุดที่ถูกจำกัด (restricted maximum likelihood: REML) สำหรับค่าประมาณที่เป็นไปได้ทั้งหมดของ $\underline{\gamma}$ (เขียนแทนด้วย $\underline{\gamma}_m$) จะได้ว่าฟังก์ชันภาวะความควรจะเป็นของพารามิเตอร์ใน Σ และ σ_ε^2 เมื่อกำหนด $\underline{\gamma}_m$ และ $\underline{Y}$ เขียนแทนด้วย $L_m(\Sigma, \sigma_\varepsilon^2 \mid \underline{\gamma}_m, \underline{y})$ ซึ่งฟังก์ชันภาวะความควรจะเป็นแบบ restricted (restricted likelihood function) คือฟังก์ชันการแจกแจงความน่าจะเป็นส่วนริม (marginal probability distribution) ดังนี้

$$\int L(\underline{\gamma}, \Sigma, \sigma_\varepsilon^2 \,|, \underline{y}) d\underline{\gamma} = L(\Sigma, \sigma_\varepsilon^2 \mid \underline{y}) \qquad (2.14)$$

ด้วยหลักการเช่นเดียวกับวิธีภาวะความควรจะเป็นสูงสุด การประมาณค่าพารามิเตอร์ส่วนประกอบความแปรปรวนจะประมาณโดยการเลือกค่าประมาณพารามิเตอร์ส่วนประกอบความแปรปรวน Σ และ σ_ε^2 ซึ่งทำให้ฟังก์ชันภาวะความควรจะเป็นแบบ restricted ร่วม (restricted joint likelihood function) มีค่าสูงสุด

เช่นเดียวกับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุด วิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted สามารถประมาณค่าโดยอาศัยอัลกอริทึมได้หลายตัว หากใช้อัลกอริทึมตามแนวคิดของวิธี IGLS จะสามารถปรับปรุงอัลกอริทึม IGLS ให้ประมาณ

ค่าตามแนวทางของวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted เรียกว่าอัลกอริทึม RIGLS โดยอัลกอริทึมในการประมาณยังเหมือนกับวิธี iterative generalized least squares แต่ได้มีการปรับสูตรในการประมาณดังนี้ (Browne, 1998; Goldstein, 1989)

เนื่องจาก $\underline{\hat{\gamma}}$ ซึ่งเป็นตัวประมาณกำลังสองน้อยสุดนัยทั่วไปของ $\underline{\gamma}$ และถ้าทราบค่าส่วนประกอบความแปรปรวน V ดังนั้นจะได้

$$\underline{E}_R = \underline{Y} - X\underline{\hat{\gamma}} \tag{2.15}$$

และจะได้

$$\underline{E}_R \underline{E}_R^T = (\underline{Y} - X\underline{\hat{\gamma}})(\underline{Y} - X\underline{\hat{\gamma}})^T \tag{2.16}$$

โดยที่

$$E(\underline{E}_R \underline{E}_R^T) = V - X(X^T V^{-1} X)^{-1} X^T \tag{2.17}$$

เรียกค่า $X(X^T V^{-1} X)^{-1} X^T$ ว่าเป็นค่าปรับความลำเอียง (bias correction)
โดยในการประมาณค่าในแต่ละรอบจะใช้

$$(Y - X\underline{\hat{\gamma}})(Y - X\underline{\hat{\gamma}})^T + X(X^T \hat{V}^{-1} X)^{-1} X^T \tag{2.18}$$

ในการประมาณค่าส่วนประกอบความแปรปรวน V ซึ่งพัฒนามาจากค่าประมาณของส่วนประกอบความแปรปรวนในรอบที่แล้ว $\hat{V}$

จากวิธีการประมาณทั้งสองจะทำให้ได้ค่าประมาณของพารามิเตอร์อิทธิพลคงที่ และอิทธิพลสุ่ม ผู้วิจัยสามารถใช้ค่าประมาณที่ได้ในส่วนนี้เพื่อประมาณค่าพารามิเตอร์สัมประสิทธิ์ความถดถอยในระดับที่ 1 (level-1 random coefficients) ตัวประมาณที่นิยมใช้เพื่อประมาณพารามิเตอร์ดังกล่าวในกรณีนี้คือ ตัวประมาณเบส์เชิงประจักษ์ (empirical bayes) ซึ่งมีรายละเอียดดังต่อไปนี้

ในการประมาณค่าพารามิเตอร์สัมประสิทธิ์ความถดถอยสุ่มในระดับที่ 1 ($\underline{\beta}_j$) อาจสามารถประมาณได้ 2 แนวทางได้แก่ การใช้ตัวประมาณกำลังสองน้อยสุดทั่วไป (ordinary least squares) และการใช้ตัวประมาณค่าเฉลี่ยของสัมประสิทธิ์ความถอยที่ได้จากวิธีภาวะความควรจะ

เป็นสูงสุดหรือวิธีภาวะความควรจะเป็นสูงสุดที่ถูกจำกัดดังที่กล่าวไว้ในข้างต้น ในกรณีที่ใช้การประมาณค่าพารามิเตอร์ $\underline{\beta}_j$ ด้วยวิธีกำลังสองน้อยสุดทั่วไปสามารถคำนวณได้จาก

$$\hat{\underline{\beta}}_j = \left(X_j^T X_j\right)^{-1}\left(X_j^T Y_j\right) \quad (2.19)$$

ส่วนในกรณีที่ใช้ตัวประมาณอิทธิพลคงที่จะสามารถประมาณค่าพารามิเตอร์ $\underline{\beta}_j$ ได้จาก

$$\underline{\beta}_j^* = W_j \hat{\gamma} \quad (2.20)$$

จากสมการ (2.19) และ (2.20) จะได้ว่าตัวประมาณเอมไพริคัลเบส์ของสัมประสิทธิ์ความถดถอยแบบสุ่มในระดับที่ 1 คือ

$$\underline{\beta}_j^{EB} = \Delta_j \hat{\underline{\beta}}_j + (I - \Delta_j) W_j \hat{\gamma} \quad (2.21)$$

เมื่อ $\Delta_j = \Sigma \cdot \left(\Sigma + V_j\right)^{-1}$

<u>*การอนุมานเกี่ยวกับพารามิเตอร์ในโมเดล*</u>

การอนุมานเกี่ยวกับพารามิเตอร์ตามแนวทางสถิติแบบดั้งเดิมสามารถกระทำได้ 2 วิธีการ ได้แก่ การประมาณค่าแบบช่วง (interval estimation) และการทดสอบสมมติฐาน (hypothesis testing) รายละเอียดเป็นดังต่อไปนี้ (ศิริชัย กาญจนวาสี, 2550; Darandari, 2004; Raudenbush and Bryk; 2002; Snijder and Bosker, 1999; Goldstein, 1995)

การประมาณค่าแบบช่วง

วิธีการประมาณค่าแบบช่วงที่นิยมใช้กันทั่วไปคือการหาช่วงความเชื่อมั่น (confidence interval) รายละเอียดมีดังต่อไปนี้

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของพารามิเตอร์อิทธิพลคงที่สามารถคำนวณได้จากสูตรดังสมการที่ (2.22) ดังนี้

$$\hat{\gamma}_{qq} \pm t_{\alpha/2} (V_{qq})^{1/2} \quad (2.22)$$

โดยที่ V_{qq} คือสมาชิกในแนวทแยงมุมของเมทริกซ์ความแปรปรวนร่วม $Cov(\underline{\hat{\gamma}})$ ในสมการที่ (2.8)

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของพารามิเตอร์สัมประสิทธิ์ความถดถอยสุ่มสามารถคำนวณได้จากสูตรดังสมการที่ (2.23) ดังนี้

$$\beta_j^{EB} \pm t_{\alpha/2}(V_{jj}^{EB})^{1/2} \quad (2.23)$$

โดยที่ V_{jj}^{EB} คือสมาชิกในแนวทแยงมุมของเมทริกซ์ความแปรปรวนร่วม V_j^{EB} ในสมการที่ (2.24) ดังนี้

$$V_j^{EB} = \left(V_j^{-1} + \Sigma^{-1}\right)^{-1} + \left(I - \Delta_j\right) Cov\left(W_j \hat{\underline{\gamma}}\right) \cdot \left(I - \Delta_j\right)^T \quad (2.24)$$

การทดสอบสมมติฐาน

ในโมเดลการวิเคราะห์แบบ 2 ระดับตามสมการที่ (2.1) ถึง (2.3) จะพบว่าสามารถทดสอบสมมติฐานของพารามิเตอร์ได้ทั้งหมด 3 ส่วนด้วยกันได้แก่ การทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์อิทธิพลคงที่ พารามิเตอร์สัมประสิทธิ์ความถดถอยสุ่มในระดับที่ 1 และพารามิเตอร์ส่วนประกอบความแปรปรวน นอกจากนี้การทดสอบสมมติฐานบังสามารถแบ่งออกได้เป็น 2 ประเภทได้แก่ การทดสอบสมมติฐานสำหรับพารามิเตอร์ตัวเดียว (single-parameter) และการทดสอบสมมติฐานสำหรับพารามิเตอร์หลายตัว (multiparameter) รายละเอียดเป็นดังต่อไปนี้ (Raudenbush and Bryk, 2002)

<u>การทดสอบสมมติฐานสำหรับพารามิเตอร์อิทธิพลคงที่</u>

กรณีที่ต้องการทดสอบสมมติฐานสำหรับพารามิเตอร์อิทธิพลคงที่ (hypothesis tests for fixed effects) แบบพารามิเตอร์เดียว สามารถตั้งสมมติฐานการทดสอบได้ดังนี้

$$H_0 : \gamma_{qs} = 0 \quad \text{vs} \quad H_1 : \gamma_{qs} \neq 0 \quad (2.25)$$

ตัวสถิติที่ใช้ในการทดสอบคือตัวสถิติที (t-test) ตามสมการดังต่อไปนี้

$$t = \frac{\hat{\gamma}_{qs}}{\left(V_{qq}\right)^{1/2}}; \qquad df = J - S_q - 1 \quad (2.26)$$

กรณีที่ต้องการทดสอบสมมติฐานสำหรับพารามิเตอร์อิทธิพลคงที่แบบหลายพารามิเตอร์ สามารถใช้การทดสอบสมมติฐานเชิงเส้นทั่วไป (general linear test) ซึ่งตั้งสมมติฐานการทดสอบในรูปของเมทริกซ์ได้ดังนี้

$$H_0 : C^T \underline{\gamma} = 0 \quad \text{vs} \quad H_1 : C^T \underline{\gamma} \neq 0 \qquad (2.27)$$

โดยที่ C คือ design matrix ขนาด $(Q \times \sum_q S_q)$

ตัวสถิติที่ใช้ในการทดสอบคือตัวสถิติไคแสควร์ (chi-square test) ตามสมการดังต่อไปนี้

$$\chi^2 = \underline{\gamma}^T C \cdot V_c^{-1} \cdot C^T \underline{\gamma} \qquad (2.28)$$

โดยที่ $V_C = Cov(C^T \hat{\underline{\gamma}}) = C^T Cov(\hat{\underline{\gamma}}) C$

ตัวสถิติในสมการที่ (2.28) จะมีองศาความเป็นอิสระ (df) เท่ากับจำนวนของคอนทราส (contrasts) ที่ต้องการทดสอบหรือเท่ากับจำนวนแถวของเมทริกซ์ C^T ข้อดีของการทดสอบสมมติฐานแบบหลายพารามิเตอร์คือเป็นการประกันความเสี่ยงประเภทที่ 1 (Type I error) ที่เกิดจากการทดสอบสมมติฐานพารามิเตอร์เดียวหลายการทดสอบ กลยุทธ์หนึ่งที่อาจใช้ในการวิเคราะห์ข้อมูลคือจะทำการทดสอบสมมติฐานแบบพารามิเตอร์เดียวก็ต่อเมื่อการทดสอบสมมติฐานแบบหลายพารามิเตอร์มีนัยสำคัญเท่านั้น

การทดสอบสมมติฐานสำหรับพารามิเตอร์อิทธิพลคงที่แบบหลายพารามิเตอร์ยังสามารถกระทำได้อีกวิธีการหนึ่งคือการทดสอบด้วยการใช้ตัวทดสอบอัตราส่วนภาวะความควรจะเป็น (likelihood ratio test) ข้อจำกัดประการหนึ่งของการทดสอบนี้คือจะสามารถทดสอบได้เฉพาะโมเดลที่ประมาณค่าพารามิเตอร์ด้วยวิธีภาวะความควรจะเป็นสูงสุดเท่านั้น หลักการของวิธีการดังกล่าวเป็นการเปรียบเทียบโมเดลคู่แข่ง (candidate models) จำนวน 2 โมเดล โมเดลแรกจะเรียกว่า โมเดลตามสมมติฐานหลัก (null model) และโมเดลที่สองจะเรียกว่า โมเดลทางเลือก (alternative model) โดยทั่วไปโมเดลตามสมมติฐานหลักจะเป็นโมเดลย่อย (sub model) ของโมเดลทางเลือก สำหรับการทดสอบสมมติฐานจำเป็นที่จะต้องคำนวณค่า deviance ของแต่ละโมเดล ค่า deviance สามารถคำนวณได้จากค่า -2log likelihood ดังนั้นความหมายของ deviance จึงหมายถึงความเหมาะสมของโมเดลกับข้อมูลเชิงประจักษ์ โมเดลที่มีค่า deviance สูงมีความสอดคล้องกับข้อมูลเชิงประจักษ์น้อยกว่าโมเดลที่มีค่า deviance ต่ำ กำหนดให้ D_0 คือค่า

deviance ของโมเดลว่าง และ D_1 คือ deviance ของโมเดลทางเลือก ตัวสถิติทดสอบเพื่อเปรียบเทียบโมเดลทั้งสองคำนวณตามสมการที่ (2.29) ดังนี้

$$H = D_0 - D_1 \tag{2.29}$$

ตัวสถิติ H จะมีการแจกแจงลู่เข้าสู่การแจกแจงแบบไคสแควร์ที่มีองศาความเป็นอิสระเท่ากับผลต่างของจำนวนพารามิเตอร์ของทั้งสองโมเดลเมื่อขนาดตัวอย่างใหญ่เพียงพอ เกณฑ์ในการตัดสินใจจะพิจารณาจากขนาดของค่า H ถ้าค่าของผลต่าง H มีค่ามากจะเป็นการแสดงหลักฐานว่าโมเดลทางเลือกสามารถอธิบายข้อมูลเชิงประจักษ์หรือมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากกว่าโมเดลว่างนั่นเอง

การทดสอบด้วยอัตราส่วนภาวะความควรจะเป็นจะได้ผลการทดสอบที่ใกล้เคียงกับการทดสอบเชิงเส้นทั่วไป อย่างไรก็ตามการทดสอบเชิงเส้นทั่วไปมีข้อดีมากกว่าการใช้การสอบอัตราส่วนภาวะความควรจะเป็นดังนี้ 1) การใช้การทดสอบเชิงเส้นทั่วไปสามารถกระทำได้สะดวกกว่าการทดสอบอัตราส่วนภาวะความควรจะเป็น เนื่องจากการทดสอบเชิงเส้นทั่วไปนั้นสามารถทดสอบสมมติฐานที่ต้องการได้โดยอาศัยค่าสถิติจากโมเดลเดียวเท่านั้น แต่การทดสอบอัตราส่วนภาวะความควรจะเป็นนั้นจำเป็นที่จะต้องประมาณพารามิเตอร์ในโมเดลใหม่ทุกครั้งเมื่อสมมติฐานการทดสอบเปลี่ยนไป 2) การทดสอบอัตราส่วนภาวะความควรจะเป็นจะไม่สามารถใช้ได้ในกรณีที่ผู้วิจัยใช้ตัวประมาณ REML เพื่อประมาณค่าพารามิเตอร์ในโมเดล 3) การทดสอบเชิงเส้นทั่วไปยอมให้ผู้วิจัยสามารถทดสอบสมมติฐานในคำถามที่เกี่ยวกับความแตกต่างระหว่างค่าของพารามิเตอร์ได้กล่าวคือสามารถทดสอบคอนทราสเชิงเส้น (linear contrast) ระหว่างพารามิเตอร์ในโมเดลได้ อย่างไรก็ตามการทดสอบอัตราส่วนภาวะความควรจะเป็นจะมีบทบาทในกรณีที่ต้องการทดสอบสมมติฐานสำหรับพารามิเตอร์ส่วนประกอบความแปรปรวนหลายพารามิเตอร์ซึ่งจะกล่าวในส่วนต่อไป

<u>การทดสอบสมมติฐานสำหรับพารามิเตอร์สัมประสิทธิ์ความถดถอยแบบสุ่มในระดับที่ 1</u>

การทดสอบสมมติฐานแบบพารามิเตอร์เดียวจะมีสมมติฐานการทดสอบดังนี้

$$H_0 : \beta_{qj} = 0 \quad \text{vs} \quad H_1 : \beta_{qj} \neq 0 \tag{2.30}$$

การเลือกตัวสถิติทดสอบสามารถกระทำได้ในสองกรณีคือการเลือกตัวสถิติทดสอบในกรณีที่ผู้วิจัยเลือกใช้ตัวประมาณกำลังสองน้อยสุด และในกรณีที่ผู้วิจัยเลือกใช้ตัวประมาณเอมไพริคัลเบส์ กรณีที่ผู้วิจัยใช้ตัวประมาณกำลังสองน้อยสุดเพื่อประมาณค่าสัมประสิทธิ์ความถดถอยแบบสุ่ม ตัวสถิติที่ใช้ในการทดสอบสมมติฐานตามสมการที่ (2.30) คือตัวสถิติที (t-test) เหมือนการทดสอบในโมเดลถดถอยเชิงเส้นแบบปกติทั่วไป แต่ในกรณีที่ผู้วิจัยใช้ตัวประมาณเอมไพริคัลเบส์ ตัวสถิติที่ใช้ในการทดสอบสมมติฐานจะมีสมการดังต่อไปนี้

$$Z = \frac{\hat{\beta}_{qj}}{\left(V_{qqj}^{EB}\right)^{1/2}} \tag{2.31}$$

การทดสอบสมมติฐานหลายพารามิเตอร์สามารถกระทำในทำนองเดียวกันกับการทดสอบสมมติฐานหลายพารามิเตอร์สำหรับพารามิเตอร์อิทธิพลคงที่ สมมติฐานการทดสอบโดยใช้การทดสอบเชิงเส้นทั่วไปเป็นดังนี้

$$H_0 : C^T \underline{\beta} = 0 \quad \text{vs} \quad H_1 : C^T \underline{\beta} \neq 0 \tag{2.32}$$

ในกรณีที่เลือกใช้ตัวประมาณกำลังสองน้อยสุดสามารถคำนวณตัวสถิติทดสอบได้จากสมการที่ (2.49) ดังนี้

$$\chi_{ols}^2 = \underline{\hat{\beta}}^T C \cdot \left(C \cdot V_{olsj} \cdot C^T\right)^{-1} \cdot C \underline{\hat{\beta}} \tag{2.33}$$

โดยที่ $\hat{V}_{olsj} = \hat{\sigma}^2 \left(X_j^T X_j\right)^{-1}$

ในกรณีที่เลือกใช้ตัวประมาณเอมไพริคัลเบส์ ตัวสถิติที่ใช้ในการทดสอบสมมติฐานสามารถคำนวณได้จากสมการที่ (2.34) ดังนี้

$$\chi_{EB}^2 = \underline{\hat{\beta}}_{EB}^T C \cdot \left(C \cdot V_{EB} \cdot C^T\right)^{-1} \cdot C \underline{\hat{\beta}}_{EB} \tag{2.34}$$

<u>การทดสอบสมมติฐานของส่วนประกอบความแปรปรวน</u>

การทดสอบสมมติฐานพารามิเตอร์เดียวมีสมมติฐานการทดสอบดังนี้

$$H_0 : \tau_{qq} = 0 \quad \text{vs} \quad H_1 : \tau_{qq} \neq 0 \tag{2.35}$$

โดยที่ τ_{qq} เป็นสมาชิกในแนวทแยงมุมของเมทริกซ์ Σ ตัวสถิติทดสอบที่ใช้ในการทดสอบสามารถคำนวณได้ตามสมการที่ (2.36) ดังนี้

$$\chi^2_{ols} = \sum_j \left(\hat{\beta}_{qj} - \hat{\gamma}_{q0} - \sum_{s=1}^{S_q} \hat{\gamma}_{qs} W_{sj} \right)^2 \Big/ \hat{V}_{olsqqj} \qquad (2.36)$$

โดยที่ $\hat{V}_{olsqqj}$ สมาชิกในแนวทแยงมุมของเมทริกซ์ $\hat{V}_{olsj}$ และ $\hat{\beta}_{qj}$ คือตัวประมาณกำลังสองน้อยสุด

การทดสอบสมมติฐานหลายพารามิเตอร์สามารถใช้การทดสอบอัตราส่วนภาวะความควรจะเป็นหรือการทดสอบ deviance เช่นเดียวกับกรณีของอิทธิพลคงที่ อย่างไรก็ตามในกรณีนี้องศาความเป็นอิสระของตัวสถิติทดสอบจะมีค่าเท่ากับผลต่างของจำนวนพารามิเตอร์ส่วนประกอบความแปรปรวนของโมเดลว่างและโมเดลทางเลือก ข้อควรระวังของการใช้ตัวทดสอบอัตราส่วนภาวะความควรจะเป็นในการทดสอบพารามิเตอร์ส่วนประกอบความแปรปรวนหลายพารามิเตอร์ คือ โมเดลที่ใช้ในการเปรียบเทียบเพื่อทดสอบสมมติฐานนั้นจะต้องมีพารามิเตอร์ในส่วนของพารามิเตอร์อิทธิพลคงที่เหมือนกันเท่านั้น

ในกรณีที่ข้อมูลที่ใช้ในการวิเคราะห์นั้นไม่เป็นไปตามข้อตกลงเบื้องต้นของโมเดล เช่นการแจกแจงของความคลาดเคลื่อนสุ่มในแต่ละระดับหรือบางระดับอาจไม่ได้มีการแจกแจงแบบปกติ หรือความแปรปรวนของความคลาดเคลื่อนสุ่มในแต่ละระดับหรือบางระดับอาจไม่ได้มีค่าคงที่ ซึ่งจะส่งผลกระทบต่อการประมาณค่าคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์ทำให้การอนุมานเกี่ยวกับพารามิเตอร์มีความผิดพลาด วิธีการหนึ่งที่ช่วยแก้ไขปัญหาคือการปรับสูตรการประมาณค่าคลาดเคลื่อนมาตรฐานเพื่อให้มีความแกร่งต่อการละเมิดข้อตกลงเบื้องต้นของโมเดล เรียกว่าตัวประมาณค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่ง (robust standard error) อาจเรียกว่า ตัวประมาณค่าคลาดเคลื่อนมาตรฐานของฮูเบอร์ไวท์ (Huber/White Estimator) (Raudenbush and Bryk, 2002; White, 1980) หลักการของการประมาณค่าคลาดเคลื่อนมาตรฐานดังกล่าวจะมีการปรับสูตรในการประมาณ $Cov(\underline{\hat{\gamma}})$ โดยใช้เมทริกซ์ทแยงมุม (Block Diagonal) ของ $E^* = \underline{E} \cdot \underline{E}^T$ ซึ่งเป็นตัวประมาณที่คงเส้นคงวาของ V แทน $Cov(\underline{\hat{\gamma}})$ และสำหรับในกรณีของพารามิเตอร์ของอิทธิพลสุ่ม หรือส่วนประกอบความแปรปรวน กระทำได้ในทำนองเดียวกันกับ ค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่งยังสามารถใช้ในการตรวจสอบความ

ถูกต้องของโมเดลได้อีกด้วย การที่ค่าคลาดเคลื่อนมาตรฐานที่ได้จากโมเดล มีความแตกต่างจากค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่งมาก อาจเป็นการบ่งบอกว่าผู้วิจัยเลือกใช้โมเดลผิดพลาด (Goldstein; 1995, Raudenbush and Bryk; 2002) ในกรณีที่ข้อตกลงเบื้องต้นของโมเดลเป็นจริงหรือใกล้เคียง ค่าประมาณค่าคลาดเคลื่อนมาตรฐานแบบปกติจะเป็นตัวประมาณที่มีความแปรปรวนต่ำสุด จึงไม่ควรใช้ค่าประมาณค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่งในการอนุมานหากข้อตกลงเบื้องต้นของโมเดลเป็นจริงหรือใกล้เคียง เนื่องจากจะทำให้ตัวสถิติทดสอบนั้นมีความไวต่ำส่งทำให้โอกาสในการปฏิเสธสมมติฐานหลักมีน้อยลง นอกจากนี้ช่วงความเชื่อมั่นที่สร้างขึ้นจะเป็นช่วงความเชื่อมั่นที่มีความยาวมากกว่าปกติส่งผลต่อการนำช่วงความเชื่อมั่นดังกล่าวไปใช้ประโยชน์ (ศิวะโชติ ศรีสุทธิยากร, 2550; Maas and Hox, 2001)

ตอนที่ 2 แนวคิดทฤษฎีเกี่ยวกับสถิติแบบเบส์

สถิติศาสตร์นั้นสามารถแบ่งออกได้เป็น 2 ส่วนได้แก่ สถิติพรรณนา (descriptive statistics) และสถิติอนุมาน (inferential statistics) สถิติพรรณนานั้นเป็นส่วนที่เกี่ยวกับ การวางแผนและออกแบบการดำเนินการ การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล รวมทั้งการประมวลผลและการคำนวณเบื้องต้น การแจกแจงความถี่ การใช้แผนภูมิและกราฟ การวัดแนวโน้มสู่ส่วนกลาง การวัดตำแหน่งของข้อมูล และการวัดการกระจายของข้อมูล ในขณะที่สถิติอนุมานเป็นส่วนที่เกี่ยวกับการวิเคราะห์ข้อมูลที่เรียกว่า การอนุมานเชิงสถิติ (statistical inference) ประกอบไปด้วย การประมาณค่า (estimation) และการทดสอบสมมติฐาน (testing of hypotheses) ในปัจจุบันนี้การอนุมานเชิงสถิตินั้นสามารถแบ่งออกได้เป็น 2 แนวคิดได้แก่ แนวคิดการอนุมานเชิงสถิติแบบดั้งเดิม (classical inference) และแนวคิดการอนุมานเชิงสถิติแบบเบส์ (Bayesian inference) (ประชุม สุวัตถี, 2545; ธีระพร วีระถาวร, 2536) แนวคิดของสถิติทั้งสองแนวมีความแตกต่างกัน สถิติแบบดั้งเดิมนั้นมีข้อสมมติพื้นฐานว่า “พารามิเตอร์ที่สนใจจะศึกษานั้นเป็นค่าคงตัวที่ไม่ทราบค่า ส่วนค่าสังเกตนั้นเป็นตัวอย่างที่สุ่มมาจากประชากรที่มีการแจกแจงความน่าจะเป็น” จากข้อสมมติของกระบวนการในการอนุมานเชิงสถิติแบบดั้งเดิมนั้นจะอาศัยการแจกแจงความน่าจะเป็นของตัวอย่างสุ่ม (sampling distribution) ที่ได้จากค่าสังเกตเพื่ออนุมานเกี่ยวกับพารามิเตอร์ที่สนใจ สารสนเทศที่ใช้นั้นจะมาจากข้อมูลเชิงประจักษ์ที่เก็บรวมรวมมาแต่

เพียงอย่างเดียว ในขณะที่สถิติแบบเบส์มีข้อสมมติพื้นฐานว่า "พารามิเตอร์ที่สนใจจะศึกษานั้นเป็นตัวแปรสุ่ม และตัวอย่างสุ่มที่มาจากประชากรนั้นเป็นค่าคงที่" จากข้อสมมติดังกล่าวกระบวนการอนุมานเชิงสถิติแบบเบส์นั้นจึงใช้การแจกแจงความน่าจะเป็นของพารามิเตอรเป็นเครื่องมือในการอนุมาน เรียกการแจกแจงความน่าจะเป็นดังกล่าวว่า การแจกแจงความน่าจะเป็นภายหลัง (posterior probability distribution) ข้อเด่นข้อหนึ่งของสถิติแบบเบส์ซึ่งเหนือกว่าสถิติแบบดั้งเดิม คือ สถิติแบบเบส์นั้นยอมให้ผู้วิจัยใช้สารสนเทศจากแหล่งอื่นๆนอกจากข้อมูลเชิงประจักษ์มามีส่วนในการวิเคราะห์ด้วย โดยมีหลักการคือ การคำนวณการแจกจแจงความน่าจะเป็นภายหลังของพารามิเตอร์ในโมเดลซึ่งจะเป็นเครื่องมือสำคัญในการอนุมานสถิติเชิงเบส์จากการแจกแจงความน่าจะเป็นก่อนหน้า (prior probability distribution) และฟังก์ชันภาวะความควรจะเป็น การรวมฟังก์ชันทั้งสองเข้าด้วยกันเพื่อหาฟังก์ชันของการแจกแจงความน่าจะเป็นภายหลังนั้นโดยหลักการสามารถกระทำได้โดยอาศัยทฤษฎีของเบส์ (bayes' theorem) อย่างไรก็ตามในกรณีที่ปัญหาที่สนใจมีความซับซ้อน จำนวนพารามิเตอร์ในโมเดลมักมีจำนวนมาก ดังนั้นฟังก์ชันการแจกแจงความน่าจะเป็นภายหลังที่ต้องการจึงมีจำนวนมิติจำนวนมาก การพิสูจน์เพื่อหาสูตรของฟังก์ชันการแจกแจงความน่าจะเป็นภายหลังในรูปปิดจึงกระทำได้ยาก ในกรณีดังกล่าวจึงมักใช้วิธีการเชิงจำลอง (simulation base method) เพื่อประมาณการแจกแจงความน่าจะเป็นภายหลัง วิธีการที่นิยมใช้กันโดยทั่วไปคือการใช้เทคนิคลูกโซ่มาร์คอฟมอนติคาร์โล (Monte Carlo Markov Chain: MCMC) (ธีระพร วีระถาวร, 2539; Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995; Iversen, 1989)

2.1 <u>แนวคิดเบื้องต้นของสถิติแบบเบส์</u>

สถิติแบบเบส์มีแนวคิดที่แตกต่างออกไปจากสถิติแบบดั้งเดิม กล่าวคือข้อสมมติพื้นฐานของสถิติแบบเบส์กำหนดให้ "พารามิเตอร์เป็นตัวแปรสุ่ม (random variables) ที่มีการแจกแจงความน่าจะเป็น ส่วนค่าสังเกตที่สุ่มมาจากประชากรนั้นถูกกำหนดให้เป็นค่าคงที่" เหตุผลเบื้องหลังในการกำหนดข้อสมมติดังกล่าวเนื่องมาจากความเชื่อที่ว่า พารามิเตอร์นั้นเป็นสิ่งที่นักวิจัยหรือนักสถิติไม่ทราบค่าแต่มีความต้องการที่จะคาดคะเนหรืออธิบาย ดังนั้นจึงมีความไม่แน่นอนในค่าพารามิเตอร์ดังกล่าวเกิดขึ้นซึ่งในทางทฤษฎีความน่าจะเป็นจะสามารถอธิบายความไม่แน่นอน

ได้ด้วยการแจกแจงความน่าจะเป็น ข้อสมมติดังกล่าวยังทำให้ความหมายของคำว่าความน่าจะเป็นในสถิติแบบเบส์นั้นมีความหมายในเชิงของระดับความเชื่อ (degree of belief) กล่าวคือมีความเป็นอัตนัย (subjective) มากกว่าสถิติแบบดั้งเดิม หัวใจของสถิติแบบเบส์มีส่วนประกอบที่สำคัญอยู่ 3 ส่วน ดังนี้ สมมติให้ $\underline{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ เป็นเวกเตอร์ของพารามิเตอร์ที่สนใจจะศึกษา (Gelman, Carlin, Stern, and Rubin, 1995)

1. การแจกแจงความน่าจะเป็นของพารามิเตอร์ $\underline{\theta}$ (เขียนแทนด้วยสัญลักษณ์ $p(\underline{\theta})$ หรือ $\pi(\underline{\theta})$) เรียกส่วนนี้ว่า “การแจกแจงความน่าจะเป็นก่อนหน้า” (prior distribution) เป็นการแจกแจงที่เก็บรวบรวมสารสนเทศเกี่ยวกับความรู้ก่อนหน้าเกี่ยวกับพารามิเตอร์ก่อนการเก็บรวบรวมข้อมูลจากงานวิจัย ซึ่งความรู้ก่อนหน้าดังกล่าวสามารถเก็บรวบรวมมาได้หลายแนวทาง เช่น ทฤษฎีที่มีมาก่อนหน้า ผลการวิจัยในอดีต ความรู้ประสบการณ์หรือความเชื่อจากผู้เชี่ยวชาญหรือจากผู้วิจัยเกี่ยวกับปัญหานั้น

2. การแจกแจงความน่าจะเป็นของข้อมูล $\underline{y}$ เมื่อกำหนดพารามิเตอร์ $\underline{\theta}$ เขียนแทนด้วยสัญลักษณ์ $p(\underline{y} \mid \underline{\theta})$ เรียกส่วนนี้ว่า “ฟังก์ชันภาวะความควรจะเป็น” (likelihood function)

3. การแจกแจงความน่าจะเป็นของพารามิเตอร์ $\underline{\theta}$ ภายหลังที่ได้จากการปรับสารสนเทศระหว่างความรู้ก่อนหน้ากับข้อมูลเชิงประจักษ์ เรียกส่วนนี้ว่า “การแจกแจงความน่าจะเป็นภายหลัง” (posterior distribution) เขียนแทนด้วยสัญลักษณ์ $p(\underline{\theta} \mid \underline{y})$

ในการหาการแจกแจงความน่าจะเป็นภายหลังนั้นสามารถหาได้จากการรวมสารสนเทศจากความรู้ก่อนหน้าที่จะอยู่ในรูปของการแจกแจงความน่าจะเป็นก่อนหน้าและข้อมูลเชิงประจักษ์ที่อยู่ในรูปของฟังก์ชันภาวะความควรจะเป็น โดยอาศัยทฤษฎีของเบส์ (Bayes’ theorem) (ธีระพร วีระถาวร, 2539; Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995; Iversen, 1989) ดังนี้

$$p(\underline{\theta} \mid \underline{y}) = \frac{p(\underline{\theta}, \underline{y})}{p(\underline{y})} = \frac{p(\underline{y} \mid \underline{\theta}) \cdot p(\underline{\theta})}{p(\underline{y})} \tag{2.52}$$

เมื่อ $p(\underline{y}) = E\left(p(\underline{y};\underline{\theta})\right)$ ซึ่งมีค่าเท่ากับ $\iint_{\Theta} ... \int p(\underline{y} \mid \underline{\theta}) p(\underline{\theta}) d\underline{\theta} = \iint_{\Theta} ... \int p(\underline{\theta}, \underline{y}) d\underline{\theta}$ ในกรณีที่เป็นตัวแปรสุ่มต่อเนื่อง หรือเท่ากับ $\sum \sum_{\Theta} ... \sum p(\underline{y} \mid \underline{\theta}) p(\underline{\theta}) = \sum \sum_{\Theta} ... \sum p(\underline{\theta}, \underline{y})$ ในกรณีที่

เป็นตัวแปรสุ่มไม่ต่อเนื่อง (ในวิทยานิพนธ์ฉบับนี้หากไม่ได้ระบุเป็นอื่นจะถือว่าพารามิเตอร์ที่สนใจเป็นตัวแปรสุ่มแบบต่อเนื่องเท่านั้น) เรียก $p(\underline{y})$ ว่า ค่าคงที่เพื่อปรับให้เป็นมาตรฐาน (normalizing constant) ของการแจกแจงความน่าจะเป็นภายหลัง นอกจากนี้ยังนิยามว่าฟังก์ชันภาวะความควรจะเป็น (likelihood function) ของพารามิเตอร์ $\underline{\theta}$ เขียนแทนด้วย $L(\underline{\theta})$ คือฟังก์ชันที่เป็นสัดส่วนกับการแจกแจงความน่าจะเป็น $p(\underline{y} \mid \underline{\theta})$ กล่าวคือ $L(\underline{\theta}) \propto p(\underline{y} \mid \theta)$ เราจึงสามารถเขียนสมการที่ (2.52) ได้ใหม่ดังนี้

$$p(\underline{\theta} \mid \underline{y}) = \frac{L(\underline{\theta}) \cdot p(\underline{\theta})}{\iint\limits_{\Theta} \ldots \int L(\underline{\theta}) \cdot p(\underline{\theta}) d\underline{\theta}} \tag{2.53}$$

พจน์ $p(\underline{y})$ จะเห็นว่าเป็นฟังก์ชันที่ขึ้นกับข้อมูลค่าสังเกต $\underline{y}$ เพียงอย่างเดียวซึ่งเป็นค่าคงที่และไม่ได้มีส่วนช่วยในการเพิ่มสารสนเทศของการแจกแจงความน่าจะเป็นภายหลังแต่อย่างใด ดังนั้นเราจึงสามารถเขียนการแจกแจงความน่าจะเป็นภายหลังให้อยู่ในรูปดังต่อไปนี้

$$p(\underline{\theta} \mid \underline{y}) \propto p(\underline{y} \mid \underline{\theta}) \times p(\underline{\theta}) \tag{2.54}$$

หรือ $$p(\underline{\theta} \mid \underline{y}) \propto L(\underline{\theta}) \times p(\underline{\theta}) \tag{2.55}$$

จากสมการที่ (2.55) จะเห็นว่าหลักการของสถิติแบบเบส์นั้นแท้จริงแล้วเป็นการปรับสารสนเทศที่ได้ระหว่างสารสนเทศที่มาจากความรู้ก่อนหน้าและสารสนเทศจากข้อมูลเชิงประจักษ์นั่นเอง นอกจากนี้หากพิจารณาในแง่มุมทางคณิตศาสตร์จะพบว่า สถิติแบบดั้งเดิมเป็นสถิติที่ใช้สารสนเทศจากการแจกแจงความน่าจะเป็นของตัวอย่างสุ่ม (sampling distribution) ที่อยู่ภายในปริภูมิตัวอย่าง (sample space) เพื่ออนุมานเกี่ยวกับพารามิเตอร์ที่อยู่ในปริภูมิพารามิเตอร์ (parameter space) ในขณะที่สถิติแบบเบส์นั้นใช้สารสนเทศจากการแจกแจงความน่าจะเป็นภายหลัง (posterior distribution) ที่อยู่ในปริภูมิพารามิเตอร์เพื่ออนุมานพารามิเตอร์ที่สนใจในปริภูมิพารามิเตอร์ จากตรรกะดังกล่าวหากได้มีการพิจารณากำหนดการแจกแจงความน่าจะเป็นก่อนหน้าได้อย่างเหมาะสมแล้ว ผลการวิเคราะห์ที่ได้จากสถิติแบบเบส์นั้นย่อมมีความสมบูรณ์มากกว่าการใช้สถิติแบบดั้งเดิม

2.2 <u>การแจกแจงความน่าจะเป็น (probability distribution)</u>

จากที่กล่าวมาในเบื้องต้นจะเห็นได้ว่าการแจกแจงความน่าจะเป็นนั้นมีบทบาทที่สำคัญต่อสถิติแบบเบส์เป็นอย่างมาก ในส่วนนี้จะนำเสนอรายละเอียดเกี่ยวกับการแจกแจงความน่าจะเป็นที่สำคัญที่ใช้ในงานวิจัยนี้ในการสร้างการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่สนใจ ในทฤษฎีความน่าจะเป็น (probability theory) การแจกแจงความน่าจะเป็นใช้ในการอธิบายค่าที่เป็นไปได้และลักษณะของความไม่แน่นอนของค่าของตัวแปรสุ่ม (random variable) โดยสามารถแบ่งออกได้เป็น 2 ประเภทตามประเภทของตัวแปรสุ่มได้แก่ การแจกแจงความน่าจะเป็นแบบไม่ต่อเนื่อง (discrete probability distribution) และการแจกแจงความน่าจะเป็นแบบต่อเนื่อง (continuous probability distribution) นอกจากนี้หากพิจารณาโดยใช้จำนวนของตัวแปรสุ่มเป็นเกณฑ์ยังสามารถแบ่งการแจกแจงความน่าจะเป็นออกได้อีกเป็น 2 ประเภทคือ การแจกแจงความน่าจะเป็นตัวแปรเดียว (univariate probability distribution) และการแจกแจงความน่าจะเป็นหลายตัวแปร (multivariate probability distribution) อย่างไรก็ตามในงานวิจัยนี้จะกล่าวรายละเอียดเฉพาะการแจกแจงความน่าจะเป็นแบบต่อเนื่องที่ใช้ในการวิจัยประกอบไปด้วย การแจกแจงความน่าจะเป็นแบบเอกรูป (uniform distribution) การแจกแจงความน่าจะเป็นแบบปกติตัวแปรเดียว (univariate normal distribution) การแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปร (multivariate normal distribution) การแจกแจงความน่าจะเป็นแกมมา (gamma distribution) การแจกแจงความน่าจะเป็นแกมมาผกผัน (inverse gamma distribution) การแจกแจงความน่าจะเป็นไคสแควร์ (chi-square distribution) การแจกแจงความน่าจะไคสแควร์ผกผัน (inverse chi-square distribution) การแจกแจงความน่าจะเป็น Wishart (Wishart distribution) และ การแจกแจงความน่าจะเป็น Wishart ผกผัน (inverse Wishart distribution) รายละเอียดเป็นดังต่อไปนี้

การแจกแจงความน่าจะเป็นแบบเอกรูป (uniform distribution)

การแจกแจงความน่าจะเป็นแบบเอกรูปแบบต่อเนื่องจะใช้ในการแทนลักษณะตัวแปรที่มีค่าอยู่ภายในช่วง และแต่ละค่าของตัวแปรภายในช่วงนั้นมีโอกาสในการเกิดเท่ากัน สัญลักษณ์ $\theta \sim U(a,b)$ แทนความหมายว่าตัวแปรสุ่มมีแจกแจงความน่าจะเป็นแบบเอกรูปที่มีพารามิเตอร์

ขอบเขตตั้งแต่ a ถึง b การแจกแจงแบบเอกรูปมักเป็นที่นิยมใช้ในกรณีที่ผู้วิจัยเลือกที่จะใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศ (noninformative prior distribution) ซึ่งมักกำหนดให้ลิมิตของ a และ b มีค่าลู่เข้าสู่ $-\infty$ และ ∞ ตามลำดับ ฟังก์ขันการแจกแจงความน่าจะเป็นแบบเอกรูปมีฟังก์ชันความหนาแน่น (density function) ดังสมการที่ (2.56) และกราฟฟังก์ชันความหนาแน่นดังรูปที่ 2.2 ก

$$p(\theta) = \frac{1}{b-a} \qquad \text{เมื่อ} \qquad \theta \in [a,b] \tag{2.56}$$

โดยที่ ค่าคาดหวัง (expected value) ของ θ คือ $E(\theta) = \frac{1}{2}(a+b)$ และความแปรปรวน (variance) คือ $Var(\theta) = \frac{1}{12}(b-a)^2$

การแจกแจงความน่าจะเป็นปกติตัวแปรเอกนาม (univariate normal distribution)

การแจกแจงความน่าจะเป็นแบบปกติ หรืออาจเรียกอีกชื่อหนึ่งว่า การแจกแจงเกาส์เซียน (Guassian distribution) เป็นการแจกแจงที่มีความสำคัญต่อการอนุมานเชิงสถิติมาก เนื่องจากการประมาณค่าพารามิเตอร์ และตัวสถิติทดสอบส่วนใหญ่ล้วนมีข้อสมมติเบื้องต้นให้ข้อมูลมีการแจกแจงความน่าจะเป็นแบบปกติ ลักษณะที่สำคัญของตัวแปรสุ่มที่มีการแจกแจงแบบปกติได้แก่ 1) ค่าที่เป็นไปได้ของข้อมูลอยู่ในช่วง $(-\infty,\infty)$ 2) การแจกแจงมีลักษณะสมมาตรรอบค่าเฉลี่ย มัธยฐาน และฐานนิยาม และ 3) ลักษณะของการแจกแจงมีลักษณะเป็นโค้งระฆังคว่ำ (bell shape curve) พารามิเตอร์ที่ใช้ในการ ระบุลักษณะของการแจกแจงมี 2 ตัว ได้แก่ พารามิเตอร์ค่าเฉลี่ย μ และพารามิเตอร์ความแปรปรวน σ^2 ซึ่งจะใช้สัญลักษณ์ $\theta \sim N(\mu,\sigma^2)$ แทนความหมายว่าตัวแปรสุ่ม θ มีการแจกแจงความน่าจะเป็นแบบปกติที่มีพารามิเตอร์ค่าเฉลี่ย μ และพารามิเตอร์ความแปรปรวน σ^2 ฟังก์ชันความหนาแน่นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติเป็นดังนี้

$$p(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(\theta-\mu)^2\right\} \text{ โดยที่ } \theta \in \Re \qquad \text{และ } \sigma > 0 \tag{2.57}$$

รูปที่ 2.2 ข แสดงกราฟของฟังก์ชันความหนาแน่นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าพารามิเตอร์ต่างๆกัน การแจกแจงแบบปกติมีค่าคาดหวังเท่ากับ $E(\theta)=\mu$ และความแปรปรวนเท่ากับ $Var(\theta)=\sigma^2$ ในกรณีที่กำหนดให้ $\mu=0$ และ $\sigma^2=1$ จะได้การแจกแจงความน่าจะเป็นที่เป็นกรณีเฉพาะของการแจกแจงแบบปกติเรียกว่า การแจกแจงแบบปกติมาตรฐาน (standard normal) เขียนแทนด้วย $\theta \sim N(0,1)$

การแจกแจงความน่าจะเป็นแบบปกติตัวแปรพหุนาม (multivariate normal distribution)

การแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรเป็นการแจกแจงความน่าจะเป็นของเวกเตอร์สุ่ม (random vector) ซึ่งเป็นกรณีทั่วไปของการแจกแจงความน่าจะเป็นแบบปกติตัวแปรเอกนามดังที่ได้กล่าวไปแล้ว พารามิเตอร์ของการแจกแจงแบบปกติตัวแปรพหุนามประกอบไปด้วยเวกเตอร์ค่าเฉลี่ย $\underline{\mu}$ ขนาด $n \times 1$ และเมทริกซ์ความแปรปรวนร่วม (variance-covariance matrix) Σ ขนาด $n \times n$ ที่มีคุณสมบัติเป็นบวกแน่นอน (positive definite matrix) เวกเตอร์สุ่ม $\underline{\theta}=(\theta_1,\theta_2,...,\theta_n)$ ที่มีการแจกแจงความน่าจะเป็นแบบปกติตัวแปรพหุนามเขียนแทนด้วย $\underline{\theta} \sim MVN(\underline{\mu},\Sigma)$ จะมีฟังก์ชันความหนาแน่นดังนี้

$$p(\underline{\theta}) = (2\pi)^{-d/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\underline{\theta}-\underline{\mu})^T \Sigma^{-1} (\underline{\theta}-\underline{\mu})\right\} \qquad (2.58)$$

ในกรณีที่พารามิเตอร์ของเวกเตอร์ค่าเฉลี่ยมีค่าเท่ากับศูนย์ และพารามิเตอร์เมทริกซ์ความแปรปรวนร่วมเป็นเมทริกซ์เอกลักษณ์จะเรียกการแจกแจงของเวกเตอร์สุ่มดังกล่าวว่า การแจกแจงความน่าจะเป็นแบบปกติมาตรฐานหลายตัวแปร รูปที่ 2.2 ค แสดงแสดงกราฟของฟังก์ชันความหนาแน่นของการแจกแจงแบบปกติมาตรฐาน 2 ตัวแปรในกรณีที่พารามิเตอร์เวกเตอร์ค่าเฉลี่ยเป็นเวกเตอร์ศูนย์และพารามิเตอร์เมทริกซ์ความแปรปรวนร่วมมีค่าเท่ากับ $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$

ในทางทฤษฎีนั้นหากเวกเตอร์สุ่ม $\underline{\theta}=(\theta_1,\theta_2,...,\theta_n)$ มีการแจกแจงแบบปกติหลายตัวแปรแล้วจะได้ว่าตัวแปรสุ่ม $\theta_1,\theta_2,...,\theta_n$ แต่ละตัวจะมีการแจกแจงแบบปกติด้วย แต่ในทางกลับกันหาก $\theta_1,\theta_2,...,\theta_n$ แต่ละตัวมีการแจกแจงความน่าจะเป็นแบบปกติแล้ว ไม่จำเป็นที่เวกเตอร์สุ่ม $\underline{\theta}=(\theta_1,\theta_2,...,\theta_n)$ จะต้องมีการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรด้วย

การแจกแจงความน่าจะเป็นแกมมา (gamma distribution)

การแจกแจงความน่าจะเป็นแกมมาเป็นการแจกแจงของตัวแปรสุ่มที่มีค่าอยู่ในช่วง $(0,\infty)$ ซึ่งมีพารามิเตอร์ระบุลักษณะการแจกแจง 2 ตัวได้แก่ พารามิเตอร์ α (shape parameter) และ β (scale parameter) โดยที่ $\alpha > 0$ และ $\beta > 0$ หากตัวแปรสุ่ม θ มีการแจกแจงความน่าจะเป็นแกมมาที่มีพารามิเตอร์ในข้างต้น เขียนแทนด้วย $\theta \sim Gamma(\alpha,\beta)$ จะได้ว่า θ มีฟังก์ชันความหนาแน่นเป็นดังนี้

$$p(\theta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}\exp\{-\beta\theta\} \tag{2.59}$$

เมื่อ Γ เป็นฟังก์ชันแกมมา $\Gamma:[0,\infty)\to\Re$ นิยามโดย $\Gamma(\alpha)=\int_0^{\infty}\theta^{\alpha-1}\exp\{-\theta\}d\theta$ กราฟของฟังก์ชันความหนาแน่นที่มีการแจกแจงแบบแกมมาที่ค่าพารามิเตอร์ต่างๆ แสดงไว้ในรูปที่ 2.2 ง

ตัวแปรสุ่ม θ ที่มีการแจกแจงแกมมาจะมีค่าคาดหวังเท่ากับ $E(\theta)=\alpha\beta$ และมีความแปรปรวนเท่ากับ $Var(\theta)=\alpha\beta^2$ การแจกแจงความน่าจะเป็นแกมมาเป็นกรณีทั่วไปของการแจกแจงความน่าจะเป็นหลายการแจกแจง เช่น

ถ้า $\theta \sim Gamma(\alpha=1,\beta=1/\lambda)$ จะได้ว่า θ จะมีการแจกแจงแบบเอกซ์โพเนนเชียล (exponential distribution) ที่มีพารามิเตอร์ λ

ถ้า $\theta \sim Gamma(\alpha=n/2,\beta=2)$ จะได้ว่า θ จะมีการแจกแจงแบบไคสแควร์ (chi-square distribution) ที่มีพารามิเตอร์องศาความเป็นอิสระ (degree of freedom) เท่ากับ n

นอกจากนี้การแจกแจงความน่าจะเป็นแกมมามีคุณสมบัติเป็นการแจกแจงความน่าจะเป็นก่อนหน้าวงศ์คู่สังยุค (conjugate prior) กับฟังก์ชันการแจกแจงภาวะความควรจะเป็นหลายตัว เช่น การแจกแจงความน่าจะเป็นปัวซงส์ (Poisson distribution) การแจกแจงเอกโพเนนเชียล (exponential distribution) การแจกแจงแบบปกติ (เมื่อสมมติว่าทราบค่าเฉลี่ย) การแจกแจงพาเรโต (Pareto distribution) การแจกแจงแกมมา (เมื่อทราบ shape parameter) และการแจกแจงแกมมาผกผัน (inverse gamma เมื่อทราบ shape parameter)

การแจกแจงความน่าจะเป็นแบบแกมมาผกผัน (inverse gamma distribution)

การแจกแจงความน่าจะเป็นแกมมาผกผัน เป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่มีค่าอยู่ในช่วง $(0,\infty)$ เช่นเดียวกับการแจกแจงแบบแกมมา ซึ่งมีคุณสมบัติที่ส่วนกลับของตัวแปรสุ่มที่มีการแจกแจงแกมมาผกผันจะมีการแจกแจงแกมมา กล่าวคือ ถ้า $\theta \sim Inv-Gamma(\alpha,\beta)$ มีการแจกแจงความน่าจะเป็นแกมมาผกผันที่มีพารามิเตอร์ α (shape parameter) และ β (scale parameter) โดยที่ $\alpha>0$ และ $\beta>0$ จะได้ว่า $1/\theta \sim Inv-Gamma(\alpha,1/\beta)$ การแจกแจงความน่าจะเป็นแกมมาผกผัน เป็นการแจกแจงความน่าจะเป็นก่อนหน้าวงศ์คู่สังยุคของพารามิเตอร์ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติ ฟังก์ชันความหนาแน่นของ θ ที่มีพารามิเตอร์ α และ β เป็นดังนี้

$$p(\theta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta^{-(\alpha+1)}\exp\{-\beta\theta\} \tag{2.60}$$

ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแกมมาผกผัน จะมีค่าคาดหวังเป็น $E(\theta)=\dfrac{\beta}{\alpha-1}$ เมื่อ $\alpha>1$ และมีความแปรปรวนเป็น $Var(\theta)=\dfrac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ เมื่อ $\alpha>2$ กราฟของฟังก์ชันความหนาแน่นของการแจกแจงความน่าจะเป็นแบบแกมมาผกผันที่ค่าพารามิเตอร์ต่างๆ แสดงไว้ในรูปที่ 2.2 จ

การแจกแจงความน่าจะเป็นไคสแควร์ (chi-square distribution)

การแจกแจงความน่าจะเป็นไคสแควร์ เป็นการแจกแจงความน่าจะเป็นที่สร้างขึ้นจากการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน กล่าวคือ ผลรวมกำลังสองของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐานที่เป็นอิสระซึ่งกันและกันจำนวน n ตัวจะมีการแจกแจงความน่าจะเป็นแบบไคสแควร์ที่มีพารามิเตอร์องศาความเป็นอิสระเท่ากับ n กล่าวคือ ถ้า $Z_i \overset{iid}{\sim} N(0,1)$ จะได้ว่า $\theta=\sum_{i=1}^{n}Z_i^2 \sim \chi^2_{(n)}$ ซึ่งมีฟังก์ชันความหนาแน่นดังนี้

$$p(\theta)=\frac{2^{-n/2}}{\Gamma(n/2)}\theta^{n/2-1}\exp\{-\theta/2\} \qquad \text{เมื่อ } \theta>0 \tag{2.61}$$

โดยที่ θ มีค่าคาดหวังเป็น $E(\theta)=n$ และความแปรปรวนเป็น $Var(\theta)=2n$ การแจกแจงความน่าจะเป็นไคสแควร์เป็นกรณีเฉพาะของการแจกแจงความน่าจะเป็นแกมมาดังที่ได้กล่าวไว้ก่อนหน้านี้ กราฟของฟังก์ชันความหนาแน่นของการแจกแจงความน่าจะเป็นแบบไคสแควร์ที่ค่าพารามิเตอร์องศาความเป็นอิสระต่างๆ แสดงไว้ในรูปที่ 2.1 ฉ

จากทฤษฎีลิมิตลู่เข้าสู่ส่วนกลาง (central limit theorem) จะได้ว่าตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์นั้นจะมีการแจกแจงที่ลู่เข้าสู่การแจกแจงความน่าจะเป็นแบบปกติเมื่อ $n\to\infty$ (ในทางปฏิบัติ $n\approx 50$) กล่าวคือ ถ้า $\theta\sim\chi^2_{(n)}$ จะได้ว่า $\sqrt{2\theta}\to N(0,1)$ เมื่อ $n\to\infty$ นอกจากนี้สมมติว่า $\theta_N=\sum_{i=1}^{n}X_i$ โดยที่ $X_i\sim N(\mu,\sigma^2)$ และ $\mu\neq 0$ จะได้ว่า $\theta_N\sim noncentral\ \chi^2_{n,\lambda}$ เมื่อ n คือพารามิเตอร์องศาความเป็นอิสระ และ λ คือพารามิเตอร์ไม่ใช่ศูนย์กลาง (noncentral parameter)

การแจกแจงความน่าจะเป็นไคสแควร์ผกผัน (inverse chi-square distribution)

ในทำนองเดียวกับการแจกแจงความน่าจะเป็นแกมมาผกผัน การแจกแจงความน่าจะเป็นไคสแควร์ผกผันสร้างจากส่วนกลับของตัวแปรสุ่มไคส กล่าวคือ สมมติว่า $\theta\sim\chi^2_{(n)}$ จะได้ว่า $1/\theta\sim Inv-\chi^2_{(n)}$ โดยที่มีฟังก์ชันความหนาแน่นดังนี้

$$p(\theta)=\frac{2^{-n/2}}{\Gamma(n/2)}\theta^{-(n/2+1)}\exp\{-1/2\theta\} \qquad \text{เมื่อ } \theta>0 \qquad (2.62)$$

ค่าคาดหวังและความแปรปรวนของตัวแปรสุ่มดังกล่าวเป็น $E(\theta)=\dfrac{1}{n-2}$ เมื่อ $n>2$ และ $Var(\theta)=\dfrac{2}{(n-2)^2(n-4)}$ เมื่อ $n>4$ ตามลำดับการแจกแจงไคสแควร์ผกผันใช้ประโยชน์มากทางสถิติวิเคราะห์แบบเบส์ในการสร้างการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ความแปรปรวน

การแจกแจงความน่าจะเป็นไคสแควร์ผกผันนั้นเป็นกรณีเฉพาะของการแจกแจงความน่าจะเป็นสเกลไคสแควร์ผกผัน (scaled inverse chisquare distribution) ในกรณีที่ $\sigma^2=1/n$ และการแจกแจงสเกลไคสแควร์ผกผันเป็นกรณีเฉพาะของการแจกแจงความน่าจะเป็นแกมมา

ผกผัน กล่าวคือ $\theta \sim Inv-Gamma(\frac{n}{2},\frac{n\sigma^2}{2})$ แล้วจะได้ว่า $\theta \sim Scale-inv-\chi^2_{n,\sigma^2}$ เมื่อ n เป็นพารามิเตอร์องศาความเป็นอิสระและ σ^2 เป็นพารามิเตอร์สเกล ฟังก์ชันการแจกแจงความน่าจะเป็นสเกลไคสแควร์ผกผัน ที่มีพารามิเตอร์ทั้งสองดังกล่าวเขียนได้ดังนี้

$$p(\theta)=\frac{(\sigma^2 n/2)^{n/2}}{\Gamma(n/2)}\theta^{-(n/2+1)}\exp\left\{-n\sigma^2/2\theta\right\} \text{ เมื่อ } \theta>0 \qquad (2.63)$$

โดยที่มีค่าคาดหวังและความแปรปรวนของตัวแปรสุ่มดังกล่าวเป็น $E(\theta)=\frac{n\sigma^2}{n-2}$ เมื่อ $n>2$ และ $Var(\theta)=\frac{2n^2\sigma^4}{(n-2)^2(n-4)}$ เมื่อ $n>4$ ตามลำดับ การแจกแจงความน่าจะเป็นสเกลไคสแควร์ผกผันมีคุณสมบัติเป็นการแจกแจงความน่าจะเป็นก่อนหน้าวงศ์คู่สังยุคในการประมาณค่าพารามิเตอร์ความแปรปรวนสำหรับการแจกแจงแบบปกติ

การแจกแจงความน่าจะเป็น Wishart (Wishart distribution)

การแจกแจงดังกล่าวเป็นการแจกแจงของเมทริกซ์สุ่ม (random matrix) สมมติว่า X เป็นเมทริกซ์ขนาด $n\times p$ โดยที่แต่ละคอลัมน์ (column) ของเมทริกซ์ดังกล่าวเป็นเวกเตอร์สุ่มที่เป็นอิสระซึ่งกันและกันและมีการแจกแจงความน่าจะเป็นแบบปกติ p ตัวแปร (p variate normal distribution) จะได้ว่าการแจกแจงความน่าจะเป็นของเมทริกซ์สุ่ม $S=X^TX$ (เรียกว่า scatter matrix) ขนาด $p\times p$ มีการแจกแจงความน่าจะเป็นแบบ Wishart ที่มีพารามิเตอร์ Σ p และ n เขียนแทนด้วย $S\sim Wishart_p(n,\Sigma)$ จากการนิยามในข้างต้นกำหนดให้ p และ n เป็นจำนวนเต็มบวก และเมทริกซ์สเกล (scale matrix) Σ เป็นเมทริกซ์ที่มีคุณสมบัติ positive semi definite โดยที่มีฟังก์ชันความหนาแน่นดังนี้

$$p(S\,|\,\Sigma)=\left(2^{np/2}\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\left(\frac{n+1-i}{2}\right)\right)^{-1}\times|\Sigma|^{-n/2}\times|S|^{(\nu-k-1)/2}\times\exp\left\{-\frac{1}{2}tr\left(\Sigma^{-1}S\right)\right\} \qquad (2.64)$$

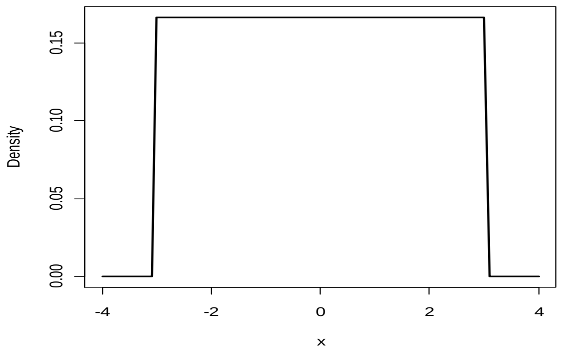
การแจกแจงความน่าจะเป็น Wishart เป็นกรณีทั่วไปของการแจกแจงความน่าจะเป็นไคสแควร์ กล่าวคือในกรณีที่ $p=\Sigma=1$ จะได้ว่าตัวแปรสุ่ม S จะมีการแจกแจงความน่าจะเป็นแบบไคสแควร์ที่มีองศาความเป็นอิสระเท่ากับ n และ การแจกแจงความน่าจะเป็น Wishart เป็นการแจกแจงความน่าจะเป็นที่มักใช้ในการอธิบาย เมทริกซ์ความแปรปรวนร่วมที่ได้จากตัวอย่าง (sample covariance matrix) ที่มีการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปร สมมติว่า $\underline{x} \sim N(\underline{\mu}, \Sigma)$ โดยที่ $\underline{x}$ เป็นเวกเตอร์สุ่มขนาด $p \times 1$ จะได้ว่า $S_\mu = (\underline{x} - \underline{\mu})(\underline{x} - \underline{\mu})^T \sim Wishart_p(n, \Sigma)$ และ $S_\mu = (\underline{x} - \underline{\bar{x}})(\underline{x} - \underline{\bar{x}})^T \sim Wishart_p(n-1, \Sigma)$

การแจกแจงความน่าจะเป็น Wishart ผกผัน (inverse Wishart distribution)

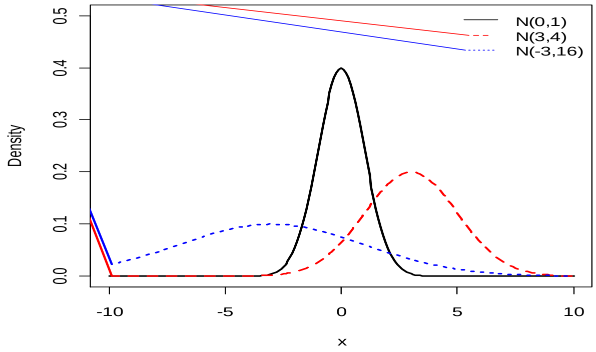
ในสถิติแบบเบส์ การแจกแจงความน่าจะเป็น Wishart ผกผันมักใช้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าวงศ์คู่สังยุค (conjugate prior) สำหรับเมทริกซ์ความแปรปรวนร่วมของเวกเตอร์สุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปร โดยเมทริกซ์สุ่ม S จะมีการแจกแจงความน่าจะเป็นแบบ Wishart ผกผัน (เขียนแทนด้วย $S \sim Inv-Wishart_p(n, \Sigma)$) เมื่อ $S^{-1} \sim Wishart_p(n, \Sigma)$ การแจกแจงความน่าจะเป็น Wishart ผกผันมีฟังก์ชันความหนาแน่นดังนี้

$$p(S \mid \Sigma^{-1}) = \left(2^{np/2} \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(\frac{n+1-i}{2} \right) \right)^{-1} \times |\Sigma|^{n/2} \times |S|^{-(n+k+1)/2} \times \exp\left\{ -\frac{1}{2} tr\left(\Sigma S^{-1} \right) \right\} \tag{2.65}$$

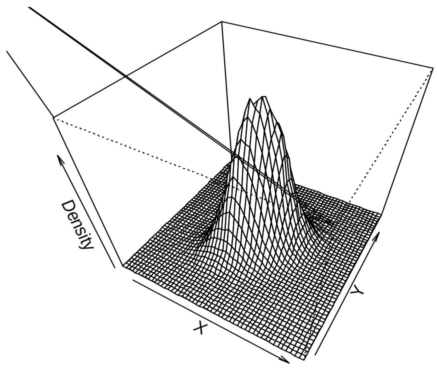
การแจกแจงความน่าจะเป็นในข้างต้นนั้นเป็นการแจกแจงที่ผู้วิจัยจะใช้ในการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ต่างๆในโมเดล ซึ่งพารามิเตอร์ในแต่ละส่วนของโมเดลนั้นมีธรรมชาติที่แตกต่างกัน ในส่วนต่อไปผู้วิจัยจึงจะนำเสนอแนวคิดในการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์แต่ละส่วนของโมเดลเชิงเส้นพหุระดับ ประกอบไปด้วย การกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์อิทธิพลคงที่ พารามิเตอร์ความแปรปรวน และพารามิเตอร์เมทริกซ์ความแปรปรวนร่วม รายละเอียดเป็นดังต่อไปนี้



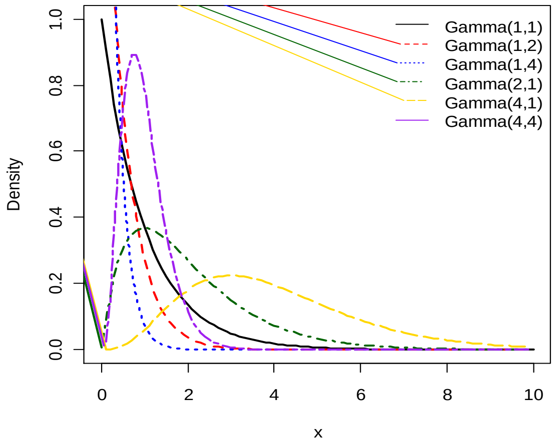
รูปที่ 2.2 ก การแจกแจงความน่าจะเป็นแบบเอกรูป (uniform distribution)



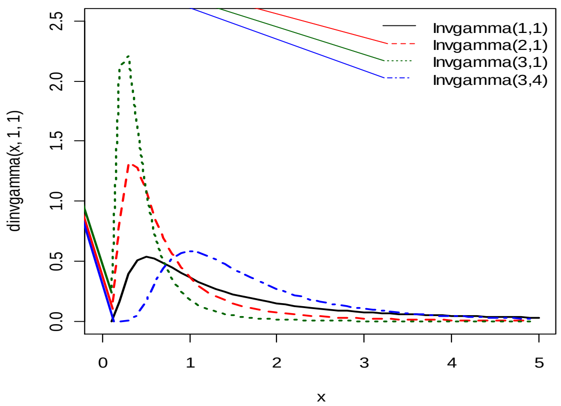
รูปที่ 2.2 ข การแจกแจงความน่าจะเป็นแบบปกติตัวแปรเดียว (univariate normal distribution)

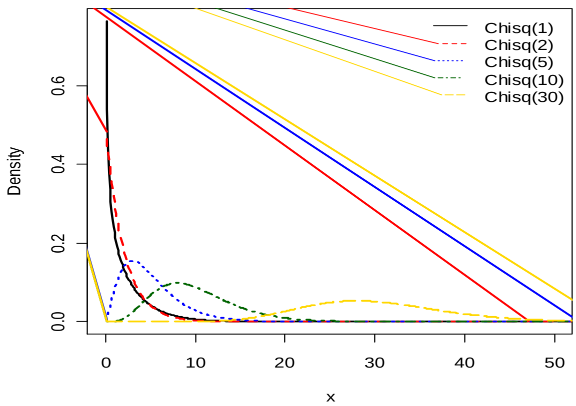


รูปที่ 2.2 ค การแจกแจงความน่าจะเป็นแบบปกติ 2 ตัวแปร (bivariate normal distribution)



รูปที่ 2.2 ง การแจกแจงความน่าจะเป็นแบบแกมมา (gamma distribution)





<u>รูปที่ 2.2 จ</u> การแจกแจงความน่าจะเป็นแกมมาผกผัน (inverse gamma distribution)

<u>รูปที่ 2.2 ฉ</u> การแจกแจงความน่าจะเป็นแบบไคสแควร์ (chi-square distribution)

2.3 <u>การแจกแจงความน่าจะเป็นก่อนหน้า (prior distributions)</u>

การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ คือการแจกแจงความน่าจะเป็นที่สร้างจากความรู้ทฤษฎีหรือความเชื่อเกี่ยวกับพารามิเตอร์ในปัญหาที่สนใจก่อนการเก็บรวบรวมข้อมูลเชิงประจักษ์ ซึ่งจะถูกนำไปรวมกับสารสนเทศจากฟังก์ชันภาวะความควรจะเป็นโดยการใช้ทฤษฎีของเบส์ทำให้สามารถคำนวณหาการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ได้ การแจกแจงความน่าจะเป็นก่อนหน้านั้นเป็นสิ่งที่ขาดไม่ได้ในสถิติแบบเบส์ การแจกแจงความน่าจะเป็นก่อนหน้าสามารถแบ่งได้ออกเป็น 2 ประเภทได้แก่ การแจกแจงความน่าจะเป็นก่อนหน้าแบบให้สารสนเทศ (informative priors) หรือการแจกแจงความน่าจะเป็นก่อนหน้าแบบเป็นอัตนัย (subjective priors) และการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศ (non-informative priors) หรือการแจกแจงความน่าจะเป็นก่อนหน้าแบบปรนัย (objective priors) (Gelman, 2002; Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995; Iversen, 1989)

การแจกแจงความน่าจะเป็นก่อนหน้าแบบให้สารสนเทศ เป็นการแจกแจงความน่าจะเป็นก่อนหน้าที่ใช้ในกรณีที่นักวิจัยมีความรู้ทฤษฎีเกี่ยวกับพารามิเตอร์ที่น่าเชื่อถือเพียงพอ จากที่มาของการกำหนดความน่าจะเป็นก่อนหน้าดังกล่าวจึงทำให้การแจกแจงความน่าจะเป็นก่อนหน้าแบบให้สารสนเทศนั้นมีความเป็นอัตนัย ซึ่งจะทำให้ผลการวิเคราะห์ในปัญหาเดียวกันนั้นมีความแตกต่างกันตามนักวิจัยแต่ละคน ดังนั้นการสร้างการแจกแจงความน่าจะเป็นก่อนหน้าแบบให้สารสนเทศจึงเป็นสิ่งที่นักวิจัยจำเป็นต้องพิจารณาอย่างรอบคอบเพื่อให้ผลการวิเคราะห์ที่ได้มีความน่าเชื่อถือ ผู้สนใจสามารถศึกษารายละเอียดเกี่ยวกับการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าแบบให้สารสนเทศได้จากเอกสารที่ระบุไว้ (Goldstein, 2006; Gelman, Carlin, Stern, and Rubin, 1995) อย่างไรก็ตามในหลายปัญหานักวิจัยเองก็อาจจะไม่ได้มีข้อมูลความรู้เกี่ยวกับพารามิเตอร์ที่สนใจมาก่อนหรือข้อมูลความรู้ที่มีอยู่ไม่เพียงพอที่จะสร้างเป็นการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ได้ ในกรณีนี้นักสถิติจึงมักแนะนำให้เลือกให้การแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศแทน(ปรางทิพย์ รัชตะปิติ, 2550; เกียรติเทพ ตั้งสันติถาวร

, 2548; วีรพา ฐานะปรัชญ์, 2542; Albert, 2009; Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995; Iversen, 1989) การแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศจะมีความเป็นปรนัยมากกว่าการแจกแจงความน่าจะเป็นก่อนหน้าแบบให้สารสนเทศ เราอาจอธิบายลักษณะของการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศโดยทั่วไปได้ว่าเป็นการแจกแจงความน่าจะเป็นก่อนหน้าที่มีอิทธิพลน้อยมากหรือไม่มีอิทธิพลต่อฟังก์ชันภาวะความควรจะเป็น ทำให้สารสนเทศที่ได้จากการวิเคราะห์เป็นสารสนเทศที่มาจากข้อมูลเชิงประจักษ์เป็นส่วนใหญ่หรืออาจมาจากข้อมูลเชิงประจักษ์ทั้งหมด เมื่อพิจารณาจากกราฟของฟังก์ชันการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารนเทศจะพบว่ามีลักษณะแบนราบเมื่อเทียบกับฟังก์ชันภาวะความควรจะเป็น ดังนั้นการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศจึงเป็นการแจกแจงความน่าจะเป็นก่อนหน้าที่ไม่มีอิทธิพลหรือไม่อิทธิพลน้อยมากต่อการสร้างการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่สนใจ การเลือกใช้การแจกแจงความน่าจะเป็นก่อนหน้าในปัญหาต่างๆนั้นมีข้อควรต้องระวังคือการแจกแจงความน่าจะเป็นก่อนหน้าที่เลือกใช้นั้นบางการแจกแจงอาจเป็น การแจกแจงความน่าจะเป็นก่อนหน้าไม่ตรงแบบ (improper prior distribution) กล่าวคือ $\iint_{\Theta}\ldots\int p(\underline{\theta})d\underline{\theta} = \infty$ ซึ่งอาจส่งผลให้ได้การแจกแจงความน่าจะเป็นภายหลังไม่ตรงแบบ (improper posterior distribution) กล่าวคือเป็นฟังก์ชันที่ไม่ได้มีคุณสมบัติของฟังก์ชันการแจกแจงความน่าจะเป็น ทำให้ไม่สามารถใช้การแจกแจงความน่าจะเป็นภายหลังดังกล่าวมาอนุมานเกี่ยวกับพารามิเตอร์ได้ ตัวอย่างเช่น สมมติว่านักวิจัยเลือกใช้การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ θ เป็นการแจกแจงแบบเอกรูปบนช่วงจำนวนจริง ซึ่งจะเป็นว่าค่าอินทิกรัลบนช่วงจำนวนจริงของฟังก์ชันความน่าจะเป็นดังกล่าวไม่สามารถหาได้ การแจกแจงความน่าจะเป็นก่อนหน้าดังกล่าวจึงเป็นการแจกแจงความน่าจะเป็นก่อนหน้าที่ไม่ตรงแบบ นอกจากนี้การแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศนี้ยังมักที่จะไม่มีคุณสมบัติไม่แปรเปลี่ยนภายใต้การแปลง (invariant under transformation) กล่าวคือการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ θ อาจเป็นการแจกแจงความน่าจะเป็นก่อนหน้าที่ไม่ให้สารสนเทศ แต่เมื่อกระทำการแปลงค่า (transformation) พารามิเตอร์ดังกล่าวด้วยฟังก์ชันการแปลง (transformation function) ฟังก์ชันการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ที่ผ่าน

การแปลงแล้วไม่จำเป็นที่จะต้องมีคุณสมบัติเป็นการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศอีก ตัวอย่างของการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศที่มักใช้กัน เช่น การแจกแจงแบบสม่ำเสมอ การแจกแจงแบบปกติที่มีค่าความแปรปรวนขนาดใหญ่ เป็นต้น (Albert, 2009; Bolstad, 2004; Browne, 1998; Gelman, Carlin, Stern, and Rubin, 1995)

ในการคำนวณค่าการแจกแจงความน่าจะเป็นภายหลัง โดยมากอาจไม่สามารถที่จะคำนวณหาการแจกแจงความน่าจะเป็นภายหลังที่ต้องการโดยการใช้วิธีเชิงวิเคราะห์ได้ เนื่องจากไม่สามารถจัดรูปให้อยู่ในรูปของการแจกแจงความน่าจะเป็นมาตรฐานได้ วิธีการหนึ่งเพื่อแก้ไขปัญหาดังกล่าวคือการใช้การแจกแจงความน่าจะเป็นก่อนหน้าที่เป็นวงศ์คู่สังยุค (conjugate prior distribution)[1] ซึ่งมีคุณสมบัติทำให้การแจกแจงความน่าจะเป็นภายหลังเป็นการแจกแจงที่เดียวกันกับการแจกแจงความน่าจะเป็นก่อนหน้า วิธีการดังกล่าวจึงทำให้นักวิจัยสามารถพิสูจน์เพื่อหารูปแบบของการแจกแจงความน่าจะเป็นภายหลังได้โดยง่าย (ธีระพร วีระถาวร, 2539; Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995; Iversen, 1989)

2.4 <u>การอนุมานเชิงสถิติแบบเบส์ (Bayesian inference)</u>

ตามที่ได้กล่าวไว้ในข้างต้นแล้วว่าการอนุมานเชิงสถิติแบบเบส์นั้นล้วนแต่ใช้การแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์เป็นเครื่องมือในการอนุมาน ซึ่งเราสามารถที่จะสรุปสารสนเทศที่เก็บรวบรวมอยู่ภายในการแจกแจงความน่าจะเป็นภายหลังได้หลายวิธีการเหมือนกับการอนุมานโดยใช้สถิติแบบดั้งเดิม ในหัวข้อนี้จะนำเสนอหลักการในการสรุปสารสนเทศจากการแจกแจงความน่าจะเป็นภายหลังในเบื้องต้นได้แก่ การประมาณค่าแบบจุด การประมาณค่าแบบช่วง และการทดสอบสมมติฐาน ซึ่งผู้ที่สนใจในรายละเอียดสามารถศึกษาเพิ่มเติมได้จากเอกสารอ้างอิงที่ได้ระบุไว้ (Albert, 2009; Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995)

---

[1] การแจกแจงความน่าจะเป็นก่อนหน้าที่มีการแจกแจงความน่าจะเป็นอยู่ในวงศ์ (family) ของการแจกแจง D ที่มีพารามิเตอร์ α จะเป็นการแจกแจงที่เป็นวงศ์คู่สังยุคกับการแจกแจง $f(\underline{y} \mid \underline{\theta})$ ถ้าการแจกแจงความน่าจะเป็นภายหลัง $f(\underline{\theta} \mid \underline{y})$ เป็นการแจกแจงที่เป็นสมาชิกในวงศ์ของ D เช่นเดียวกัน กล่าวคือ ถ้า $\underline{\theta} \sim D(\alpha)$ แล้ว $\underline{\theta} \mid \underline{y} \sim D(\tilde{\alpha})$

การประมาณค่าแบบจุด (point estimation)

ในการหาค่าประมาณแบบจุดโดยใช้แนวทางเบส์เซียนนั้น หากผู้วิจัยต้องการใช้ค่าเฉลี่ยเป็นค่าประมาณแบบจุดจะสามารถหาได้โดยใช้ค่าคาดหวังแบบมีเงื่อนไขของการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ เรียกชื่อว่า ค่าเฉลี่ยภายหลัง (posterior mean) ซึ่งหาได้ดังนี้

$$E(\theta | y) = \int \theta \cdot p(\theta | \underline{y}) d\theta \tag{2.66}$$

นอกจากนี้ยังสามารถหาค่ามัธยฐาน เรียกว่า มัธยฐานภายหลัง (posterior median) ได้จาก

$$P(\theta \le median | y) = P(\theta \ge median | y) = \frac{1}{2} \tag{2.67}$$

และฐานนิยม เรียกว่า ฐานนิยมภายหลัง (posterior mode) ได้จากการหาค่าของ θ ที่ทำให้ $p(\theta | y)$ มีค่าสูงสุดหรือเขียนแทนด้วย

$$\max_{\theta} \{p(\theta | y)\} \tag{2.68}$$

ในกรณีที่ต้องการหาค่าความแปรปรวนจากการแจกแจงความน่าจะเป็นภายหลังนั้น จะเรียกว่า ความแปรปรวนภายหลัง (posterior variance) ความหมายของค่าความแปรปรวนภายหลังมีความหมายคือ เป็นค่าที่ใช้วัดความไม่แน่นอนของค่าพารามิเตอร์ที่สนใจ ซึ่งถูกกำหนดให้เป็นตัวแปรสุ่มตามแนวคิดของสถิติแบบเบส์ ในทำนองเดียวกันนี้เราสามารถหาค่าส่วนเบี่ยงเบนมาตรฐานภายหลัง (posterior standard deviation) ความแปรปรวนร่วมภายหลัง (posterior covariance) ฯลฯ ได้จากการแจกแจงความน่าจะเป็นภายหลังนี้

การประมาณค่าแบบช่วง (interval estimation)

ในการวิเคราะห์โดยใช้สถิติแบบเบส์นั้นช่วงของการประมาณค่าพารามิเตอร์นั้นจะไม่ได้เรียกว่าช่วงความเชื่อมั่นเหมือนในสถิติแบบดั้งเดิม แต่จะเรียกว่า ช่วงความน่าเชื่อถือ (credible intervals) โดยที่ A จะเป็นช่วงความน่าเชื่อถือของพารามิเตอร์ θ ถ้า

$$P(\theta \in A | y) = \int_A p(\theta | y) d\theta \tag{2.69}$$

ยกตัวอย่างเช่น สมมติว่าต้องการสร้างช่วงความน่าเชื่อถือ 95% (95% credible interval) ของพารามิเตอร์ θ จะสามารถสร้างได้โดยหาช่วง A ที่ทำให้ $\int_A p(\theta | y) d\theta = 0.95$ การสร้างช่วง

ความน่าเชื่อถือตามวิธีนี้จะทำให้สามารถหาช่วง A ที่ทำให้มีความน่าเชื่อถือเท่ากับระดับที่กำหนดไว้ได้หลายช่วงจากการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ θ อันเดียวกัน ดังนั้นนักสถิติบางท่านอาจทำการสร้างช่วงความน่าเชื่อถือนี้โดยเลือกช่วง A ที่ทำให้ความน่าจะเป็นในส่วนที่น้อยกว่าขอบเขตล่างและมากกว่าขอบเขตบนของช่วงความน่าเชื่อถือนั้นมีขนาดเท่ากัน ช่วงความน่าเชื่อถือที่สร้างโดยแนวคิดนี้มีข้อดีคือมีความไม่แปรเปลี่ยนภายใต้การแปลงอีกด้วย

นอกจากวิธีการสร้างช่วงความน่าเชื่อถือตามที่ได้กล่าวไว้ในข้างต้นแล้ว ยังมีวิธีการสร้างช่วงความน่าเชื่อถือที่เป็นที่นิยมใช้กันเรียกว่า ฟังก์ชันความหนาแน่นภายหลังสูงสุด (highest posterior density (HPD)) ช่วงความน่าเชื่อถือที่ได้จากวิธีนี้จะเรียกว่า ช่วงความหนาแน่นภายหลังสูงสุด (highest posterior density interval) หรือเรียกสั้นๆว่า ช่วง HPD $100(1-\alpha)\%$ ช่วง HPD คือบริเวณที่สอดคล้องกับเงื่อนไข 2 ประการต่อไปนี้

1. ความน่าจะเป็นภายหลังของบริเวณนั้นมีค่าเท่ากับ $100(1-\alpha)\%$
2. ความหนาแน่นที่ต่ำที่สุดในบริเวณ HPD จะมีค่ามากกว่าหรือเท่ากับความหนาแน่น ณ จุดอื่นๆนอกเหนือบริเวณ HPD เท่านั้น

จากเงื่อนไขของช่วง HPD ที่กล่าวมานั้นจะเห็นว่าช่วง HPD เป็นช่วงครอบคลุมความหนาแน่นส่วนใหญ่ของการแจกแจงความน่าจะเป็นภายหลัง ซึ่งช่วงดังกล่าวมีคุณสมบัติที่ดีคือช่วงความน่าเชื่อถือที่ได้จะเป็นช่วงที่มีความยาวช่วงที่สั้นที่สุดเสมอเมื่อเทียบกับวิธีการสร้างช่วงความน่าเชื่อถืออื่นๆ

สิ่งที่ต้องคำนึงถึงในการวิเคราะห์ด้วยสถิติแบบเบส์คือ การแปลความหมายของสถิติแบบเบส์และแบบดั้งเดิม มีความแตกต่างกัน ในกรณีการแปลผลจาก $100(1-\alpha)\%$ ช่วงความน่าเชื่อถือนั้น นักวิจัยหรือนักสถิติสามารถอ้างหรือยืนยันได้ว่าความน่าจะเป็นที่ค่าพารามิเตอร์ที่สนใจจะอยู่ในช่วงความน่าเชื่อถือที่ได้สร้างขึ้นมานั้นมีค่าเท่ากับ $100(1-\alpha)\%$ ซึ่งเป็นระดับของความน่าเชื่อถือที่ได้กำหนดไว้ จะเห็นว่าการแปลความหมายของช่วงความน่าเชื่อถือดังกล่าวเป็นการแปลความหมายไปที่ตัวพารามิเตอร์ที่สนใจโดยตรง ที่ทำเช่นนี้ได้เพราะสถิติแบบเบส์นั้นสร้างโมเดลความน่าจะเป็นเพื่ออธิบายพารามิเตอร์ที่สนใจโดยตรง ในขณะที่เมื่อเปรียบเทียบกับการ

แปลความหมายของ $100(1-\alpha)\%$ ช่วงความเชื่อมั่นนั้น เป็นการยืนยันในเชิงความถี่สัมพัทธ์ของการสุ่มตัวอย่างว่าช่วงความเชื่อมั่นที่สร้างขึ้นมานั้นจะครอบคลุมค่าพารามิเตอร์ที่สนใจด้วยความน่าจะเป็นเท่ากับ $100(1-\alpha)\%$

การทดสอบสมมติฐาน (hypothesis testing)

สมมติว่านักสถิติต้องการที่จะทดสอบสมมติฐานทางสถิติโดยสมมติฐานว่าง (null hypothesis) คือ $H_0:\theta\in\Theta_0$ และสมมติฐานทางเลือก (alternative hypothesis) คือ $H_1:\theta\in\Theta_0^C$ เมื่อ Θ_0 เป็นเซตย่อยของปริภูมิพารามิเตอร์ (subset of the parameter space) และ Θ_0^C เป็นส่วนเติมเต็ม (complement) ของ Θ_0 จากการใช้การแจกแจงความน่าจะเป็นภายหลัง $p(\theta\,|\,y)$ จะทำให้สามารถคำนวณค่าความน่าจะเป็นภายหลัง $P(\theta\in\Theta_0\,|\,y)$ และ $P(\theta\in\Theta_0^C\,|\,y)$ ได้ ซึ่งก็คือความน่าจะเป็นที่สมมติฐานว่างและสมมติฐานทางเลือกจะเป็นจริงตามลำดับ วิธีการหนึ่งที่ใช้ในการทดสอบสมมติฐานแบบเบส์เซียนคือการเปรียบเทียบอัตราส่วนระหว่างความน่าจะเป็นภายหลัง $P(\theta\in\Theta_0\,|\,y)$ และ $P(\theta\in\Theta_0^C\,|\,y)$ ถ้า $P(\theta\in\Theta_0\,|\,y)\geq P(\theta\in\Theta_0^C\,|\,y)$ จะยอมรับสมมติฐานว่าง หรือนักสถิติอาจตั้งเกณฑ์ไว้ก่อนหน้าการทดสอบว่าจะยอมรับสมมติฐานว่างเมื่อความน่าจะเป็น $P(\theta\in\Theta_0\,|\,y)$ มีค่ามากกว่าเกณฑ์ที่ได้กำหนดไว้ เป็นต้น

ในการทดสอบสมมติฐานว่างแบบจุด (point null hypothesis) $H_0:\theta=\theta_0$ กับสมมติฐานทางเลือก $H_1:\theta\neq\theta_0$ จะต้องพิจารณาก่อนว่าการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์นั้นเป็นการแจกแจงแบบไม่ต่อเนื่อง (discrete) หรือต่อเนื่อง (continuous) โดยถ้าหากเป็นการแจกแจงแบบไม่ต่อเนื่องนั้นจะสามารถหาความน่าจะเป็นของสมมติฐานว่างได้โดยตรง ในขณะที่หากการแจกแจงความน่าจะเป็นภายหลังเป็นการแจกแจงแบบต่อเนื่องจะทำให้ความน่าจะเป็นของสมมติฐานว่างมีค่าเท่ากับ 0 ซึ่งจะไม่สามารถทำการทดสอบสมมติฐานได้ วิธีการหนึ่งที่ใช้แก้ไขปัญหาในกรณีที่ต้องทำการทดสอบสมมติฐานในกรณีนี้คือการปรับสมมติฐานว่างใหม่ให้เป็นสมมติฐานว่างแบบช่วงโดยปรับให้อยู่ในช่วงแคบๆดังนี้ $H_0:\theta\in\Theta_0=(\theta_0-a,\theta_0+a)$ เมื่อ a เป็นค่าคงที่ขนาดเล็ก ซึ่งจะทำให้สามารถทดสอบสมมติฐานต่อไปได้

ตัวสถิติที่สำคัญในการวิเคราะห์ข้อมูลแบบเบส์คือ Bayes factor (Berger, 1985; Kass and Raftery, 1995) ซึ่งเป็นสถิติที่ใช้ในการเปรียบเทียบโมเดล (model comparison) ซึ่งมีความยืดหยุ่นสามารถนำไปประยุกต์ใช้กับโมเดลต่างๆได้มากมาย หลักการของ Bayes factor สามารถอธิบายได้โดยสังเขปดังนี้

สมมติให้ $\underline{y}$ เป็นเวกเตอร์ของค่าสังเกตที่มีขนาดตัวอย่าง n หน่วย M_0 กับ M_1 เป็นโมเดล 2 โมเดลที่ผู้วิจัยต้องการทดสอบสมมติฐานเพื่อเปรียบเทียบโมเดลทั้งสองว่าโมเดลใดมีความเหมาะสมกับข้อมูลค่าสังเกต และ $p(M_0)$ เป็นความน่าจะเป็นก่อนหน้า (prior distribution) ที่โมเดล M_0 เป็นโมเดลที่ความเหมาะสมกับข้อมูลค่าสังเกต ดังนั้นจะได้ว่า $p(M_1) = 1 - p(M_0)$ และจากทฤษฎีของเบส์ในทฤษฎีความน่าจะเป็นจะได้ว่า ความน่าจะเป็นที่โมเดล M_k จะเป็นโมเดลที่ถูกต้องเมื่อกำหนดข้อมูลค่าสังเกต $\underline{y}$ สามารถคำนวณได้ดังนี้

$$p(M_k \mid \underline{y}) = \frac{p(\underline{y} \mid M_k)\,p(M_k)}{p(\underline{y} \mid M_1)\,p(M_1) + p(\underline{y} \mid M_0)\,p(M_0)} \quad \text{โดยที่ } k = 0,1 \qquad (2.70)$$

ดังนั้นจะได้ว่าอัตราส่วนความน่าจะเป็นภายหลังระหว่างความน่าจะเป็นที่โมเดล M_1 จะเป็นโมเดลที่เหมาะสมเมื่อกำหนดข้อมูลค่าสังเกต $\underline{y}$ เทียบกับความน่าจะเป็นที่โมเดล M_0 จะเป็นโมเดลที่เหมาะสมเมื่อกำหนดข้อมูลค่าสังเกต $\underline{y}$ คือ

$$\frac{p(M_1 \mid \underline{y})}{p(M_0 \mid \underline{y})} = \frac{p(\underline{y} \mid M_1)\,p(M_1)}{p(\underline{y} \mid M_0)\,p(M_0)} \qquad (2.71)$$

จากสมการที่ (2.71) จะนิยามว่า Bayes factor สำหรับการเปรียบเทียบโมเดล M_1 และ M_0 เขียนแทนด้วย B_{10} คือ

$$B_{10} = \frac{p(\underline{y} \mid M_1)}{p(\underline{y} \mid M_0)} \qquad (2.72)$$

จากสมการที่ (2.71) และ (2.72) จะได้ว่า

$$\frac{p(M_1 \mid \underline{y})}{p(M_0 \mid \underline{y})} = posterior\ \ odds = Bayes\ \ factor \times prior\ odds \qquad (2.73)$$

พิจารณาจากสมการที่ (2.73) สมมติว่ากำหนดให้ $p(M_0) = p(M_1) = 0.5$ จะได้ว่า Bayes factors จะมีค่าเท่ากับ posterior odds สมมติว่าผู้วิจัยต้องการเปรียบเทียบโมเดลโดยการทดสอบ

สมมติฐานโดยที่ สมมติฐานหลัก H_0 คือโมเดล M_0 เป็นโมเดลที่เหมาะสม เปรียบเทียบกับสมมติฐานทางเลือก H_1 คือโมเดล M_1 เป็นโมเดลที่เหมาะสม แนวทางการทดสอบสมมติฐานดังกล่าวมีความแตกต่างจากการทดสอบสมมติฐานเพื่อเปรียบเทียบโมเดลแบบดั้งเดิมด้วยตัวสถิติทดสอบไคสแควร์ กล่าวคือตัวสถิติทดสอบดังกล่าวสร้างจากวิธีการอัตราส่วนภาวะความควรจะเป็นซึ่งจะมีหลักการที่จะต้องสมมติให้สมมติฐานหลักที่ต้องการทดสอบเป็นจริงก่อน จากนั้นจึงพิจารณาเกณฑ์การปฏิเสธสมมติฐานหลักจากอัตราส่วนภาวะความควรจะเป็น ซึ่งการเปรียบเทียบโมเดลแบบเบส์นั้นไม่จำเป็นต้องมีการสมมติดังกล่าว นอกจากนี้เมื่อพิจารณาสมการที่ (2.73) จะเห็นว่าการเปรียบเทียบความเหมาะสมระหว่างสองโมเดลนั้นเป็นการเปรียบเทียบโดยใช้ข้อมูลค่าสังเกตชุดเดียวกันดังนั้นการที่ขนาดตัวอย่างมีขนาดใหญ่ขึ้นจะไม่ได้ทำให้โอกาสที่จะปฏิเสธสมมติฐานหลักนั้นมีเพิ่มมากขึ้นซึ่งเป็นข้อดีที่แตกต่างจากสถิติแบบดั้งเดิม นอกจากนี้แนวคิดดังกล่าวยังสามารถใช้กับการเปรียบเทียบโมเดลที่ไม่ได้ซ้อนกันได้อีกด้วย

เมื่อพิจารณาค่า B_{10} หากมีค่ามากกว่า 1 จะแปลความหมายได้ว่า M_1 เป็นโมเดลสอดคล้องกับข้อมูลค่าสังเกตมากกว่าโมเดล M_0 Harold Jeffreys ได้นำเสนอเกณฑ์การแปลผลในการทดสอบสมมติฐานเพื่อเปรียบเทียบโมเดลแสดงใน ตารางที่ 2.1

<u>ตารางที่ 2.1</u> เกณฑ์การแปลผลค่า Bayes factor

B_{10}	$2\log B_{10}$	**ระดับของหลักฐานในการปฏิเสธสมมติฐานหลัก** (strength of evidence)
<1	<0	คัดค้าน H_1 (สนับสนุน H_0)
$1-3$	$0-2$	สนับสนุน H_1 น้อยมาก
$3-20$	$2-6$	มั่นใจใน H_1
$20-150$	$6-10$	ยืนยัน H_1 มาก
>150	>10	สนับสนุน H_1 มากที่สุด

2.5 ข้อดีและข้อด้อยของสถิติแบบเบส์

สถิติแบบเบส์และสถิติแบบดั้งเดิมล้วนเป็นแนวทางที่มีทั้งข้อดีและข้อด้อยทั้งนั้น นอกจากนี้ยังมีบางส่วนที่มีความคล้ายคลึงกันอีกด้วย ในกรณีที่ขนาดตัวอย่างมีขนาดใหญ่จะพบว่าผลการวิเคราะห์ที่ได้จากวิธีการทางสถิติแบบเบส์นั้นจะมีความใกล้เคียงหรือเหมือนกับผลการวิเคราะห์ที่ได้จากวิธีการทางสถิติแบบดั้งเดิม เราอาจสรุปข้อดีของการใช้สถิติแบบเบส์ได้ดังต่อไปนี้ (Bolstad, 2004; Gelman, Carlin, Stern, and Rubin, 1995)

1. สถิติแบบเบส์เป็นแนวทางที่ยอมให้นักวิจัยหรือนักสถิติสามารถนำความรู้หรือความเชื่อเกี่ยวกับพารามิเตอร์ที่สนใจในงานวิจัย ไม่ว่าความรู้หรือความเชื่อนั้นจะได้จากประสบการณ์ของผู้วิจัยเองหรือเป็นความรู้ที่ได้จากงานวิจัยในอดีต มารวมกับสารสนเทศที่ได้จากข้อมูลเชิงประจักษ์เพื่อพัฒนาเป็นสารสนเทศซึ่งเก็บรวบรวมอยู่ในการแจกแจงความน่าจะเป็นภายหลัง นอกจากนี้นักสถิติยังสามารถนำการแจกแจงความน่าจะเป็นภายหลังที่ได้มาเป็นการแจกแจงความน่าจะเป็นก่อนหน้าในการวิจัยครั้งต่อไปได้อีก ซึ่งกระบวนการดังกล่าวจะทำให้สามารถพัฒนาสารสนเทศเกี่ยวกับพารามิเตอร์ที่สนใจได้อย่างสมบูรณ์มากยิ่งขึ้นเรื่อยๆ

2. การอนุมานเกี่ยวกับพารามิเตอร์ที่สนใจนั้นเป็นการอนุมานจากการแจกแจงของพารามิเตอร์โดยตรง ไม่ได้เป็นการอนุมานผ่านการแจกแจงของตัวสถิติ ดังนั้นจึงไม่จะเป็นที่จะต้องคำนึงถึงปัจจัยจากขนาดตัวอย่างเล็กซึ่งจะส่งผลต่อการแจกแจงของตัวสถิติเหมือนในการใช้สถิติแบบดั้งเดิม

3. สถิติแบบเบส์เป็นแนวทางที่สอดคล้องกับหลักภาวะหลักความควรจะเป็น (likelihood principle) กล่าวคือสมมติว่ามีค่าสังเกตสองชุดคือ $x_1, x_2, ..., x_n$ และ $y_1, y_2, ..., y_n$ โดยค่าสังเกตทั้งสองชุดนี้อาจได้มาจากการทดลองเดียวกันหรือได้มาจากสองการทดลองที่แตกต่างกัน ค่าสังเกตทั้งสองชุดนี้มีสมบัติในการกำหนดฟังก์ชันภาวะความควรจะเป็นเหมือนกันสำหรับพารามิเตอร์ θ หรือในการกำหนดฟังก์ชันภาวะความควรจะเป็นที่เป็นสัดส่วนซึ่งกันและกัน นั่นคือค่าสังเกตทั้งสองชุดจะให้สารสนเทศเกี่ยวกับพารามิเตอร์ที่ไม่ทราบค่า θ เหมือนกัน ในขณะที่สถิติแบบดั้งเดิม ไม่ได้เป็นไปตามหลักการดังกล่าวเสมอไป

4. สถิติแบบเบส์มีการแปลผลที่ตรงไปตรงมา เนื่องจากเป็นการอนุมานจากการแจกแจงของพารามิเตอร์ที่สนใจโดยตรง ยกตัวอย่างเช่น 95% ช่วงความน่าเชื่อถือ จะหมายความว่าความน่าจะเป็นที่ค่าพารามิเตอร์ที่แท้จริงจะอยู่ในช่วงนี้มีค่าเท่ากับ 0.95

5. หลักการของสถิติแบบเบส์ทำให้สามารถแก้ปัญหาในโมเดลที่มีความซับซ้อนได้ง่ายและมีความสะดวก เช่น โมเดลเชิงเส้นพหุระดับ (multi-level model) โมเดลเชิงเส้นพหุระดับที่มีโครงสร้างของความแปรปรวนซับซ้อน (multi-level model with complex variation) เป็นต้น

แต่อย่างไรก็ตามสถิติแบบเบส์ก็ยังมีข้อด้อยดังต่อไปนี้

1. ในการเลือกการแจกแจงความน่าจะเป็นก่อนหน้านั้นไม่ได้มีกฎเกณฑ์หรือหลักการที่แน่นอน ดังนั้นการใช้สถิติแบบเบส์นักสถิติที่ใช้จึงจะเป็นที่จะต้องมีทักษะในการแปลงความเชื่อก่อนหน้าเกี่ยวกับพารามิเตอร์ที่สนใจนั้น ในอยู่ในรูปแบบของฟังก์ชันทางคณิตศาสตร์ ซึ่งก็คือการแจกแจงความน่าจะเป็นก่อนหน้านั่นเอง หากนักสถิติไม่ได้มีความระมัดระวังต่อการเลือกการแจกแจงความน่าจะเป็นก่อนหน้าแล้ว ผลการวิเคราะห์ที่ได้นั้นมีโอกาสที่จะมีความผิดพลาดได้สูง

2. ในบางกรณีการแจกแจงความน่าจะเป็นภายหลังที่คำนวณได้นั้น อาจได้รับอิทธิพลจากการแจกแจงความน่าจะเป็นก่อนหน้ามากเกินไป ทำให้สารสนเทศที่ได้ภายหลังนั้นมีแต่ส่วนที่เป็นความเชื่อก่อนหน้าของผู้วิจัยโดยที่ไม่ได้พิจารณาข้อมูลเชิงประจักษ์หรือพิจารณาแต่เพียงเล็กน้อยเท่านั้น

3. ในกรณีที่โมเดลมีความซับซ้อนและจำเป็นที่จะต้องใช้เทคนิคการจำลองเพื่อประมาณการแจกแจงความน่าจะเป็นภายหลังนั้น ผลการวิเคราะห์ที่ได้อาจมีความแตกต่างกันในแต่ละครั้งที่ทำการวิเคราะห์ยกเว้นแต่ว่าจะใช้เลขสุ่มตัวแรก (random seed) ในการจำลองตัวเดียวกัน จะเห็นว่าข้อด้อยข้อนี้ไปขัดแย้งกับข้อดีที่กล่าวไว้ว่าวิธีการแบบเบส์เซียนนั้นเป็นไปตามหลักภาวะความควรจะเป็นสูงสุด ในความเป็นจริงแล้วหลักการดังกล่าวจะเป็นจริงได้ในกรณีที่การแจกแจงความน่าจะเป็นภายหลังนั้นได้มาจากวิธีการเชิงวิเคราะห์เท่านั้น หากได้มาจากวิธีการจำลองแล้วค่าประมาณของพารามิเตอร์ที่ได้จะมีความผันแปรไปตามชุดของเลขสุ่ม (random number) ที่ใช้ในแต่ละกระบวนการ

จากที่ได้กล่าวมาจะพบว่าสถิติแบบเบส์จะมีวิธีการอนุมานค่าพารามิเตอร์ที่แตกต่างจากสถิติแบบดั้งเดิมกล่าวคือการอนุมานเชิงสถิติทั้งหมดซึ่งได้แก่ การประมาณค่าแบบจุด การประมาณค่าแบบช่วง และการทดสอบสมมติฐาน สถิติแบบเบส์นั้นจะใช้การแจกแจงความน่าจะเป็นภายหลังซึ่งสามารถคำนวณได้จากการใช้ทฤษฎีของเบส์ในทฤษฎีความน่าจะเป็นเพื่อการอนุมานทั้งหมด อย่างไรก็ตามในทางปฏิบัติมีปัญหาจำนวนน้อยมากที่สามารถพิสูจน์หารูปแบบของการแจกแจงความน่าจะเป็นภายหลังให้อยู่ในรูปปิด (closed form) ได้โดยวิธีเชิงวิเคราะห์ (analytical methods)ในทางคณิตศาสตร์ เพราะว่าในปัญหาในทางปฏิบัติมีความซับซ้อนดังนั้นโมเดลที่ใช้ในการอธิบายปัญหาดังกล่าวย่อมต้องมีจำนวนพารามิเตอร์ในโมเดลจำนวนมาก การที่จะหาค่าอินทิกรัลของพจน์ $p(y)$ ซึ่งเป็นอินทิกรัลหลายชั้นตามจำนวนพารามิเตอร์ในโมเดลนั้นจึงกระทำได้ยากหรือไม่สามารถกระทำได้เลย ดังนั้นในปัญหาในทางปฏิบัติโดยส่วนใหญ่นักสถิติจึงเลือกใช้วิธีการประมาณการแจกแจงความน่าจะเป็นภายหลังโดยใช้เทคนิคการจำลอง (simulation) เข้ามาช่วย หลักการโดยทั่วไปของวิธีการนี้คือ หากสามารถจำลองค่าสังเกตจากการแจกแจงภายหลังของพารามิเตอร์ที่สนใจจะประมาณค่าได้จำนวนมากเพียงพอ เราจะสามารถประมาณค่าพารามิเตอร์ที่สนใจนั้น รวมทั้งฟังก์ชันต่างๆของพารามิเตอร์นั้นได้จากค่าสังเกตที่จำลองมาเหล่านั้นโดยตรง วิธีการที่เป็นรู้จักและใช้กันอย่างแพร่หลายคือวิธี Markov Chain Monte Carlo (MCMC)

2.6 <u>ลูกโซ่มาร์คอฟและลูกโซ่มาร์คอฟมอนติคาร์โล (Markov Chain and Markov Chain Monte Carlo Method: MCMC)</u>

ข้อจำกัดของสถิติแบบเบส์คือการคำนวณการแจกแจงความน่าจะเป็นภายหลัง เนื่องจากการแจกแจงความน่าจะเป็นภายหลังนั้นมักมีรูปแบบเป็นอินทิกรัลหลายชั้นเทียบกับพารามิเตอร์ในโมเดลซึ่งในทางปฏิบัตินั้นการอินทิเกรตพจน์ดังกล่าวเพื่อหาการแจกแจงความน่าจะเป็นภายหลังนั้นกระทำได้ยากหรือในบางกรณีอาจเป็นไปไม่ได้เลย วิธีการหนึ่งในการประมาณการแจก แจงความน่าจะเป็นภายหลังดังกล่าวคือการใช้ลูกโซ่มาร์คอฟมอนติคาร์โล ซึ่งเป็นวิธีการเชิงจำลองมีจุดมุ่งหมายที่จะสุ่มตัวอย่างพารามิเตอร์ที่สนใจจากการแจกแจงความน่าจะเป็นที่มีความซับซ้อน

ลูกโซ่มาร์คอฟ (Markov Chain)

วิธีลูกโซ่มาร์คอฟมอนติคาร์โลนั้นเป็นวิธีการเชิงจำลองซึ่งจะสุ่มตัวอย่างของพารามิเตอร์ที่สนใจโดยที่ค่าสังเกตที่สุ่มขึ้นมานั้นจะขึ้นอยู่กับค่าสังเกตก่อนหน้าเพียงค่าเดียวเท่านั้น เราสามารถนิยามระบบลูกโซ่มาร์คอฟได้ดังต่อไปนี้

กำหนดให้ θ_t เป็นค่าของตัวแปรสุ่ม Θ ณ เวลา t จะได้ว่า ปริภูมิสถานะ (state space) คือพิสัย (range) ของค่า Θ ทั้งหมดที่เป็นไปได้ "กระบวนการมาร์คอฟ (Markov process)" เป็นกระบวนการสุ่ม (stochastic process)[2] ประเภทหนึ่งที่ใช้ในการอธิบายความไม่แน่นอนหรือความน่าจะเป็นของการเปลี่ยนสถานะ (transition probability) จากสถานะหนึ่งไปยังอีกสถานะหนึ่ง ความน่าจะเป็นของการเปลี่ยนสถานะนั้นมีคุณสมบัติที่ความน่าจะเป็นในการที่ระบบจะเปลี่ยนไปยังสถานะต่อไปนั้นจะขึ้นกับสถานะในอดีตเพียงสถานะเดียวเท่านั้น ซึ่งสามารถเขียนเป็นสัญลักษณ์ทางคณิตศาสตร์ได้ดังนี้

$$P(\theta_{n+1} = i_{n+1} \mid \theta_n = i_n, \theta_{n-1} = i_{n-1}, ..., \theta_1 = i_1) = P(\theta_{n+1} = i_{n+1} \mid \theta_n = i_n) \qquad (2.74)$$

สำหรับทุกจำนวนนับ n ซึ่ง $n+1 \in T$

"ลูกโซ่มาร์คอฟ (Markov chain)" เป็นลำดับของตัวแปรสุ่ม $\theta_1, \theta_2, \theta_3, ..$ ที่เกิดขึ้นจากกระบวนการมาร์คอฟในข้างต้น ลักษณะของแต่ละลูกโซ่มาร์คอฟแต่ละลูกโซ่นั้นจะมีลักษณะที่แตกต่างกันขึ้นกับปริภูมิสถานะและความน่าจะเป็นของการเปลี่ยนแปลงสถานะ เขียนแทนด้วย $P(i, j) = P(i \rightarrow j)$ ซึ่งมีความหมายถึงความน่าจะเป็นที่กระบวนการจะเปลี่ยนจากสถานะจากสถานะ θ_i ไปยังสถานะ θ_j ภายใน 1 ขั้นกล่าวคือ

$$P(i, j) = P(i \rightarrow j) = P(\theta_{t+1} = j \mid \theta_t = i) \qquad (2.75)$$

หากระบบมาร์คอฟใดมีความน่าจะเป็นในการเปลี่ยนแปลงสถานะจาก θ_i ไปยังสถานะ θ_j ที่คงที่ในทุกช่วงเวลาจะเรียกว่า ระบบมาร์คอฟนั้นเป็นลูกโซ่มาร์คอฟแบบเอกพันธ์ (homogeneous Markov Chain) กล่าวคือเป็นระบบลูกโซ่มาร์คอฟซึ่ง

$$P(\theta_{t+1} = j \mid \theta_t = i) = P(\theta_2 = j \mid \theta_1 = i) \qquad \forall i, j \in S \qquad (2.76)$$

---

[2] กระบวนการสุ่ม (stochastic process) คือระบบที่ตัวแปรสุ่มแปรผันค่าไปตามเวลา

ในกรณีที่ต้องการพิจารณาระบบมาร์คอฟเมื่อมีการเปลี่ยนแปลงมากกว่า 1 ขั้นตอน จะสามารถพิจารณาได้จากความน่าจะเป็นในการเปลี่ยนแปลงสถานะ t ขั้นตอน กำหนดให้ความน่าจะเป็นที่ระบบมาร์คอฟจะเปลี่ยนแปลงจากสถานะ i ไปยังสถานะ j ใน t ขั้นตอน คือ

$$p_{ij}(t) = P(\theta_t = j \mid \theta_1 = i) \tag{2.77}$$

การหาค่าความน่าจะเป็นในเปลี่ยนแปลงสถานะ t ขั้นตอน สามารถกระทำได้โดยการใช้ทฤษฎี Chapman-Kolomogorov ช่วยในการหาความน่าจะเป็นดังกล่าว ดังนี้ ความน่าจะเป็นที่ระบบมาร์คอฟจะเปลี่ยนแปลงสถานะมาที่สถานะ j จากสถานะเริ่มต้นใดๆสามารถหาได้จาก

$$\begin{aligned}\pi_j(t+1) &= P(\theta_{t+1} = j) \\ &= \sum_k P(\theta_{t+1} = j \mid \theta_t = k) \cdot P(\theta_t = k) \\ &= \sum_k P(k \to j) \cdot \pi_k(t)\end{aligned} \tag{2.78}$$

จากสมการที่ (2.78) จะเห็นว่าการใช้ทฤษฎีดังกล่าวอาจกระทำได้ยาก ซึ่งสามารถทำให้ง่ายขึ้นได้โดยการเขียนสมการ Chapman-Kolomogorov ในรูปของเมทริกซ์ได้ดังนี้

$$\pi(t+1) = \pi(t) \cdot P \tag{2.79}$$

เมื่อกำหนดให้ เมทริกซ์ P เป็นเมทริกซ์ของความน่าจะเป็นในการเปลี่ยนแปลงสถานะ (probability transition matrix) ซึ่งเป็นเมทริกซ์ที่มีสมาชิกเป็น $P(i, j) = P(i \to j)$ ซึ่งจะได้ว่าผลรวมของสมาชิกในแต่ละแถวจะมีค่าเท่ากับ 1 กล่าวคือ $\sum_i P(i, j) = \sum_i P(i \to j) = 1$

จากสมการที่ (2.79) จะได้ว่า

$$\pi(t+1) = \pi(t) \cdot P = \pi(t-1) \cdot P^2 = \pi(t-2) \cdot P^3 = \dots = \pi(0) \cdot P^t \tag{2.80}$$

จากทฤษฎีของ Chapman-Kolomogorov ในข้างต้นจะได้ว่า หากทราบเมทริกซ์ของความน่าจะเป็นของการเปลี่ยนแปลงสถานะ P เราจะสามารถหาความน่าจะเป็นของการเปลี่ยนสถานะ t ขั้น และเมทริกซ์ของการเปลี่ยนแปลงสถานะ t ขั้นได้เสมอ ซึ่งจากเมทริกซ์ของความน่าจะเป็นของการเปลี่ยนแปลงสถานะในข้างต้นจะได้ว่า ความน่าจะเป็นในการเปลี่ยนแปลงสถานะ n ขั้นจากสถานะ i ไปยังสถานะ j เขียนแทนด้วย $p_{ij}^{(n)} = P(\theta_{t+n} = j \mid \theta_t = i) \in P^n$

นอกจากนี้ยังมีคุณสมบัติที่สำคัญของลูกโซ่มาร์คอฟซึ่งมีความจำเป็นต่อเทคนิคลูกโซ่มาร์คอฟ มอนติคาร์โล คุณสมบัติแรกคือ "คุณสมบัติลดทอนไม่ได้ (irreducicible)" ลูกโซ่มาร์คอฟจะมีคุณสมบัติลดทอนไม่ได้ ถ้า $\exists n \in Z$ ซึ่ง $p_{ij}^{(n)} > 0 \;\; \forall i, j$ กล่าวคือจากแต่ละสถานะในปริภูมิสถานะนั้นสามารถเปลี่ยนแปลงไปยังสถานะทุกสถานะภายในปริภูมิสถานะได้เสมอ หรืออาจจะกล่าวว่าปริภูมิสถานะไม่สามารถลดทอนได้นั่นเอง คุณสมบัติที่สองคือ "คุณสมบัติไม่เป็นคาบ (aperiodic)" ลูกโซ่มาร์คอฟจะมีคุณสมบัติไม่เป็นคาบเมื่อ จำนวนขั้นตอนในการเปลี่ยนแปลงสถานะระหว่าง 2 สถานะไม่จำเป็นต้องเป็นพหุคูณของจำนวนเต็มใดๆ กล่าวคือต้องไม่มีวัฏจักร (cycle) เกิดขึ้นในระบบของลูกโซ่มาร์คอฟ

ในระบบลูกโซ่มาร์คอฟเราจะพบว่าการแจกแจงความน่าจะเป็นของสถานะจะมีการเปลี่ยนแปลงไปตามระยะเวลา อย่างไรก็ตามในบางระบบเราจะพบว่าเมื่อระยะเวลาในการดำเนินระบบมีมากเพียงพอการแจกแจงความน่าจะเป็นของสถานะจะลู่เข้าไปสู่การแจกแจงคงที่ ซึ่งเรียกว่า การแจกแจงสถานะอยู่ตัว (stationary distribution) กล่าวคือเป็นความน่าจะเป็นของการเปลี่ยนแปลงสถานะที่เป็นอิสระจากการแจกแจงความน่าจะเป็นของการเปลี่ยนแปลงสถานะเริ่มต้น (initial) ดังนี้ สมมติว่า π^* เป็นการแจกแจงสถานะอยู่ตัวจะได้ว่า $\pi^* = \pi^* \cdot P$ ในกรณีที่ลูกโซ่มาร์คอฟมีลักษณะเป็นคาบนั้นจะได้ว่าระบบลูกโซ่จะมีลักษณะการเปลี่ยนแปลงระหว่างสถานะภายในปริภูมิสถานะที่เป็นวัฏจักรตายตัว ซึ่งการแจกแจงความน่าจะเป็นในการเปลี่ยนแปลงสถานะของระบบลูกโซ่ดังกล่าวนี้จะไม่สามารถลู่เข้าไปสู่การแจกแจงสถานะอยู่ตัวได้

เทคนิคลูกโซ่มาร์คอฟมอนติคาร์โล (MCMC) นั้นเป็นวิธีการจำลองลูกโซ่มาร์คอฟ[3] ที่มีการออกแบบให้ลูกโซ่มาร์คอฟที่จำลองนั้นมีคุณสมบัติที่ดีบางประการซึ่งจะทำให้ลูกโซ่มาร์คอฟที่สร้างขึ้นมานั้นจะลู่เข้าไปสู่การแจกแจงที่เรียกว่า การแจกแจงสถานะอยู่ตัว (stationary distribution) จากข้อดีของเทคนิค MCMC ดังกล่าวนี้จึงได้มีการประยุกต์ใช้เทคนิค MCMC ในการ

---

[3] กำหนดให้ $\theta_1, \theta_2, \theta_3, \ldots$ เป็นลำดับของตัวแปรสุ่มที่มีปริภูมิพารามิเตอร์ T เราจะกล่าวว่าลำดับดังกล่าวมีคุณสมบัติของลูกโซ่มาร์คอฟ ก็ต่อเมื่อสำหรับทุกจำนวนนับ n ซึ่ง $n+1 \in T$

$$P(\theta_{n+1} = i_{n+1} \mid \theta_n = i_n, \theta_{n-1} = i_{n-1}, \ldots, \theta_1 = i_1) = P(\theta_{n+1} = i_{n+1} \mid \theta_n = i_n)$$

เมื่อ $i_1, i_2, \ldots, \in S$ ค่าที่เป็นไปได้ของลูกโซ่มาร์คอฟจะเรียกว่าสถานะ (state) และค่าที่เป็นไปได้ทั้งหมดของลูกโซ่มาร์คอฟจะถูกเก็บไว้ในเซต S ที่เรียกว่าปริภูมิสถานะ (state space)

หาการแจกแจงความน่าจะเป็นภายหลัง ในกรณีที่ไม่สามารถใช้วิธีเชิงวิเคราะห์หาได้ดังที่ได้กล่าวไว้ในหัวข้อที่แล้ว (Albert, 2009; Gamerman and Lopes, 2006; Tanner, 1996; Gelman, Carlin, Stern, and Rubin, 1995)

MCMC เป็นเทคนิคการจำลองเทคนิคหนึ่งที่นิยมใช้กันมาในการจำลองตัวอย่างจากการแจกแจงความน่าจะเป็นภายหลังเพื่อประมาณการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์เมื่อไม่สามารถใช้วิธีการเชิงวิเคราะห์ได้ เนื่องจากเป็นเทคนิคที่ออกแบบการจำลองโดยการสุ่มตัวอย่างจากลูกโซ่มาร์คอฟที่มีคุณสมบัติลดทอนไม่ได้และไม่เป็นคาบซึ่งทำให้มั่นใจได้ว่าในสุดท้ายหากจำนวนรอบที่ทำการสุ่มตัวอย่างมีมากเพียงพอการแจกแจงสถานะของลูกโซ่มาร์คอฟดังกล่าวลู่เข้าไปสู่การแจกแจงสถานะอยู่ตัว ซึ่งจะเทียบเท่ากับการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่สนใจในมุมมองของสถิติแบบเบส์ (Albert, 2009; Gamerman and Lopes, 2006; Gelman, Browne, 1998; Carlin, Stern, and Rubin, 1995) ดังนั้นหากเราดำเนินการจำลองด้วยวิธี MCMC ในจำนวนรอบที่มากเพียงพอแล้วเราจะสามารถมั่นใจได้ว่าตัวอย่างที่จำลองขึ้นมานั้นจะมีการแจกแจงใกล้เคียงหรือเหมือนกับการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่สนใจ วิธีการหรืออัลกอริทึมที่ใช้ในการสร้างลูกโซ่มาร์คอฟนั้นมีอยู่ด้วยกันหลายวิธี วิธีการที่เป็นที่นิยมได้แก่ วิธี Metropolis sampling วิธี Metropolis-Hasting sampling และวิธี Gibbs sampling รายละเอียดของวิธีต่างๆเป็นดังนี้

วิธี Metropolis sampling

อัลกอริทึม Metropolis (Metropolis algorithm) นี้ถูกพัฒนาโดย Metropolis และคณะ (Metropolis et al.,1953) วิธีนี้เป็นวิธีที่กระทำได้ง่ายและมีประสิทธิภาพในการจำลองตัวอย่างสุ่มจากการแจกแจงความน่าจะเป็นที่ต้องการถึงแม้ว่าการแจกแจงดังกล่าวจะมีความซับซ้อน สมมติว่าจุดประสงค์ต้องการสุ่มตัวอย่างจากการแจกแจงความน่าจะเป็นที่เป็นเป้าหมาย $p(\theta)$ ซึ่ง $p(\theta) = f(\theta)/K$ โดยที่ K คือค่าคงที่ เรียกว่า normalized constant ที่ไม่ทราบค่า และไม่สามารถคำนวณค่าได้โดยตรง อัลกอริทึม Metropolis จะสุ่มลำดับของตัวแปรสุ่มจากการแจกแจงความน่าจะเป็นที่มีรายละเอียดดังต่อไปนี้

1. กำหนดให้ค่าเริ่มต้น θ_0 ใดๆ ซึ่ง $f(\theta_0) > 0$

2. ทำการสุ่มตัวอย่าง θ^* จากการแจกแจงความน่าจะเป็นโครงร่าง $q(\theta^* \mid \theta_{t-1})$ ซึ่งมีเงื่อนไขกำหนดคือต้องมีคุณสมบัติสมมาตร กล่าวคือ $q(\theta_1 \mid \theta_2) = q(\theta_2 \mid \theta_1)$

3. คำนวณค่าอัตราส่วนความน่าจะเป็นภายหลังซึ่งเป็นเกณฑ์ที่ใช้ในการคัดเลือกตัวอย่างสุ่มดังนี้

$$\alpha_t = \min\left\{ \frac{f(\theta^*)}{f(\theta_{t-1})}, 1 \right\} \tag{2.81}$$

4. สุ่มตัวอย่าง u จากการแจกแจงสม่ำเสมอ $U \sim (0,1)$

5. ถ้า $u < \alpha_t$ จะกำหนดให้ $\theta_t = \theta^*$ นอกจากนั้นจะกำหนดให้ $\theta_t = \theta_{t-1}$

6. เริ่มวนซ้ำใหม่ตั้งแต่ขั้นตอนที่ 2 จนกว่าจำนวนรอบของการวนซ้ำจะครบตามจำนวนที่ต้องการ

ขั้นตอนที่ 3 ถึง 5 ของอัลกอริทึมเป็นขั้นตอนที่ใช้ในการคัดกรองตัวอย่างสุ่มที่สุ่มได้จากการแจกแจงความน่าจะเป็นที่ใช้เป็นโครงร่าง โดยจะยอมรับตัวอย่างสุ่มที่ดำเนินการสุ่มมาด้วยความน่าจะเป็นเท่ากับ α_t โดยอาจเรียก α_t ความน่าจะเป็นในการเคลื่อนที่หรือเปลี่ยนแปลงสถานะ อัลกอริทึมดังกล่าวจะสร้างลูกโซ่มาร์คอฟ $(\theta_0, \theta_1, ..., \theta_k, ...)$ ซึ่งมีความน่าจะเป็นในการเปลี่ยนแปลงสถานะจาก θ_t ไปยัง θ_{t+1} จะขึ้นอยู่กับสถานะ θ_t เท่านั้น ไม่ได้ขึ้นกับ $\theta_0, \theta_1, ..., \theta_{t-1}$ จากการสร้างลำดับของลูกโซ่มาร์คอฟด้วยวิธีการในข้างต้น หากมีจำนวนการวนรอบและมีการพิจารณาตัดลูกโซ่ช่วงแรก (burn-in) ที่เพียงพอ จะได้ว่าตัวอย่างสุ่มที่เหลือจากการสุ่มจากระบบลูกโซ่มาร์คอฟดังกล่าวจะมีการแจกแจงเข้าไปสู่การแจกแจงสถานะอยู่ตัวซึ่งเท่ากับการแจกแจงความน่าจะเป็น $p(\theta)$ ที่ต้องการนั่นเอง

วิธี Metropolis-Hastings sampling

จากข้อกำหนดเบื้องต้นของวิธี Metropolis sampling ที่กำหนดให้การแจกแจงความน่าจะเป็นโครงร่างจะต้องมีคุณสมบัติสมมาตรซึ่งไม่ค่อยสะดวกในการนำอัลกอริทึมดังกล่าวมาประยุกต์ใช้จริง ดังนั้นจึงได้มีการพัฒนาอัลกอริทึมใหม่โดยอาศัยพื้นฐานเดิมของวิธี Metropolis sampling แต่ขยายให้มีความเป็นกรณีทั่วไปมากขึ้น คือยอมให้การแจกแจงความน่าจะเป็นโครง

ร่างมีลักษณะที่ไม่สมมาตรได้ (Hastings, 1970) โดยมีการปรับอัตราส่วนความน่าจะเป็นภายหลัง α_t เป็นดังนี้

$$\alpha_t = \min\left\{\frac{f(\theta^*)/q(\theta_{t-1}\mid\theta^*)}{f(\theta_{t-1})/q(\theta^*\mid\theta_{t-1})},1\right\} \tag{2.82}$$

ดังนั้นจะเห็นว่าวิธี Metropolis sampling เป็นกรณีเฉพาะกรณีหนึ่งของวิธี Metropolis-Hastings sampling เราสามารถพิสูจน์ได้ว่าวิธี Metropolis sampling และ Metropolis-Hasting sampling เป็นวิธีการในการสร้างลูกโซ่มาร์คอฟซึ่งจะลู่เข้าสู่การแจกแจงความน่าจะเป็นที่เป็นเป้าหมายได้ นอกจากนี้วิธี Metropolis-Hastings sampling นอกจากจะทำให้สามารถขยายกรอบการใช้งานออกไปได้แล้ว ยังช่วยในการเพิ่มอัตราการลู่เข้าของ Markov Chain ได้อีกด้วย ผู้สนใจสามารถศึกษาบทพิสูจน์และทฤษฎีต่างๆในรายละเอียดได้ในเอกสารอ้างอิงที่ได้ระบุไว้ (Walsh, 2004; Browne, 1998; Gelman, Carlin, Stern, and Rubin, 1995)

วิธี Gibbs sampling

เทคนิค Gibbs sampling (Geman and Geman (1984)) เป็นเทคนิคที่พัฒนาขึ้นโดยเป็นกรณีเฉพาะของวิธี Metropolis-Hastings sampling กล่าวคือเป็นวิธี Metropolis-Hasting sampling ที่มีการแจกแจงโครงร่างเป็นการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขภายหลัง (conditional posterior probability distribution) ดังนั้นวิธีการนี้จึงเป็นวิธีที่มีประสิทธิภาพสูงในด้านความเร็วเนื่องจากจะไม่มีการทิ้งตัวอย่างสุ่มที่สุ่มได้จากในแต่ละรอบเลย ในการดำเนินการใช้เทคนิค Gibbs sampling นี้ ขั้นแรกจะต้องพิจารณาแบ่งส่วนการแจกแจกความน่าจะเป็นภายหลังของพารามิเตอร์ออกเป็นการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ที่สนใจเมื่อกำหนดให้พารามเตอร์ตัวอื่นที่เหลือคงที่ ดังนั้นวิธีการนี้จะมีประสิทธิภาพสูงมากหากพารามิเตอร์ในโมเดลที่สนใจนั้นไม่มีความสัมพันธ์กัน และรูปแบบของการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขนั้นสามารถสุ่มตัวอย่างได้ง่าย เราอาจเขียนวิธีการของเทคนิค Gibbs sampling ให้เป็นทางการได้ดังต่อไปนี้

สมมติให้ $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_K)^T$ คือเวกเตอร์ของพารามิเตอร์ที่สนใจในโมเดล โดยที่พารามิเตอร์ θ_k ; $t = 1,2,\dots,K$ อาจเป็นสเกลาร์ (scalar) เวกเตอร์ หรือเมทริกซ์ก็ได้ $p\left(\underline{y}|\underline{\theta}\right)$ คือฟังก์ชันภาวะความควรจะเป็น (*likelihood function)* และ $p(\underline{\theta})$ คือการแจกแจงความน่าจะเป็นก่อนหน้าของเวกเตอร์พารามิเตอร์ ในกรณีที่การสุ่มตัวอย่างพารามิเตอร์จากการแจกแจงความน่าจะเป็นภายหลังร่วม $p\left(\underline{\theta}|\underline{y}\right)$ กระทำได้ยาก ผู้วิเคราะห์อาจสุ่มตัวอย่างพารามิเตอร์ θ_t จากการแจกแจงความน่าจะเป็นส่วนริม (marginal distribution) ของพารามิเตอร์โดยใช้คุณสมบัติความเป็นอิสระดังนี้ $p\left(\underline{\theta}|\underline{y}\right) = p\left(\theta_1|\underline{y}\right)p\left(\theta_2|\underline{y}\right)\dots p\left(\theta_K|\underline{y}\right)$ ในกรณีที่การสุ่มตัวอย่างจากการแจกแจงความน่าจะเป็นส่วนริมในข้างต้นกระทำได้ยาก แต่การแจกแจงความน่าจะเป็นแบบมีเงื่อนไข (conditional distribution) ของพารามิเตอร์ θ_k เมื่อกำหนดพารามิเตอร์ $\theta_{k'}$ และข้อมูลค่าสังเกต $\underline{y}$ เขียนแทนด้วย $p\left(\theta_k|\theta_1, \theta_2, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_K, \underline{y}\right)$สามารถจัดรูปให้อยู่ในรูปแบบมาตรฐานได้ อัลกอริทึมการสุ่มตัวอย่างแบบกิบส์จะมีความเหมาะสำหรับการแก้ไขปัญหาดังกล่าว โดยใช้คุณสมบัติของทฤษฎีความน่าจะเป็น

$$p\left(\theta_k|\theta_1, \theta_2, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_K, \underline{y}\right) = \frac{p\left(\theta_1, \theta_2, \dots, \theta_K|\underline{y}\right)}{p\left(\theta_1, \theta_2, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_K|\underline{y}\right)} \propto p\left(\theta_1, \theta_2, \dots, \theta_K|\underline{y}\right) \qquad (2.83)$$

จากสมการที่ (2.83) จะเห็นว่า $p\left(\theta_k|\theta_1, \theta_2, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_K, \underline{y}\right)$ เป็นสัดส่วนโดยตรงกับ $p\left(\underline{\theta}|\underline{y}\right)$ อัลกอริทึมของวิธี Gibbs sampling เป็นดังนี้

1. กำหนดให้ $n = 0$ แทนรอบของการสุ่ม เลือกค่าเริ่มต้นของ $\underline{\theta}$ ก่อน ใช้สัญลักษณ์ $\underline{\theta}^0$
2. ใช้ค่าเริ่มต้นของ $\underline{\theta}^0$ จำลองค่าของพารามิเตอร์ θ_i แต่ละตัวจากการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของ θ_i ดังนี้
   - สุ่ม θ_1^{t+1} จาก $p(\theta_1 | \theta_2^t, \dots, \theta_k^t, y)$
   - สุ่ม θ_2^{t+1} จาก $p(\theta_2 | \theta_1^{t+1}, \theta_3^t \dots, \theta_k^t, y)$
   - …
   - สุ่ม θ_k^{t+1} จาก $p(\theta_k | \theta_1^{t+1}, \dots, \theta_{k-1}^{t+1}, y)$
3. ถ้า $n < N$ จะกำหนดให้ $n = n + 1$ แล้วเริ่มวนซ้ำใหม่ตั้งแต่ขั้นตอนที่ 2

2.7 สิ่งสำคัญที่จำเป็นต้องพิจารณาในการใช้เทคนิค MCMC

การใช้เทคนิค MCMC นั้น ต้องมีการพิจารณาในรายละเอียดของการสร้างลูกโซ่ในหลายประเด็น เพื่อให้ผลที่ได้นั้นมีความถูกต้องและน่าเชื่อถือ สิ่งที่ผู้ใช้เทคนิค MCMC ควรที่จะต้องพิจารณานั้นมีดังต่อไปนี้ (Hara, and Sillanpää, 2009; Draper, 2005; Browne, 1998)

ค่าเริ่มต้น (initial value)

ในการสร้างลูกโซ่มาร์คอฟนั้นสิ่งแรกที่จะต้องกำหนดคือ “การกำหนดค่าเริ่มต้น (initial value)” ของลูกโซ่ การกำหนดค่าเริ่มต้นที่ดีนั้นจะทำให้ลูกโซ่มาร์คอฟที่สร้างขึ้นนั้นสามารถลู่เข้าสู่การแจกแจงสถานะอยู่ตัวได้อย่างรวดเร็ว แนวคิดหนึ่งในการกำหนดค่าเริ่มต้นที่สมเหตุสมผลคือการเลือกค่าเริ่มต้นให้ใกล้เคียงกับจุดศูนย์กลางของการแจกแจงความน่าจะเป็นภายหลังที่ต้องการให้มากที่สุด โดยมากนักสถิติมักเลือกใช้ค่าประมาณพารามิเตอร์ที่ได้จากวิธีประมาณแบบดั้งเดิม เช่น วิธีกำลังสองน้อยสุด หรือวิธีภาวะความควรจะเป็นสูงสุดเป็นค่าเริ่มต้น

Burn-in period

ลูกโซ่มาร์คอฟที่สร้างขึ้นนั้นจำเป็นที่จะต้องใช้ระยะเวลาช่วงหนึ่งในเพื่อที่จะเข้าสู่สถานะอยู่ตัว จะเห็นได้ว่าค่าสังเกตในช่วงแรกของการสร้างลูกโซ่นั้นจะยังไม่ได้เป็นค่าสังเกตที่ถูกสุ่มมาจากการแจกแจงความน่าจะเป็นภายหลังที่ต้องการ ดังนั้นในการใช้เทคนิค MCMC เพื่อประมาณการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่สนใจนั้นจึงจำเป็นที่จะต้องมีการตัดค่าสังเกตในช่วงแรกทิ้งเพื่อไม่ให้ค่าสังเกตในช่วงที่ลูกโซ่ยังไม่เข้าสู่สถานะอยู่ตัวนั้นมีอิทธิพลต่อการประมาณการแจกแจงความน่าจะเป็นภายหลัง ค่าสังเกตที่ตัดทิ้งไปนั้นจะเรียกว่า Burn-in โดยทั่วไปในการสร้างลูกโซ่นักสถิติมักจะกำหนดให้ burn-in มีจำนวนประมาณ 1000 ถึง 5000 ค่าสังเกต อย่างไรก็ตามในการใช้เทคนิค MCMC นั้นควรที่จะต้องมีการประเมินการลู่เข้าของลูกโซ่ด้วย เพื่อใช้เป็นข้อมูลในการกำหนด Burn-in ที่เหมาะสม เครื่องมือที่ใช้ในการพิจารณาจำนวน burn-in ที่เหมาะสมนั้นอาจใช้ Trace plot หรือ Cumulative estimates of the quantiles plot ช่วยในการพิจารณา ซึ่งรายละเอียดของกราฟทั้งสองจะกล่าวถึงในหัวข้อการวินิจฉัยการลู่เข้าของลูกโซ่

การลู่เข้าของลูกโซ่ (convergence of chain)

การใช้เทคนิค MCMC นั้นหากลูกโซ่ที่สร้างขึ้นมาสามารถกระจายไปได้ทั่วปริภูมิของพารามิเตอร์ที่สนใจนั้นได้ในช่วงระยะเวลาหนึ่ง จะเรียกว่า "ลูกโซ่ดังกล่าวผสมผสานกันได้ดี (mixing well)" ในทางกลับกันหากลูกโซ่ที่สร้างขึ้นนั้นไม่สามารถกระจายไปทั่วปริภูมิของพารามิเตอร์ได้หรือกระจายได้ทั่วในอัตราที่ช้ามาก จะเรียกว่า "ลูกโซ่ดังกล่าวผสมผสานกันไม่ดี (mixing poorly)" ซึ่งเหตุการณ์ดังกล่าวสามารถเกิดขึ้นในกรณีที่การแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่สนใจมีลักษณะเป็นหลายฐานนิยม[4] (multi-modal) การกำหนดค่าเริ่มต้นที่ไม่ดีอาจทำให้ลูกโซ่ที่สร้างขึ้นมาติดอยู่บริเวณฐานนิยมใดฐานนิยมหนึ่ง ซึ่งหากกำหนดจำนวนรอบในการสร้างลูกโซ่ไม่เพียงพอก็จะทำให้การแจกแจงความน่าจะเป็นภายหลังที่ประมาณได้จากตัวอย่างนั้นไม่ใชการแจกแจงที่ต้องการอย่างแท้จริงซึ่งจะส่งผลเสียต่อการอนุมานเกี่ยวกับพารามิเตอร์นั้นต่อไป จากปัญหาดังกล่าวจึงได้มีการพัฒนาวิธีการเพื่อประเมินการลู่เข้าของลูกโซ่ วิธีการหนึ่งที่เป็นที่นิยมใช้กันคือ วิธี Multiple highly dispersed initial values (Gelman and Rubin, 1992) ซึ่งเป็นวิธีการที่สามารถใช้ได้ง่ายในทางปฏิบัติ มีหลักการคือ 1) สร้างลูกโซ่มาร์คอฟหลายๆ ลูกโซ่โดยแต่ละลูกโซ่จะกำหนดค่าเริ่มต้นให้มีความแตกต่างกันและมีความหลากหลาย 2) พิจารณาว่าแต่ละลูกโซ่นั้นลู่เข้าสู่ฐานนิยมตัวเดียวกันหรือไม่ หากแต่ละลูกโซ่มีความสอดคล้องกันนั้น แสดงว่าสามารถมั่นใจในระดับหนึ่งได้ว่าลูกโซ่ที่สร้างขึ้นมานั้นน่าจะลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังที่ต้องการได้ อย่างไรก็ตามในความเป็นจริงแล้วผู้วิเคราะห์จะไม่ได้ทางที่จะสามารถตรวจสอบและยืนยันได้ว่าลูกโซ่มาร์คอฟที่สร้างนั้นสามารถลู่เข้าไปสู่การแจกแจงสถานะอยู่ตัวที่ต้องการแล้ว ทำได้แต่เพียงตรวจสอบว่าลูกโซ่ดังกล่าวยังไม่ลู่เข้าสู่การแจกแจงสถานะคงตัว เครื่องมือที่ใช้ในการตรวจสอบการลู่เข้าอาจแบ่งออกได้เป็น 2 แนวทางคือ แนวทางการใช้กราฟเพื่อตรวจสอบ และแนวทางการใช้ตัวสถิติเพื่อตรวจสอบ ซึ่งรายละเอียดของวิธีการต่างๆจะกล่าวถึงในหัวข้อการวินิจฉัยการลู่เข้าของลูกโซ่

---

[4] การแจกแจงความน่าจะเป็นภายหลังแบบหลายฐานนิยมสามารถเกิดขึ้นได้ในกรณีที่นักสถิติกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าที่มีสารสนเทศขัดแย้งกับสารสนเทศที่ได้จากข้อมูลจริงมาก

อัตสหสัมพันธ์ (Autocorrelation)

ค่าสังเกตที่สุ่มได้จากการใช้เทคนิค MCMC นั้น เนื่องจากเป็นการสุ่มจากระบบลูกโซ่มาร์คอฟดังนั้นค่าสังเกตจึงมีความสัมพันธ์กันเองตามรอบที่ทำการจำลอง ค่าสังเกตที่สร้างจากระบบลูกโซ่มาร์คอฟนั้นจึงเกิดปัญหาในการวิเคราะห์ที่เรียกว่า “อัตตสหสัมพันธ์” (Autocorrelation) ขึ้น จากปรากฏการณ์ดังกล่าวค่าสังเกตของการแจกแจงความน่าจะเป็นภายหลังที่สร้างขึ้นจากเทคนิคลูกโซ่มาร์คอฟจึงมีระบบที่สามารถอธิบายได้ด้วยโมเดลอนุกรมเวลา (time-series model) เรียกว่า โมเดลออโตรีเกรตซีฟลำดับที่ k (k th order autoregressive model (AR(p)) ซึ่งมีโมเดลทั่วไปดังนี้ สมมติว่า ว่า $\theta_1, \theta_2, ..., \theta_n$ เป็นลำดับของค่าสังเกตที่สุ่มได้ในรอบที่ i จากทั้งหมด n รอบของการใช้เทคนิค MCMC จะได้ว่าโมเดล AR(p) คือ

$$\theta_t = \alpha_1\theta_{t-1} + \alpha_2\theta_{t-2} + ... + \alpha_p\theta_{t-p} + \varepsilon_t \quad \text{เมื่อ } \varepsilon_t \sim N(0, \sigma^2) \qquad (2.84)$$

โมเดลดังกล่าวเป็นโมเดลที่มีลักษณะเหมือนกับโมเดลการถดถอย (regression model) แต่มีความแตกต่างกันคือในโมเดล AR(p) ค่าสังเกต ณ ช่วงเวลา t จะสามารถทำนายได้จากค่าสังเกต ณ ช่วงเวลาก่อนหน้า $t-p$ โดยที่ขนาดของอิทธิพลในแต่ละช่วงเวลาจะระบุโดยค่าสัมประสิทธิ์ความถดถอย α_i

อัตตสหสัมพันธ์ลำดับที่ k (k th order autocorrelation) เขียนแทนด้วยสัญลักษณ์ ρ_k สามารถประมาณได้จาก

$$\hat{\rho}_k = \frac{Cov(\theta_t, \theta_{t+k})}{Var(\theta_t)} = \frac{\sum_{t=1}^{n-k}(\theta_t - \bar{\theta})(\theta_{t-k} - \bar{\theta})}{\sum_{t=1}^{n-k}(\theta_t - \bar{\theta})^2} \quad \text{เมื่อ } \bar{\theta} = \frac{1}{n}\sum_{t=1}^{n}\theta_t \qquad (2.85)$$

เครื่องมือที่ใช้ในการพิจารณาอัตตสหสัมพันธ์ในค่าสังเกตที่ได้จากเทคนิค MCMC จะใช้ฟังก์ชันอัตตสหสัมพันธ์ (autocorrelation function: ACF) เพื่อตรวจสอบอัตตสหสัมพันธ์ระหว่างค่าสังเกต θ_t กับ θ_{t-p} ซึ่งจะเรียกว่า อัตตสหสัมพันธ์ ณ ช่วงเวลา p (autocorrelation at lag p) และจะใช้ฟังก์ชันอัตตสหสัมพันธ์บางส่วน (partial autocorrelation function: PACF) เพื่อตรวจสอบค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง θ_t กับ θ_{t-k} โดยที่ขจัดสหสัมพันธ์ของ $\rho_1, \rho_2, ..., \rho_{k-1}$ ออกไป ซึ่งจะช่วยในการพิจารณาลำดับของโมเดล AR(p) ยกตัวอย่างเช่น ถ้า

ลูกโซ่มาร์คอฟ θ_t มี PACF ณ lag ที่ p จะมีค่าไม่เท่ากับ 0 อย่างมีนัยสำคัญทางสถิติ กล่าวคือค่าสังเกตที่สำคัญในการสร้างค่าสังเกต θ_t คือ $\theta_{t-1}, \theta_{t-2}, ..., \theta_{t-p}$ ส่วนค่าสังเกตที่เกิน p พจน์แรกจะไม่ได้มีส่วนเกี่ยวข้องหรือมีอิทธิพลในการสร้าง θ_t หรือกล่าวได้ว่าในการวิเคราะห์ลูกโซ่มาร์คอฟดังกล่าวผู้วิจัยควรมีการเลือกเก็บค่าสังเกตเฉพาะค่าสังเกตที่ห่างกันเกิน p ช่วงเวลาเพื่อให้ค่าสังเกตที่ได้นั้นใกล้เคียงข้อสมมติของความเป็นอิสระ เรียกวิธีการดังกล่าวว่า “Thin” ยกตัวอย่างเช่น สมมติว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าสังเกต $\rho_1 = 0.99$ หากกำหนดให้ใช้ตัวอย่างสุ่มเฉพาะตัวที่ 50 จะได้ว่าค่าสัมประสิทธิ์สหสัมพันธ์จะลดลงเป็นมีค่าเท่ากับ $0.99^{50} = 0.605$ อาจเลือกเก็บค่าสังเกตเฉพาะค่าที่ 10 เท่านั้น ซึ่งจะช่วยลดสหสัมพันธ์ระหว่างค่าสังเกตลงได้ ในทำนองเดียวกันหากกำหนดให้ใช้ตัวอย่างสุ่มเฉพาะตัวที่ 100 หรือ 500 จะได้ว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าสังเกตจะมีค่าเท่ากับ $0.99^{100} = 0.366$ และ $0.99^{500} = 0.007$ ตามลำดับ

อีกวิธีการหนึ่งที่ใช้ในการลดทอนปัญหาข้างต้นคือการสร้างลูกโซ่มาร์คอฟให้มีขนาดตัวอย่างมากเพียงพอ หลักการคือถ้าค่าสังเกตที่สุ่มมาจากการแจกแจงความน่าจะเป็นภายหลังนั้นมีความสัมพันธ์กัน จะได้ว่าค่าสังเกตชุดที่สุ่มมานั้นยังสามารถใช้ในการประมาณการแจกแจงความน่าจะเป็นภายหลังที่สนใจได้โดยที่ไม่ยังไม่มีความคลาดเคลื่อนได้อยู่ อย่างไรก็ตามการที่ข้อมูลมีความสัมพันธ์กันนั้นจะทำให้เกิดความซ้ำซ้อนของข้อมูลค่าสังเกต หากผู้วิจัยต้องการนำลูกโซ่มาร์คอฟดังกล่าวไปใช้ในการประมาณการแจกแจงความน่าจะเป็นภายหลังที่ต้องการแล้วขนาดตัวอย่างหรือจำนวนรอบที่ใช้ในการสร้างลูกโซ่มาร์คอฟนั้นจะต้องมีจำนวนที่มากเพียงพอ วิธีการนี้จึงจำเป็นต้องมีเกณฑ์เพื่อพิจารณาว่าควรใช้จำนวนค่าสังเกตเท่าใดจึงจะเพียงพอ วิธีการหนึ่งที่ใช้ในการกำหนดขนาดตัวอย่างที่เหมาะสมนั้นสามารถใช้โมเดล AR(p) เป็นเครื่องมือในการช่วยตัดสินใจถึงจำนวนค่าสังเกตที่เหมาะสมได้ สมมติว่าจากการตรวจสอบแล้วพบวาลูกโซ่มาร์คอฟที่สร้างขึ้นมามีลักษณะที่สอดคล้องกับโมเดล AR(1) ซึ่งเขียนเป็นโมเดลได้ดังนี้

$$\theta_t = \alpha\theta_{t-1} + \varepsilon_t \tag{2.86}$$

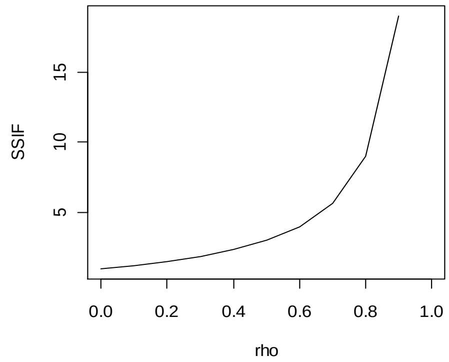
จากโมเดลจะได้ว่า $\rho_1 = \alpha$ และค่าคลาดเคลื่อนมาตรฐานในการประมาณค่าเฉลี่ยของพารามิเตอร์จะมีค่าเท่ากับ

$$S\hat{E}\left(\bar{\theta}\right) = \frac{\hat{\sigma}}{\sqrt{n}}\sqrt{\frac{1+\hat{\rho}_1}{1-\hat{\rho}_1}} \tag{2.87}$$

ในการกำหนดจำนวนรอบที่เหมาะสมนั้นสามารถทำได้โดยกำหนดค่าขอบเขตของค่าคลาดเคลื่อนมาตรฐานในข้างต้นเรียกว่า “ค่า Tolerance” เขียนแทนด้วยสัญลักษณ์ T ดังนี้ $SE(\bar{\theta}) \le T$ ซึ่งเมื่อจัดรูปสมการใหม่จะได้ว่า

$$n \ge \frac{\hat{\sigma}^2}{T^2} \cdot \left(\frac{1+\hat{\rho}_1}{1-\hat{\rho}_1} \right) \tag{2.88}$$

พจน์ $\left(\frac{1+\hat{\rho}_1}{1-\hat{\rho}_1} \right)$ ในอสมการข้างต้น เรียกว่า “Sample size inflation factor (SSIF)” เป็นพจน์ที่บ่งบอกถึงขนาดตัวอย่างหรือจำนวนรอบที่ต้องเพิ่มขึ้นเพื่อที่จะทำให้ได้ความแม่นยำในการประมาณให้อยู่ในระดับที่ต้องการ ซึ่งจะแปลความหมายเป็นจำนวนเท่าของขนาดตัวอย่างที่ต้องเพิ่มขึ้นเมื่อเปรียบเทียบกับขนาดตัวอย่างในสถานการณ์ค่าสังเกตนั้นไม่เกิดปัญหาอัตตสหสัมพันธ์ ยกตัวอย่างเช่นหากได้ค่า SSIF = 10 แสดงว่าในกาสร้างลูกโซ่มาร์คอฟนี้จำเป็นที่จะต้องใช้จำนวนรอบมากกว่ากรณีที่ค่าสังเกตเป็นอิสระกัน 10 เท่าเพื่อที่จะได้ความแม่นยำเท่ากับ T ตามที่ได้ตั้งไว้ รูปที่ 3 แสดงพล๊อตระหว่างค่า SSIF กับอัตตสหสัมพันธ์ ลำดับที่ 1 จะเห็นว่าเมื่อค่า อัตตสหสัมพันธ์มีค่าเพิ่มขึ้น ค่า SSIF จะมีค่าเพิ่มขึ้นด้วย จะเห็นว่าเมื่อค่าอัตตสหสัมพันธ์มีค่าเข้าใกล้ 1 ค่า SSIF จะมีค่าลู่เข้าสู่อนันต์



<u>รูปที่ 2.3</u> กราฟพล๊อตระหว่างค่า SSIF กับค่า อัตตสหสัมพันธ์ ลำดับที่ 1

การวินิจฉัยการลู่เข้าของลูกโซ่ (Convergence Diagnostic)

เราสามารถแบ่งวิธีการวินิจฉัยการลู่เข้าของลูกโซ่ออกได้เป็น 2 แนวทางคือ 1) แนวทางการวินิจฉัยด้วยกราฟ (graphical diagnostic) และ 2) แนวทางการวินิจฉัยด้วยตัวสถิติ (statistical diagnostic) รายละเอียดของแต่ละวิธีการเป็นดังนี้

1) แนวทางการวินิจฉัยด้วยกราฟ (graphical diagnostic)

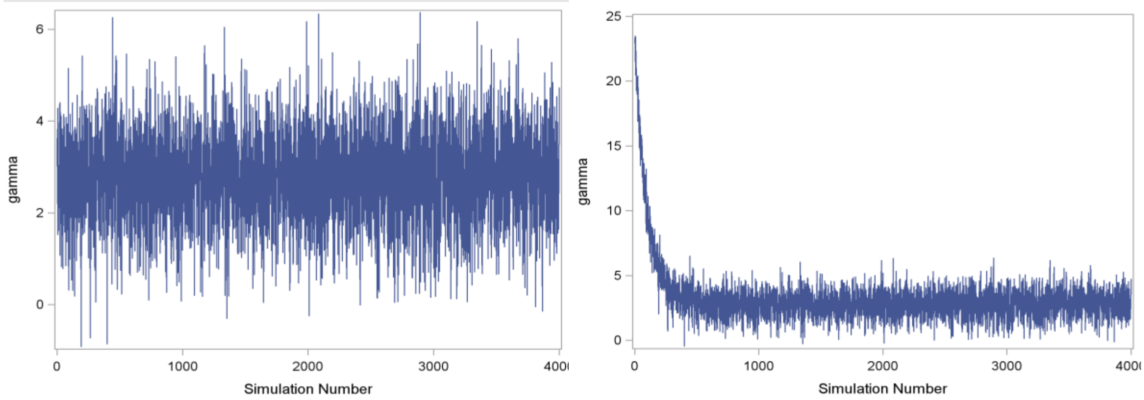
กราฟที่เป็นที่นิยมเพื่อใช้ในการวินิจฉัยการลู่เข้าของลูกโซ่มาร์คอฟนั้นมีอยู่ 2 ประเภท ได้แก่ trace plot และ cumulative estimates of the quantiles plot รายละเอียดมีดังต่อไปนี้

1.1) Trace plot

เป็นกราฟที่พล๊อตค่าของค่าสังเกตที่สุ่มได้ในแต่ละรอบเรียงกันตั้งแต่รอบแรกไปจนถึงรอบสุดท้าย ใช้พิจารณาการไม่ลู่เข้าของลูกโซ่มาร์คอฟ การพิจารณาจำนวน burn-in ที่ต้องใช้ และการพิจารณาว่าลูกโซ่ที่สร้างขึ้นมานั้นผสมผสานกันดีหรือไม่ หาก trace plot ที่ได้มีลักษณะที่คงที่ไม่เปลี่ยนแปลงแล้ว เป็นไปได้ที่ลูกโซ่มาร์คอฟนั้นจะเข้าสู่สถานะอยู่ตัวแล้ว อย่างไรก็ตามในการวินิจฉัยจะไม่สามารถยืนยันการลู่เข้าของลูกโซ่ได้อย่างแน่นอนเนื่องจากเป็นสิ่งที่ไม่อาจทราบได้ รูปที่ 2.3 แสดง trace plot ของลูกมาร์คอฟใน 4 ลักษณะ ดังนี้ รูปที่ 2.3 ก แสดง trace plot ที่แสดงถึงการผสมผสานกันได้ดีของลูกโซ่ ซึ่งอาจสามารถใช้ลูกโซ่ชุดนี้ในการประมาณการแจกแจงความน่าจะเป็นภายหลังเพื่ออนุมานเกี่ยวกับพารามิเตอร์ที่สนใจได้เลย รูปที่ 2.3 ข แสดง trace plot ของลูกโซ่ที่มีการกำหนดค่าเริ่มต้นที่ไม่ดี ทำให้ลูกโซ่ดังกล่าวจะเริ่มผสมผสานกันได้ดีเมื่อจำนวนรอบผ่านไปจำนวนหนึ่งแล้ว ซึ่งในการใช้ลูกโซ่ชุดนี้ในการประมาณการแจกแจงความน่าจะเป็นภายหลังจำเป็นที่จะต้องมีการตัดค่าสังเกตในช่วงแรกทิ้งไปจำนวนหนึ่ง (burn-in) รูปที่ 2.3 ค แสดง trace plot ของลูกโซ่ที่แสดงให้เห็นว่าจำเป็นที่จะต้องใช้จำนวนรอบในการสร้างลูกโซ่ให้มากยิ่งขึ้นเพื่อที่จะทำให้ลูกโซ่ที่สร้างขึ้นเข้าสู่สถานะอยู่ตัว และรูปที่ 2.4 ง แสดง trace plot ของลูกโซ่ที่มีปัญหาในการลู่เข้าสู่สถานะอยู่ตัว จากรูปจะเห็นว่าลูกโซ่ไม่มีแนวโน้มที่จะลู่เข้าสู่ค่าใดค่าหนึ่งหรือบริเวณใดบริเวณหนึ่งเลย

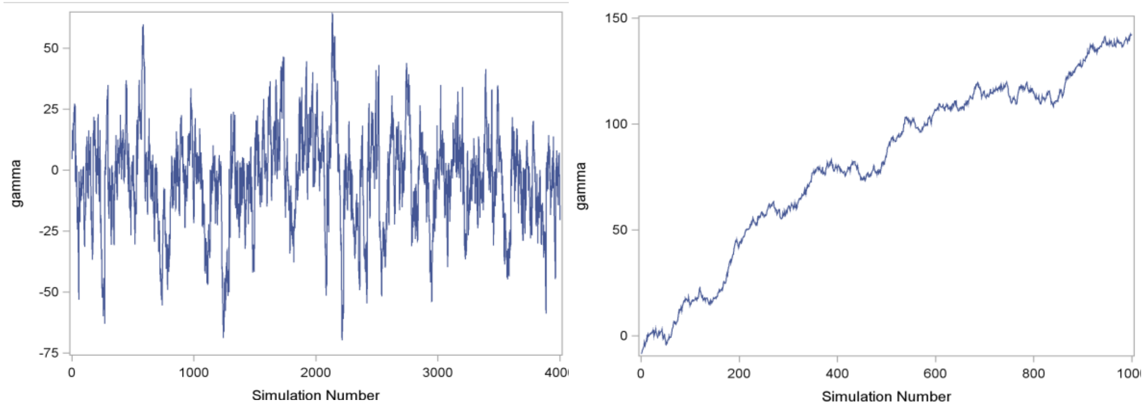
1.2) Cumulative estimates of the quantiles plot

อาจเรียกสั้นๆว่า “Cumulative quantile plot” เป็นกราฟที่พล็อตระหว่างค่าควอนไทล์ที่ประมาณจากตัวอย่างกับจำนวนรอบที่ตั้งแต่รอบแรกไปจนถึงรอบสุดท้าย กราฟดังกล่าวมีวัตถุประสงค์ในการใช้เหมือนกับ trace plot ตัวอย่างของ cumulative quantile plot แสดงไว้ในรูปที่ 2.4



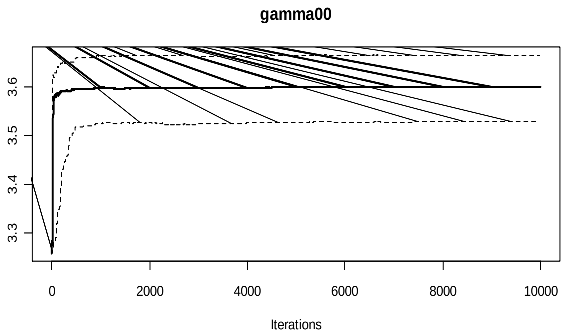
ก ข

<u>รูปที่ 2.4</u> trace plot ของลูกโซ่มาร์คอฟในลักษณะต่างๆ



ค ง

<u>รูปที่ 2.4</u> trace plot ของลูกโซ่มาร์คอฟในลักษณะต่างๆ (ต่อ)



<u>รูปที่ 2.5</u> Cumulative quantile plot

2) แนวทางการวินิจฉัยด้วยตัวสถิติ (statistical diagnostic)

ในการวินิจฉัยการลู่เข้าของลู่โซ่นั้นได้มีการพัฒนาวิธีการทางสถิติเพื่อตรวจสอบการลู่เข้าของลูกโซ่ไว้หลายวิธีการ ซึ่งแต่ละวิธีการค่อนข้างที่จะมีความซับซ้อนในส่วนนี้จึงจะนำเสนอเฉพาะหลักการแนวคิดและวิธีการแปลผลของตัวทดสอบแต่ละตัวอย่างคร่าวๆเท่านั้น รายละเอียดของแต่ละวิธีการมีดังต่อไปนี้

2.1) Geweke’s diagnostic

วิธีการของ Geweke (Geweke, 1992) เป็นวิธีการเพื่อทดสอบว่าลูกโซ่เข้าสู่สถานะอยู่ตัวแล้วหรือไม่ แนวคิดของวิธีการนี้มีแนวคิดว่า ถ้าลูกโซ่มาร์คอฟที่สร้างขึ้นนั้นอยู่ในสถานะอยู่ตัวแล้วค่าเฉลี่ยของลูกโซ่ในช่วงแรกและค่าเฉลี่ยของลูกโซ่ในช่วงหลังควรที่จะต้องไม่แตกต่างกัน สถิติที่ใช้ในการทดสอบคือสถิติทดสอบ Z เรียกว่า “Geweke Z-score” เพื่อเปรียบเทียบค่าเฉลี่ยระหว่างลูกโซ่ใน 2 ส่วน วิธีการในการแปลและสรุปผลนั้นกระทำเหมือนสถิติทดสอบ Z ในการทดสอบเพื่อเปรียบเทียบค่าเฉลี่ย 2 กลุ่มประชากรทุกประการ

2.2) Heidelberger-Welch diagnostic

วิธีการของ Heidelberger และ Welch (Heidelberger and Welch, 1981, 1983) ประกอบไปด้วย 2 ส่วน ส่วนแรกเป็นส่วนที่ใช้ในการทดสอบสถานะอยู่ตัวของลูกโซ่ (stationary test) ซึ่งเป็นการทดสอบโดยใช้ตัวสถิติ Cramer-von-Mises เพื่อทดสอบสมมติฐานว่าลูกโซ่ที่สุ่มมานั้นมาจาก covariance stationary process การปฏิเสธสมมติฐานจะหมายความว่าลูกโซ่ดังกล่าวต้องการจำนวนรอบในการสร้างลูกโซ่เพิ่มขึ้น ส่วนที่สองเรียกว่า “Half-width test” เป็นส่วนที่ใช้ตรวจสอบว่าจำนวนรอบที่ใช้ในการสร้างลูกโซ่นั้นเพียงพอต่อการประมาณค่าเฉลี่ยของลูกโซ่ที่แม่นยำหรือไม่ ในการแปลผลการทดสอบการทดสอบสถานะอยู่ตัวของลูกโซ่จะซึ่งทดสอบด้วยตัวสถิติ Cramer-von-Mises จะเป็นการทดสอบทางเดียว การปฏิเสธสมมติฐานหลักแสดงว่ามีหลักฐานที่น่าเชื่อได้ว่าลูกโซ่ยังไม่เข้าสู่สถานะอยู่ตัว ส่วนการทดสอบ half-width นั้นจะพิจารณาจากค่าสถิติ half-width (half-width statistic) ถ้ามีค่ามากกว่าเกณฑ์ความแม่นยำที่ได้กำหนดไว้ในเบื้องต้น จะแสดงว่าจำนวนรอบที่ใช้ในการสร้างลูกโซ่ในปัจจุบันนั้นยังไม่เพียงพอที่จะได้ความแม่นยำในการประมาณค่าเฉลี่ยของค่าสังเกตได้ตามระดับที่ต้องการ

2.3) Gelman-Rubin diagnostic

วิธีการทั้งสองวิธีการที่ได้เสนอไว้ก่อนหน้าเป็นวิธีการวินิจฉัยลูกโซ่โดยพิจารณาจากลูกโซ่เพียงเส้นเดียวเท่านั้น Gelman และ Rubin (Gelman and Rubin, 1992) ได้นำเสนอวิธีการที่แตกต่างออกไป โดยใช้เทคนิคการวิเคราะห์ความแปรปรวนมาช่วยในการพิจารณาลูกโซ่หลายลูกโซ่ที่สร้างจากค่าเริ่มต้นที่แตกต่างกันและมีความหลากหลาย แนวคิดนี้เป็นแนวคิดที่ใช้ได้ดีกับปัญหาที่มีความซับซ้อนและการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่สนใจมีการแจกแจงมีเป็นลักษณะหลายฐานนิยม (multi-modal) เพราะการกำหนดค่าเริ่มต้นหลายค่านั้นจะช่วยในการพิจารณาว่ามีลูกโซ่ที่ติดอยู่ในบริเวณฐานนิยมใดฐานนิยมหนึ่งหรือไม่ ซึ่งการสรุปที่ได้จากตัวอย่างลูกโซ่ชุดนั้นมีความผิดพลาด สถิติที่ใช้ในกาวินิจฉัยนั้นจะใช้สารสนเทศจากความแปรปรวนภายในลูกโซ่ (within chain variance) กับความแปรปรวนระหว่างลูกโซ่ (between-chain variance) มาสร้างเป็นตัวสถิติที่เรียกว่า “Shrink factor” ซึ่งหากตัวสถิติมีค่าเข้าใกล้ 1 จะเป็นตัวชี้ว่าลูกโซ่ลู่เข้าสู่สถานะอยู่ตัวแล้ว อย่างไรก็ตามการพิจารณาจากค่าสถิติ shrink factor แต่เพียงอย่างเดียวไม่เพียงพอเนื่องจากมีโอกาสที่ลูกโซ่จะไม่ลู่เข้า แต่ค่าสถิติ shrink factor นั้นได้ค่าเข้าใกล้ 1 ดังนั้นเพื่อเป็นการป้องกันความผิดพลาดดังกล่าวจึงควรมีการพิจารณากราฟของ shrink factor ที่พล๊อตระหว่างค่า shrink factor กับ จำนวนรอบที่ได้ทำการสุ่มตัวอย่างลูกโซ่เพื่อตรวจสอบความคงที่ของค่า shrink factor หากพิจารณากราฟแล้วค่า shrink factor ไม่ได้มีแนวโน้มที่คงที่ แสดงว่าลูกโซ่ที่สร้างขึ้นนั้นยังไม่ได้ลู่เข้าสู่สถานะอยู่ตัว

สถิติที่ใช้ในการสรุปสารสนเทศที่ได้จากลูกโซ่ (summary statistics)

ค่าสังเกตที่สุ่มจากระบบของลูกโซ่มาร์คอฟด้วยเทคนิค MCMC จะใช้ตัวอย่างเพื่อประมาณการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ ดังนั้นจึงสามารถใช้ตัวอย่างค่าสังเกตดังกล่าวในการประมาณค่าพารามิเตอร์และค่าสถิติอื่นๆที่สนใจได้โดยตรง ในหัวข้อนี้จะนำเสนอสถิติที่ใช้ในการสรุปสารสนเทศของพารามิเตอร์จากชุดของตัวอย่างสุ่ม รายละเอียดมีดังต่อไปนี้

กำหนดให้ $\Theta = \{\underline{\theta}_1, \underline{\theta}_2, ,,.., \underline{\theta}_p\}$ เป็นอาเรย์ (array) ที่เก็บเวกเตอร์ของพารามิเตอร์ที่สนใจจำนวน p ตัว โดยที่ในแต่ละเวกเตอร์ของพารามิเตอร์ $\underline{\theta}_i$; $i = 1,2,..., p$ จะมีสังเกตที่ได้

จากการสุ่มตัวอย่างลูกโซ่จำนวน n ค่าสังเกต กล่าวคือ $\underline{\theta}_i = \{\theta_i^t \mid t = 1,2,...,n\}$ ค่าสถิติที่ใช้ในการสรุปสารสนเทศที่ได้จากตัวอย่างสุ่มของพารามิเตอร์ที่สนใจมีดังต่อไปนี้

1) ค่าเฉลี่ยภายหลัง (posterior mean)

ค่าเฉลี่ยภายหลังโดยใช้สารสนเทศจากตัวอย่างที่สุ่มมาคำนวณได้ดังนี้

$$E(\underline{\theta}_i \mid \underline{y}) \approx \bar{\theta}_i = \frac{1}{n}\sum_{t=1}^{n}\theta_i^t \qquad \text{เมื่อ } i = 1,2,...,p \tag{2.89}$$

2) ความแปรปรวนภายหลัง (posterior variance)

ความแปรปรวนภายหลังโดยใช้สารสนเทศจากตัวอย่างที่สุ่มมาคำนวณได้ดังนี้

$$Var(\underline{\theta}_i \mid \underline{y}) \approx s_i^2 = \frac{1}{n-1}\sum_{t=1}^{n}\left(\theta_i^t - \bar{\theta}_i\right)^2 \tag{2.90}$$

3) ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยภายหลัง[5] (standard error of posterior mean)

ในกรณีที่เป็นการใช้เทคนิค Monte Carlo แบบปกติการคำนวณค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยนั้นสามารถกระทำได้โดยตรงดังนี้ $\hat{\sigma}_i / \sqrt{n}$ แต่ในกรณีที่เป็นตัวอย่างสุ่มที่มาจากเทคนิค MCMC ต้องไม่ลืมว่าตัวอย่างสุ่มที่ได้นั้นมีความสัมพันธ์กันเอง กล่าวคือเกิดอัตตสหสัมพันธ์ขึ้น การใช้วิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานแบบปกติจะทำให้ประมาณค่าคลาดเคลื่อนมาตรฐานได้ต่ำกว่าปกติ (underestimate) ดังนั้นในการประมาณค่าคลาดเคลื่อนมาตรฐานที่ถูกต้องนั้นจะต้องมีการปรับสูตรการประมาณให้มีการคำนึงถึงอัตตสหสัมพันธ์ด้วยดังนี้

$$SE(\bar{\theta}_i) = \sqrt{\frac{1 + 2\sum_{k=1}^{\infty}\rho_k(\theta_i)}{n} \cdot \frac{1}{n-1}\sum_{t=1}^{n}\left(\theta_i^t - \bar{\theta}_i\right)^2} \tag{2.91}$$

พิจารณาพจน์ $\dfrac{1 + 2\sum_{k=1}^{\infty}\rho_k(\theta_i)}{n} = \dfrac{1}{ESS}$ เมื่อ ESS คือ “effective sample size” ซึ่งเป็นค่าที่ใช้ในการพิจารณาการลู่เข้าของลูกโซ่ได้อีกตัวหนึ่ง ถ้าค่า ESS มีค่าสูงแสดงว่าลูกโซ่มีแนวโน้มว่าจะอยู่ในสถานะอยู่ตัวแล้ว แต่หากมีค่าต่ำแสดงว่าลูกโซ่ดังกล่าวต้องเพิ่มจำนวนรอบให้มากขึ้น

---

[5] อาจเรียกว่า Monte Carlo standard error (MCSE)

4) ช่วงความน่าเชื่อถือแบบหางเท่ากัน (equal-tail credible interval)

กำหนดให้ $\pi(\theta_i \mid \underline{y})$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็นส่วนริมสะสมของพารามิเตอร์ θ_i (marginal posterior cumulative distribution function) ช่วงความน่าเชื่อถือแบบหางเท่าสำหรับพารามิเตอร์ θ_i คือช่วง $\left(\theta_i^{\alpha/2}, \theta_i^{1-\alpha/2}\right)$ เมื่อ $\pi(\theta_i^{\alpha/2} \mid \underline{y}) = \dfrac{\alpha}{2}$ และ $\pi(\theta_i^{1-\alpha/2} \mid \underline{y}) = 1 - \dfrac{\alpha}{2}$ ในทางปฏิบัติการหาช่วงความน่าเชื่อถือจากตัวอย่างจะใช้ตัวอย่าง ณ เปอร์เซนไทล์ที่ $\dfrac{\alpha}{2}$ และ $1-\dfrac{\alpha}{2}$ เป็นขอบเขตล่างและขอบเขตบนของช่วงความน่าเชื่อถือตามลำดับ

5) Deviance Information Criterion (DIC)

ค่าสถิติ DIC (Spiegelhalter et al. 2002) เป็นค่าสถิติเพื่อใช้ในการเปรียบเทียบโมเดล ค่า DIC ที่มีขนาดเล็กจะแสดงว่าโมเดลนั้นมีความเหมาะสมกับข้อมูล กล่าวคือโมเดลที่เหมาะสมกับข้อมูลมากกว่าจะมีค่า DIC น้อยกว่าโมเดลอื่นๆ กำหนดให้ θ เป็นเวกเตอร์ของพารามิเตอร์ในโมเดล สูตรที่ใช้ในการคำนวณ DIC เป็นดังนี้

$$DIC = \overline{D(\theta)} + p_D = D\left(\bar{\theta}\right) + 2p_D \tag{2.92}$$

เมื่อ $D(\theta) = 2\left(\log\left(f(\underline{y})\right) - \log(p(\underline{y} \mid \theta)\right)$ เป็นค่า deviance ของโมเดล โดยที่ $p(\underline{y} \mid \theta)$ คือฟังก์ชันภาวะความควรจะเป็น และ $f(\underline{y})$ คือพจน์มาตรฐาน (standardized term) ที่เป็นฟังก์ชันที่ขึ้นกับข้อมูลค่าสังเกตเพียงอย่างเดียว

$\overline{D(\theta)}$ เป็นค่าเฉลี่ยภายหลังของ deviance คำนวณได้จาก $\dfrac{1}{n}\sum_{t=1}^{n} D\left(\theta^t\right)$ ซึ่งเป็นค่าที่ใช้วัดความเหมาะสมของโมเดลกับข้อมูลโดยเฉลี่ย

$D\left(\bar{\theta}\right)$ เป็นค่า deviance ที่ใช้ประเมินการประมาณค่า $\bar{\theta}$ มีค่าเท่ากับ $-2\log(p(\underline{y} \mid \bar{\theta})$

p_D เรียกว่า effective number of parameters มีค่าเท่ากับ $\overline{D(\theta)} - D\left(\bar{\theta}\right)$ เป็นค่าที่ใช้อธิบายความซับซ้อนของโมเดล

ตอนที่ 3 การประมาณค่าพารามิเตอร์ในโมเดลพหุระดับด้วยสถิติแบบเบส์

จากที่ได้กล่าวไว้ในหัวข้อก่อนหน้าแล้วว่า แนวทางในการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับนั้นสามารถแบ่งออกได้เป็น 2 แนวทางได้แก่ แนวทางแบบดั้งเดิมและแนวทางแบบเบส์ ในหัวข้อนี้จะนำเสนอแนวคิดการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับโดยอาศัยแนวทางแบบเบส์ซึ่งจำเป็นที่จะต้องใช้เทคนิค MCMC มาช่วยในการประมาณการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ในโมเดล และการแจกแจงความน่าจะเป็นภายหลังที่ได้นั้นจะใช้เป็นเครื่องมือในการอนุมานต่อไป โดยในหัวข้อนี้จะนำเสนอการประยุกต์ใช้หลักการของเบส์เซียนในการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับ

โดยทั่วไปในการวิเคราะห์ข้อมูลโดยอาศัยสถิติแบบเบส์นั้นอาจสามารถแบ่งออกได้เป็น 4 ขั้นตอนได้แก่ 1) ขั้นตอนการสร้างโมเดล (model building): เป็นขั้นตอนในการนิยามโมเดลที่สนใจประกอบไปด้วย การนิยามตัวแปร รูปแบบของโมเดล การแจกแจงความน่าจะเป็นก่อนหน้าและฟังก์ชันภาวะความควรจะเป็น 2) ขั้นตอนการคำนวณการแจกแจงความน่าจะเป็นภายหลัง (calculation of the posterior distribution): เป็นขั้นตอนในการเลือกวิธีการคำนวณการแจกแจงความน่าจะเป็นภายหลัง (วิธีการเชิงวิเคราะห์ (analytic method) วิธีการเชิงตัวเลข (numerical method) และวิธีลูกโซ่มาร์คอฟมอนติคาร์โล (MCMC)) รวมไปถึงการคำนวณการแจกแจงความน่าจะเป็นภายหลังดังกล่าว 3) ขั้นตอนการวิเคราะห์การแจกแจงความน่าจะเป็นภายหลัง (analysis of the posterior distribution): เป็นขั้นตอนการวิเคราะห์การแจกแจงความน่าจะเป็นภายหลังที่สร้างขึ้นเพื่อพิจารณาความเหมาะสมและความน่าเชื่อถือของการแจกแจงความน่าจะเป็นภายหลังดังกล่าว และ 4) ขั้นตอนการอนุมาน (inference): เป็นขั้นตอนการนำเอาสารสนเทศที่ได้จากการแจกแจงความน่าจะเป็นภายหลังไปอนุมานเกี่ยวกับพารามิเตอร์ที่สนใจ (Browne, 1998; Gelman, Carlin, Stern, and Rubin, 1995) ในส่วนนี้จะนำเสนอวิธีการประมาณค่าพารามิเตอร์ของโมเดลเชิงเส้นพหุระดับได้แก่ การประมาณพารามิเตอร์ในโมเดลว่างแบบ 2 ระดับ โมเดลสัมประสิทธิ์ความถดถอยสุ่มแบบ 2 ระดับ และโมเดลสัมประสิทธิ์ความถดถอยสุ่มในกรณีทั่วไป อัลกอริทึมที่ใช้ในการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับที่นำเสนอจะใช้อัลกอริทึมการสุ่มตัวอย่างแบบกิบบ์ (Gibbs sampling algorithm) รายละเอียดเป็นดังนี้

3.1 <u>การกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าในโมเดลเชิงเส้นพหุระดับ</u>

ในโมเดลเชิงเส้นพหุระดับสามารถแบ่งพารามิเตอร์ในโมเดลออกได้เป็น 2 ส่วนได้แก่ ส่วนของพารามิเตอร์อิทธิพลคงที่ (fixed effect parameters) ซึ่งได้แก่ ความสัมประสิทธิ์ความถดถอยในโมเดล และส่วนของพารามิเตอร์อิทธิพลสุ่ม (random effect parameters) ซึ่งประกอบไปด้วยค่าความคลาดเคลื่อน และส่วนประกอบความแปรปรวน ถ้าพิจารณาธรรมชาติของพารามิเตอร์ในแต่ละส่วนจะพบว่ามีลักษณะตามธรรมชาติที่แตกต่างกัน ดังนั้นการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในแต่ละส่วนจึงมีการกำหนดรูปแบบการแจกแจงความน่าจะเป็นที่แตกต่างกัน (Browne and Draper, 2006; Gelman, 2006; Browne, Goldstein, Woodhouse and Yang, 2001; Browne, 1995; Gelman, Carlin, Stern, and Rubin, 1995)

1) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์อิทธิพลคงที่ (prior distribution for fixed effect parameters)

เมื่อพิจารณาจากธรรมชาติของค่าเฉลี่ยหรือค่าพารามิเตอร์สัมประสิทธิ์ถดถอยจะพบว่าสามารถมีค่าเป็นค่าใดก็ได้ในเซตของจำนวนจริง ดังนั้นการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศที่นิยมเลือกใช้กันได้แก่ การแจกแจงความน่าจะเป็นก่อนหน้าแบบเอกรูป (uniform prior distribution) การแจกแจงความน่าจะเป็นก่อนหน้าแบบปกติการแจกแจงความน่าจะเป็นก่อนหน้าแบบปกติหลายตัวแปร (multivariate normal prior distribution) รายละเอียดเป็นดังนี้

การแจกแจงความน่าจะเป็นก่อนหน้าแบบเอกรูป (uniform prior for fixed effect parameter) ในกรณีนี้มักจะกำหนดให้เป็นการแจกแจงความน่าจะเป็นแบบเอกรูปบนจำนวนจริงตามธรรมชาติของค่าที่เป็นไปได้ของพารามิเตอร์อิทธิพลคงที่ ($p(\theta) \propto 1$) การแจกแจงดังกล่าวถือเป็นการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศ เพราะถือว่าทุกๆค่าที่เป็นไปได้ของพารามิเตอร์อิทธิพลคงที่นั้นมีโอกาสในการเกิดเท่ากันหมด การแจกแจงความน่าจะเป็นก่อนหน้านี้เป็นการแจกแจงความน่าจะเป็นที่ไม่ตรงแบบ (improper prior distribution) เพราะอินทิเกรตของฟังก์ชันความหนาแน่นนั้นมีค่าไม่เท่ากับ 1 อย่างไรก็ตามการแจกแจงความน่าจะเป็นภายหลังที่ได้จากการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้านี้ยังเป็นการแจกแจงความ

น่าจะเป็นที่ตรงแบบอยู่ (proper posterior distribution) ด้วยเหตุผลดังกล่าวการแจกแจงความน่าจะเป็นแบบเอกรูปนี้ดังกล่าวนี้เป็นการแจกแจงหนึ่งที่นักสถิติมักนิยมเลือกใช้ เพราะสามารถหารูปแบบของการแจกแจงความน่าจะเป็นภายหลังได้ง่าย และการแจกแจงความน่าจะเป็นภายหลังที่ได้นั้นยังเป็นการแจกแจงความน่าจะเป็นที่ตรงแบบ (proper posterior distribution)

การแจกแจงความน่าจะเป็นอีกรูปแบบหนึ่งที่นิยมเลือกใช้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์อิทธิพลคงที่คือ การแจกแจงความน่าจะเป็นแบบปกติหรือการแจกแจงความน่าจะเป็นก่อนหน้าแบบปกติหลายตัวแปร (normal prior for fixed effect parameters) การแจกแจงดังกล่าวเป็นที่นิยมเพราะเป็นการแจกแจงที่อยู่ในวงศ์ชี้กำลัง (exponential family) ซึ่งมีคุณสมบัติที่เรียกว่าวงศ์คู่สังยุค (conjugacy) ซึ่งทำให้การแจกแจงความน่าจะเป็นก่อนหน้า และการแจกแจงความน่าจะเป็นภายหลังนั้นมีรูปแบบเดียวกัน คุณสมบัติดังกล่าวสะดวกต่อการหารูปแบบของการแจกแจงความน่าจะเป็นภายหลัง นอกจากนี้การแจกแจงความน่าจะเป็นแบบปกติยังเอื้อให้ผู้วิจัยปรับความเชื่อเกี่ยวกับความรู้ก่อนหน้าของพารามิเตอร์ได้พอสมควร กล่าวคือหากผู้วิจัยไม่มีความมั่นใจหรือมีความรู้ก่อนหน้าเกี่ยวกับพารามิเตอร์ที่สนใจน้อย ผู้วิจัยสามารถปรับให้การแจกแจงความน่าจะเป็นก่อนหน้านี้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศโดยการกำหนดให้ไฮเปอร์พารามิเตอร์ (hyperparameter) ค่าเฉลี่ยก่อนหน้ามีค่าเท่ากับ 0 และความแปรปรวนก่อนหน้ามีขนาดใหญ่ รูปแบบของการแจกแจงแบบปกติที่มีการกำหนดค่าไฮเปอร์พารามิเตอร์ดังกล่าวจะมีรูปแบบที่คล้ายคลึงกับการแจกแจงแบบเอกรูปดังที่ได้กล่าวไปในข้างต้น ซึ่งจะเห็นตัวอย่างได้ในรูปที่ 2.1 ก

2) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวน

เมื่อพิจารณาจากธรรมชาติของพารามิเตอร์ความแปรปรวนนั้นพบว่ามีข้อจำกัดคือพารามิเตอร์ความแปรปรวนนั้นไม่สามารถติดลบได้ ดังนั้นการแจกแจงความน่าจะเป็นก่อนหน้าแบบปกตินั้นจึงไม่สามารถใช้ได้กับพารามิเตอร์ความแปรปรวนได้ การแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศของพารามิเตอร์ความแปรปรวนที่นักสถิตินิยมเลือกใช้และเป็นการแจกแจงที่มีคุณสมบัติวงศ์คู่สังยุคได้แก่ การแจกแจงความน่าจะเป็นแบบเอกรูป (uniform distribution) สำหรับพารามิเตอร์ความแปรปรวน σ^2 การแจกแจงความน่าจะเป็นแบบไคสแควร์

(chi-square distribution) หรือการแจกแจงแบบแกมมา (gamma distribution) โดยกำหนดให้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าของส่วนกลับของพารามิเตอร์ความแปรปรวน $1/\sigma^2$ หรือผู้วิจัยอาจกำหนดการแจกแจงความน่าจะเป็นให้เป็นแบบไคสแควร์ผกผัน (inverse chi-square distribution) หรือการแจกแจงความน่าจะเป็นแบบแกมมาผกผัน (inverse gamma distribution) ของพารามิเตอร์ความแปรปรวน σ^2 โดยตรงก็ได้

การใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบเอกรูปกับพารามิเตอร์ความแปรปรวนนั้นสามารถกระทำได้ 2 รูปแบบคือ การกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอของพารามิเตอร์ความแปรปรวน σ^2 กล่าวคือ $p(\sigma^2) \propto 1$ และการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอของพารามิเตอร์ $\log\sigma^2$ วิธีการแรกนั้นจะให้การแจกแจงความน่าจะเป็นภายหลังที่ตรงแบบ แต่มีข้อเสียคือไม่มีความสมเหตุสมผลเนื่องจากเป็นการกำหนดให้ความน่าจะเป็นที่ค่าพารามิเตอร์ความแปรปรวนจะมีค่ามากนั้นมีความเท่าเทียมกับความน่าจะเป็นที่ค่าพารามิเตอร์ความแปรปรวนจะมีค่าน้อย ส่วนวิธีการที่สองนั้นเป็นวิธีการที่พัฒนาขึ้นโดยมีแนวคิดที่จะใช้การแปลงข้อมูลเพื่อให้ได้การแจกแจงความน่าจะเป็นก่อนหน้าที่ไม่ให้สารสนเทศที่มีความเหมาะสม (Box and Tiao, 1992) แต่อย่างไรก็ตามในกรณีที่ของโมเดลเชิงเส้นพหุระดับนั้นการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนดังกล่าวอาจทำให้การแจกแจงความน่าจะเป็นภายหลังที่ได้นั้นเป็นการแจกแจงไม่ตรงแบบ (Dumouchel and Waternaux, 1992) ดังนั้นจึงไม่ควรใช้การแจกแจงดังกล่าวในการพัฒนาการประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับ (Browne, 1998)

การแจกแจงความน่าจะเป็นก่อนหน้าแบบแกมมาสำหรับพารามิเตอร์ส่วนกลับของความแปรปรวน $1/\sigma^2$ อาจเรียกว่าพารามิเตอร์ความแม่นยำ (precision parameters) การแจกแจงความน่าจะเป็นแบบแกมมาที่มีพารามิเตอร์ α และ β ของพารามิเตอร์ $1/\sigma^2$ สามารถใช้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าได้ โดยการแจกแจงดังกล่าวจะลู่เข้าสู่การแจกแจงแบบสม่ำเสมอของพารามิเตอร์ $\log\sigma^2$ เมื่อ $\alpha = \beta \to 0$ ทำให้ในกรณีที่ผู้วิจัยมีสารสนเทศเกี่ยวกับพารามิเตอร์ดังกล่าวต่ำ อาจเลือกกำหนดให้ค่าพารามิเตอร์ของการแจกแจงความน่าจะเป็นทั้งสองมีค่าน้อยๆ การแจกแจงความน่าจะเป็นก่อนหน้าแบบแกมมาเป็นการแจกแจงความน่าจะเป็นที่มีคุณสมบัติวงศ์คู่สังยุค ทำให้การแจกแจงความน่าจะเป็นภายหลังที่สร้างจากการแจกแจงความ

น่าจะเป็นก่อนหน้าแบบแกมมามีการแจกแจงแบบแกมมาเช่นเดียวกัน ข้อควรระวังคือการแจกแจงความน่าจะเป็นดังกล่าวเป็นส่วนกลับของพารามิเตอร์ความแปรปรวน ดังนั้นในกรณีที่ผู้วิจัยต้องการอนุมานเกี่ยวกับพารามิเตอร์ความแปรปรวน ผู้วิจัยจำเป็นต้องหาส่วนผกผันของพารามิเตอร์ความแม่นยำที่ได้เพื่อให้ได้พารามิเตอร์ความแปรปรวนที่ต้องการ นอกจากนี้ผู้วิจัยสามารถกำหนดให้การแจกแจงความน่าจะเป็นก่อนหน้าของส่วนกลับพารามิเตอร์ความแปรปรวนดังกล่าวเป็นการแจกแจงแบบไคสแควร์ได้ อย่างไรก็ตามการแจกแจงทั้งสองล้วนเป็นการแจกแจงความน่าจะเป็นเดียวกัน กล่าวคือ การแจกแจงความน่าจะเป็นแบบไคสแควร์เป็นกรณีเฉพาะของการแจกแจงความน่าจะเป็นแบบแกม ที่กำหนดพารามิเตอร์ $\alpha = n/2$ และ $\beta = 2$ ตามลำดับ นอกจากนี้ผู้วิจัยยังสามารถกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าให้กับพารามิเตอร์ความแปรปรวน σ^2 โดยตรง จากการใช้การแจกแจงความน่าจะเป็นแกมมาผกผัน หรือการแจกแจงความน่าจะเป็นไคสแควร์ผกผัน ซึ่งทำให้มีความสะดวกมากขึ้นเพราะผู้วิจัยไม่ต้องหาส่วนกลับของพารามิเตอร์ความแม่นยำในภายหลัง

3) การแจกแจงความน่าจะเป็นของพารามิเตอร์เมทริกซ์ความแปรปรวนร่วม

จากแนวคิดของการกำหนดความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนนั้นสามารถขยายเป็นการแจกแจงความน่าจะเป็นของเมทริกซ์ความแปรปรวนร่วมได้เช่นเดียวกัน โดยที่การแจกแจงที่ใช้จะเป็นกรณีทั่วไปของการแจกแจงความน่าจะเป็นก่อนหน้าที่กำหนดให้กับพารามิเตอร์ความแปรปรวนในหัวข้อที่แล้ว การแจกแจงที่นิยมใช้คือ การแจกแจงความน่าจะเป็นแบบเอกรูป และการแจกแจงความน่าจะเป็นของเมทริกซ์แบบ Wishart

การแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอสำหรับพารามิเตอร์เมทริกซ์ความแปรปรวนร่วมนั้น สามารถกำหนดให้ในทำนองเดียวกันกับการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าให้กับพารามิเตอร์ความแปรปรวนที่นำเสนอไว้ก่อนหน้า คือ การกำหนดแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอให้กับเมทริกซ์ความแปรปรวนร่วม Σ หรือ $p(\Sigma) \propto 1$ และการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอให้กับ $\log \Sigma$ กล่าวคือ $p(\log \Sigma) \propto 1$ อย่างไรก็ตามการแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอของพารามิเตอร์เมทริกซ์ความแปรปรวนร่วมนั้นไม่เป็นที่นิยมใช้กัน เนื่องด้วยการแจกแจงความน่าจะเป็นภายหลังที่ได้อาจไม่ได้

เป็นการแจกแจงความน่าจะเป็นที่ตรงแบบ และยังไม่สอดคล้องกับธรรมชาติของพารามิเตอร์ความแปรปรวนอีกด้วย

การแจกแจงความน่าจะเป็นของเมทริกซ์แบบ Wishart สามารถใช้ในการกำหนดให้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์อินเวอร์สของเมทริกซ์ความแปรปรวนร่วม Σ^{-1} เรียกว่าเมทริกซ์ความแม่นยำ (precision matrix) การแจกแจงดังกล่าวเป็นการแจกแจงที่นักสถิติมักนิยมเลือกใช้เป็นการแจกแจงความน่าจะเป็นก่อนหน้าเนื่องจากเป็นการแจกแจงที่เป็นวงศ์คู่สังยุคกับการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปร ทำให้การแจกแจงความน่าจะเป็นภายหลังของเมทริกซ์ความแม่นยำ Σ^{-1} มีการแจกแจงความน่าจะเป็นแบบ Wishart และในทำนองเดียวกับการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวน ผู้วิจัยอาจกำหนดให้การแจกแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์เมทริกซ์ความแปรปรวนร่วม Σ เป็นการแจกแจงความน่าจะเป็น Wishart ผกผันซึ่งมีคุณสมบัติเป็นวงศ์คู่สังยุคกับการแจกแจงความน่าจะเป็นแบบปกติเช่นเดียวกัน

3.2 <u>โมเดล 2 ระดับแบบไม่มีเงื่อนไขสมบูรณ์</u> (two-level fully unconditional model or two – level null model)

การประมาณค่าพารามิเตอร์ด้วยสถิติแบบเบส์สามารถดำเนินการได้ตามขั้นตอน 4 ขั้นตอนตามที่ได้กล่าวไว้ในข้างต้น รายละเอียดเป็นดังต่อไปนี้

<u>ขั้นตอนที่1</u>: ขั้นตอนการสร้างโมเดล (model building)

โมเดลที่ต้องการประมาณค่าพารามิเตอร์มีรายละเอียดดังสมการที่ (2.93) ถึง (2.95) ดังนี้

<u>โมเดลระดับที่ 1</u>

$$y_{ij} = \beta_{0j} + \varepsilon_{ij} \tag{2.93}$$

<u>โมเดลระดับที่ 2</u>

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{2.94}$$

<u>โมเดลรวม</u>

$$y_{ij} = \gamma_{00} + u_{0j} + \varepsilon_{ij} \tag{2.95}$$

เมื่อ $\varepsilon_{ij} \sim Normal(0,\sigma_\varepsilon^2)$ และ $u_{0j} \sim Normal(0,\sigma_u^2)$, $i=1,2,...,n_j$ และ $j=1,2,...,J$

จากโมเดลในข้างต้นจะพบว่ามีพารามิเตอร์ที่ต้องการประมาณค่าจำนวน 3 พารามิเตอร์ ได้แก่ พารามิเตอร์อิทธิพลคงที่ (γ_{00}) พารามิเตอร์ส่วนประกอบความแปรปรวน (σ_ε^2 และ σ_u^2)

ในการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้านั้นจะกำหนดให้ พารามิเตอร์ที่เป็นอิทธิพลคงที่ γ_{00} มีการแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอ ($p(\gamma_{00}) \propto 1$) พารามิเตอร์ความแปรปรวน σ_ε^2 และ σ_u^2 กำหนดให้มีการแจกแจงแบบ scaled inverse chi-square ที่มีพารามิเตอร์ ν_ε กับ S_ε^2 และ ν_u กับ S_u^2 ตามลำดับ

<u>ขั้นตอนที่2</u>: ขั้นตอนการคำนวณการแจกแจงความน่าจะเป็นภายหลัง (calculation of the posterior distribution)

เมื่อกำหนดโมเดลและการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์เรียบร้อยแล้ว ขั้นตอนต่อไปคือการประมาณการแจกแจงความน่าจะเป็นภายหลังที่ต้องการ ในกรณีนี้คือการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ γ_{00}, σ_ε^2 และ σ_u^2 เมื่อกำหนดข้อมูลค่าสังเกต $\underline{y}$ ในขั้นตอนนี้จะใช้อัลกอริทึมการสุ่มตัวอย่างแบบกิบบ์ อัลกอริทึมประกอบไปด้วย 4 ขั้นตอน ดังต่อไปนี้ (Browne, 1998)

<u>ขั้นที่ 1</u>: สุ่มตัวอย่างพารามิเตอร์ γ_{00} จาก $p(\gamma_{00} \mid \underline{y},\sigma_\varepsilon^2,\sigma_u^2,\underline{u})$

เนื่องจาก $p(\gamma_{00} \mid \underline{y},\sigma_\varepsilon^2,\sigma_u^2,\underline{u}) \propto \prod_{i,j}\left(\frac{1}{\sigma_\varepsilon^2}\right)^{1/2} \cdot \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(y_{ij}-u_j-\gamma_{00}\right)^2\right\}$

ซึ่งเมื่อจัดรูปแล้วจะได้ว่า $\gamma_{00} \mid \underline{y},\sigma_\varepsilon^2,\sigma_u^2,\underline{u} \sim Normal(\hat{\gamma}_{00},\hat{D}_0)$

โดยที่ $\hat{D}_0 = \frac{\sigma_\varepsilon^2}{N}$ และ $\hat{\gamma}_{00} = \frac{\sum_j \sum_i (y_{ij}-u_{0j})}{N}$

<u>ขั้นที่ 2</u>: สุ่มตัวอย่าง u_{0j} จาก $p(u_{0j} \mid \underline{y},\sigma_\varepsilon^2,\sigma_u^2,\gamma_{00})$

เนื่องจาก

$$p(u_{0j} \mid \underline{y}, \sigma_\varepsilon^2, \sigma_u^2, \gamma_{00}) \propto \prod_{i=1}^{n_j} \left(\frac{1}{\sigma_\varepsilon^2} \right)^{1/2} \cdot \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(y_{ij} - u_j - \gamma_{00} \right)^2 \right\} \cdot \left(\frac{1}{\sigma_u^2} \right)^{1/2} \cdot \exp\left\{ -\frac{u_j^2}{2\sigma_u^2} \right\}$$

ซึ่งเมื่อจัดรูปแล้วจะได้ว่า $u_{0j} \mid \underline{y}, \sigma_\varepsilon^2, \sigma_u^2, \gamma_{00} \sim Normal(\hat{u}_j, \hat{D}_j)$

โดยที่ $\hat{D}_j = \left(\frac{n_j}{\sigma_\varepsilon^2} + \frac{1}{\sigma_u^2} \right)^{-1}$ และ $\hat{u}_{0j} = \dfrac{\hat{D}_j \cdot \sum_i (y_{ij} - \gamma_{00})}{\sigma_\varepsilon^2}$

<u>ขั้นที่ 3</u>: สุ่มตัวอย่างพารามิเตอร์ $1/\sigma_u^2$ จาก $p(1/\sigma_u^2 \mid \underline{y}, \sigma_\varepsilon^2, \underline{u}, \gamma_{00})$

เนื่องจาก

$$p(1/\sigma_u^2 \mid \underline{y}, \sigma_\varepsilon^2, \underline{u}, \gamma_{00}) \propto \prod_{j=1}^{J} \left(\frac{1}{\sigma_u^2} \right)^{1/2} \cdot \exp\left\{ -\frac{u_j^2}{2\sigma_u^2} \right\} \cdot \left(\sigma_u^2 \right)^{-(\nu_u/2+1)} \exp\left\{ -\frac{\nu_u S_u^2}{2\sigma_u^2} \right\}$$

เมื่อจัดรูปแล้วจะได้ว่า $1/\sigma_u^2 \mid \underline{y}, \sigma_\varepsilon^2, \underline{u}, \gamma_{00} \sim Gamma\left(\frac{J + \nu_u}{2}, \frac{1}{2}\left(\nu_u S_u^2 + \sum_j u_j^2 \right) \right)$

<u>ขั้นที่ 4</u>: สุ่มตัวอย่างพารามิเตอร์ $1/\sigma_\varepsilon^2$ จาก $p(1/\sigma_\varepsilon^2 \mid \underline{y}, \sigma_u^2, \underline{u}, \gamma_{00})$

เนื่องจาก

$$p(1/\sigma_\varepsilon^2 \mid \underline{y}, \sigma_u^2, \underline{u}, \gamma_{00}) \propto \prod_{i,j} \left(\frac{1}{\sigma_\varepsilon^2} \right)^{1/2} \cdot \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(y_{ij} - u_j - \gamma_{00} \right)^2 \right\} \cdot \left(\sigma_\varepsilon^2 \right)^{-(\nu_\varepsilon/2+1)} \exp\left\{ -\frac{\nu_\varepsilon S_\varepsilon^2}{2\sigma_\varepsilon^2} \right\}$$

โดยที่ $1/\sigma_\varepsilon^2 \mid \underline{y}, \sigma_u^2, \underline{u}, \gamma_{00} \sim Gamma\left(\frac{N + \nu_\varepsilon}{2}, \frac{1}{2}\left(\nu_\varepsilon S_\varepsilon^2 + \sum_j \sum_i e_{ij}^2 \right) \right)$

อัลกอริทึมในข้างต้นจะกระทำแบบทวนซ้ำ (iterate) ไปเรื่อยๆ จนกระทั่งตัวอย่างที่สุ่มจากการแจกแจงความน่าจะเป็นในข้างต้นจะมีจำนวนเพียงพอที่จะลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ในโมเดลที่ต้องการ หลังจากสุ่มตัวอย่างพารามิเตอร์ในจำนวนรอบที่เพียงพอแล้วผู้วิจัยจะดำเนินการวิเคราะห์ในขั้นตอนที่ 3 คือการวิเคราะห์การแจกแจงความน่าจะ

เป็นภายหลัง และขั้นตอนที่ 4 คือการอนุมานเกี่ยวกับพารามิเตอร์จากการแจกแจงความน่าจะเป็นภายหลังที่ได้ต่อไป

3.3 <u>โมเดลสัมประสิทธิ์ความถดถอยแบบสุ่มแบบ 2 ระดับอย่างง่าย</u> (simple two-levels random coefficients model)

การประมาณค่าพารามิเตอร์ในกรณีนี้สามารถขยายจากแนวคิดที่ใช้ในโมเดลว่างแบบ 2 ระดับในหัวข้อที่แล้ว รายละเอียดมีดังต่อไปนี้

<u>ขั้นตอนที่1</u>: ขั้นตอนการสร้างโมเดล (model building)

โมเดลที่ต้องการประมาณค่าพารามิเตอร์มีรายละเอียดดังสมการที่ (2.96) ถึง (2.97) ดังนี้

<u>โมเดลระดับที่ 1</u>

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + \varepsilon_{ij} \qquad (2.96)$$

<u>โมเดลระดับที่ 2</u>

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (2.97)$$

$$\beta_{1j} = \gamma_{01} + u_{1j} \qquad (2.98)$$

<u>โมเดลรวม</u>

$$y_{ij} = \gamma_{00} + \gamma_{01} x_{ij} + u_{0j} + u_{1j} x_{ij} + \varepsilon_{ij} \qquad (2.99)$$

เมื่อ $\varepsilon_{ij} \sim Normal(0, \sigma_\varepsilon^2)$ และ $\underline{u}_j = (u_{oj}, u_{1j})^T \sim MVN(0, V_u)$, $i = 1,2,...,n_j$ และ $j = 1,2,...,J$ และเพื่อความสะดวกจะใช้สัญลักษณ์ X_{ij} แทนเวกเตอร์ $(X_{0ij}, X_{1ij})^T$ เมื่อ $X_{0ij} = 1$ และ $X_{1ij} = x_{ij}$

ในการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้านั้นจะกำหนดให้ พารามิเตอร์ที่เป็นอิทธิพลคงที่ $\underline{\gamma} = (\gamma_{00}, \gamma_{01})^T$ มีการแจกแจงความน่าจะเป็นก่อนหน้าแบบสม่ำเสมอ ($p(\underline{\gamma}) \propto 1$) พารามิเตอร์ความแปรปรวน σ_ε^2 กำหนดให้มีการแจกแจงแบบ scaled inverse chi-square ที่มีพารามิเตอร์ ν_ε กับ s_ε^2 และพารามิเตอร์เมทริกซ์ความแปรปรวนร่วม $V_u = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{bmatrix}$ กำหนดให้มีการแจกแจงแบบ Wishart ที่มีพารามิเตอร์ ν_p กับ S_p

<u>ขั้นตอนที่2</u>: ขั้นตอนการคำนวณการแจกแจงความน่าจะเป็นภายหลัง (calculation of the posterior distribution)

เมื่อกำหนดโมเดลและการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์เรียบร้อยแล้ว ขั้นตอนต่อไปคือการประมาณค่าพารามิเตอร์ในโมเดล ซึ่งจะมีทั้งหมด 4 ขั้นตอนเช่นเดียวกับโมเดลว่างแบบ 2 ระดับในข้างต้น เหตุผลเนื่องจากพารามิเตอร์ในทั้งสองโมเดลมีพารามิเตอร์ชนิดเดียวกันเพียงแต่โมเดลสัมประสิทธิ์ความถดถอยสุ่มจะมีความซับซ้อนมากกว่า จำนวนพารามิเตอร์จึงมากกว่า ดังนั้นการสุ่มตัวอย่างในแต่ละขั้นตอนจะเปลี่ยนเป็นการสุ่มตัวอย่างของเวกเตอร์ของพารามิเตอร์จากการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขแทน อัลกอริทึมที่ใช้จะเป็นอัลกอริทึมการสุ่มตัวอย่างแบบกิบบ์เช่นเดิม รายละเอียดเป็นดังนี้ (Browne, 1998)

<u>ขั้นที่ 1</u>: สุ่มตัวอย่างพารามิเตอร์ $\underline{\gamma}$ จาก $p(\underline{\gamma} \mid \underline{y}, \sigma_\varepsilon^2, V_u, \underline{u})$

เนื่องจาก $p(\underline{\gamma} \mid \underline{y}, \sigma_\varepsilon^2, \sigma_u^2, \underline{u}) \propto \prod_{i,j} \left(\frac{1}{\sigma_\varepsilon^2} \right)^{1/2} \cdot \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(y_{ij} - X_{ij}\underline{u}_j - X_{ij}\underline{\gamma} \right)^2 \right\}$

ซึ่งเมื่อจัดรูปแล้วจะได้ว่า $\underline{\gamma} \mid \underline{y}, \sigma_\varepsilon^2, V_u, \underline{u} \sim MVN(\hat{\underline{\gamma}}, \hat{D})$

โดยที่ $\hat{D} = \sigma_\varepsilon^2 \left[\sum_{i,j} X_{ij}^T X_{ij} \right]^{-1}$ และ $\hat{\underline{\gamma}} = \frac{\hat{D}}{\sigma_\varepsilon^2} \sum_{i,j} X_{ij}^T (y_{ij} - X_{ij}\underline{u}_j)$

<u>ขั้นที่ 2</u>: สุ่มตัวอย่าง $\underline{u}_j$ จาก $p(u_{0j} \mid \underline{y}, \sigma_\varepsilon^2, V_u, \underline{\gamma})$

เนื่องจาก

$$p(\underline{u}_j \mid \underline{y}, \sigma_\varepsilon^2, V_u, \underline{\gamma}) \propto \prod_{i=1}^{n_j} \left(\frac{1}{\sigma_\varepsilon^2} \right)^{1/2} \cdot \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(y_{ij} - X_{ij}u_j - X_{ij}\underline{\gamma} \right)^2 \right\} \cdot |V_u|^{-\frac{1}{2}} \cdot \exp\left\{ -\frac{1}{2} \underline{u}_j V_u^{-1} \underline{u}_j \right\}$$

ซึ่งเมื่อจัดรูปแล้วจะได้ว่า $\underline{u}_j \mid \underline{y}, \sigma_\varepsilon^2, V_u, \underline{\gamma} \sim MVN(\hat{u}_j, \hat{D}_j)$

โดยที่ $\hat{D}_j = \left(\frac{\sum_{i=1}^{n_j} X_{ij}^T X_{ij}}{\sigma_\varepsilon^2} + V_u^{-1} \right)^{-1}$ และ $\hat{u}_{0j} = \frac{\hat{D}_j}{\sigma_\varepsilon^2} \sum_{i=1}^{n_j} X_{ij}^T (y_{ij} - X_{ij}\underline{\gamma})$

<u>ขั้นที่ 3</u>: สุ่มตัวอย่างพารามิเตอร์ V_u จาก $p(V_u \mid \underline{y}, \sigma_\varepsilon^2, \underline{u}, \underline{\gamma})$

เนื่องจาก $p(V_u \mid \underline{y}, \sigma_\varepsilon^2, \underline{u}, \underline{\gamma}) \propto \prod_{j=1}^{J} |V_u|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}\underline{u}_j^T V_u^{-1} \underline{u}_j\right\} \cdot p(V_u^{-1})$

โดยที่ $p(V_u^{-1}) \propto \left|V_u^{-1}\right|^{(\nu_p-3)/2} \exp\left\{-\frac{1}{2} tr(S_p^{-1} V_u^{-1})\right\}$

เมื่อจัดรูปแล้วจะได้ว่า $V_u \mid \underline{y}, \sigma_\varepsilon^2, \underline{u}, \underline{\gamma} \sim Wishart\left(J + \nu_p, \left(\sum_{j=1}^{J} \underline{u}_j^T \underline{u}_j + S_p^{-1} \right)^{-1} \right)$

<u>ขั้นที่ 4</u>: สุ่มตัวอย่างพารามิเตอร์ $1/\sigma_\varepsilon^2$ จาก $p(1/\sigma_\varepsilon^2 \mid \underline{y}, V_u, \underline{u}, \underline{\gamma})$

เนื่องจาก

$$p(1/\sigma_\varepsilon^2 \mid \underline{y}, V_u, \underline{u}, \underline{\gamma}) \propto \prod_{i,j} \left(\frac{1}{\sigma_\varepsilon^2}\right)^{1/2} \cdot \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(y_{ij} - X_{ij} u_j - X_{ij} \underline{\gamma}\right)^2\right\} \cdot \left(\frac{1}{\sigma_\varepsilon^2}\right)^{-2} p(\sigma_\varepsilon^2)$$

โดยที่ $1/\sigma_\varepsilon^2 \mid \underline{y}, \sigma_u^2, \underline{u}, \gamma_{00} \sim Gamma\left(\frac{N + \nu_\varepsilon}{2}, \frac{1}{2}\left(\nu_\varepsilon S_\varepsilon^2 + \sum_j \sum_i e_{ij}^2\right)\right)$

เมื่อทำการสุ่มตัวอย่างค่าสังเกตของพารามิเตอร์ด้วยอัลกอริทึมของกิบบ์ตามจำนวนที่ได้กำหนดไว้แล้ว ผู้วิจัยจะดำเนินการวิเคราะห์การแจกแจงความน่าจะเป็นภายหลังและอนุมานพารามิเตอร์โมเดลโดยใช้ค่าสังเกตที่ได้จากขั้นตอนที่ 2 ต่อไป

3.4 <u>โมเดลเชิงเส้นพหุระดับในกรณีทั่วไป (general multi-level model)</u>

การประยุกต์ใช้แนวคิดของเบส์เซียนในการประมาณค่าพารามิเตอร์ในโมเดลพหุระดับในกรณีทั่วไปนั้น สามารถทำได้โดยอาศัยแนวคิดเดียวกับอัลกอริทึมที่ใช้ในการประมาณค่าพารามิเตอร์ในโมเดลว่าง 2 ระดับที่ได้นำเสนอไปในหัวข้อที่แล้ว สิ่งที่แตกต่างกันคือในโมเดลทั่วไปนั้นส่วนประกอบของโมเดลในส่วนที่เป็นอิทธิพลสุ่มนั้นจะมีความซับซ้อนทำให้การกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของส่วนประกอบความแปรปรวน และการแจกแจงความน่าจะเป็น

ภายหลังของส่วนประกอบความแปรปรวนนั้นจำเป็นที่จะต้องใช้การแจกแจงความน่าจะเป็นของเมทริกซ์เข้ามาช่วย รายะเอียดเกี่ยวกับอัลกอริทึมในการประมาณผู้ที่สนใจสามารถศึกษาได้จากเอกสารอ้างอิงที่ได้ระบุไว้ (Browne and Draper, 2000; Browne, 1998)

การประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับในกรณีทั่วไปที่มี N ระดับ ก่อนการพัฒนาอัลกอริทึมในการประมาณนั้นจำเป็นที่จะต้องแบ่งประเภทของพารามิเตอร์ในโมเดลก่อนเช่นเดียวกับอัลกอริทึมในโมเดลที่ได้นำเสนอไปในข้างต้น จากการพิจารณาพบว่าจะสามารถแบ่งประเภทของพารามิเตอร์ในโมเดลออกได้เป็น 4 ประเภทได้แก่ 1) พารามิเตอร์ที่เป็นอิทธิพลคงที่ (fixed effects parameters) 2) เศษเหลือ (residual) ใน N-1 ระดับ (ยกเว้นเศษเหลือในระดับที่ 1) 3) พารามิเตอร์ความแปรปรวนในระดับที่ 1 (level-1 variance parameter) และ 4) พารามิเตอร์ความแปรปรวนและความแปรปรวนร่วม N-1 ระดับที่เหลือจากระดับที่ 1 (N-1 higher level variance-covariance parameters)

ในหัวข้อนี้จะกำหนดสัญลักษณ์ที่ใช้ใหม่ เพื่อให้มีความสอดคล้องกับลักษณะของโมเดลดังนี้ กำหนดให้

M_T แทนจำนวนค่าสังเกตทั้งหมดในระดับที่ 1

$M_{l,j}$ แทนจำนวนค่าสังเกตในระดับที่ l กลุ่มที่ j

ยกตัวอย่างเช่น ในโมเดลอย่างง่าย 2 ระดับ M_T อาจเป็นจำนวนนักเรียนทั้งหมดในทุกโรงเรียนที่ได้เก็บรวบรวมข้อมูลมา และ $M_{2,j}$ คือจำนวนนักเรียนทั้งหมดในโรงเรียนที่ j

X_{li} แทนเวกเตอร์ของตัวแปรในระดับที่ l ของค่าสังเกตที่ i

$\beta_{l,j}$ แทนพารามิเตอร์ที่เป็นอิทธิพลสุ่มในระดับที่ l กลุ่มที่ j (ในกรณีที่ $l=1$ จะไม่เขียน supscript j เนื่องจากมีความหมายถึงพารามิเตอร์ที่เป็นอิทธิพลคงที่)

และกำหนดให้ $d_{lj} = e_i - X_{li}\beta_{lj}$

ในการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดลจะกำหนดให้ พารามิเตอร์ที่เป็นอิทธิพลคงที่มีการแจกแจงแบบปกติหลายตัวแปรที่มีเวกเตอร์ค่าเฉลี่ย $\underline{\mu}_p$ และเมทริกซ์ความแปรปรวนร่วม S_p $(\beta_1 \sim MVN(\underline{\mu}_p, S_p))$ พารามิเตอร์ความแปรปรวนในระดับที่ 1 มีการแจกแจงแบบสเกลอินเวอรส์ไคสแควร์ที่มีพารามิเตอร์ ν_ε และ s_ε

$(\sigma^2 \sim SI_{\chi^2}(\nu_\varepsilon, s_\varepsilon))$ และพารามิเตอร์เมทริกซ์ความแปรปรวนร่วมในระดับที่ l เขียนแทนด้วย V_l กำหนดให้มีการแจกแจงแบบอินเวอร์สวิชชาตที่มีพารามิเตอร์ ν_{Pl} และ S_{Pl} $(V_l \sim IW(\nu_{Pl}, S_{Pl}))$

อัลกอริทึมที่ใช้ในการประมาณค่าพารามิเตอร์ประกอบไปด้วย 4 ขั้นตอนเช่นเดียวกับอัลกอริทึมที่ใช้กับโมเดลที่ได้นำเสนอไว้ในก่อนหน้านี้

<u>ขั้นที่ 1</u>: สุ่มตัวอย่างพารามิเตอร์ที่เป็นอิทธิพลคงที่ β_1

เนื่องจาก $$p(\beta_1 \mid y,...) \propto p(y \mid \beta_1,...)p(\beta_1)$$

จะได้ว่า $$\beta_1 \sim MVN(\hat{\beta}_1, \hat{D}_1)$$

โดยที่ $\hat{D}_1 = \left[\sum_{i \in M_T} \frac{X_{1i}^T X_{1i}}{\sigma^2} + S_p^{-1}\right]^{-1}$ และ $\hat{\beta}_1 = \hat{D}_1\left[\sum_{i \in M_T} \frac{X_{1i}^T d_{1i}}{\sigma^2} + S_p^{-1}\underline{\mu}_p\right]$

<u>ขั้นที่ 2</u>: สุ่มตัวอย่างเศษเหลือในระดับที่ l β_l

เนื่องจาก $$p(\beta_l \mid y,...) \propto p(y \mid \beta_l,...)p(\beta_l \mid V_l)$$

จะได้ว่า $$\beta_{lj} \sim MVN(\hat{\beta}_{lj}, \hat{D}_{lj})$$

โดยที่ $\hat{D}_{lj} = \left[\sum_{i \in M_{l,j}} \frac{X_{1i}^T X_{1i}}{\sigma^2} + V_l^{-1}\right]^{-1}$ และ $\hat{\beta}_{lj} = \frac{\hat{D}_{lj}}{\sigma^2}\sum_{i \in M_{l,j}} X_{li}^T d_{li}$

<u>ขั้นที่ 3</u>: สุ่มตัวอย่างพารามิเตอร์ความแปรปรวนในระดับที่ 1 σ^2

เนื่องจาก $$p(1/\sigma^2 \mid y,...) \propto p(y \mid \sigma^2,...)p(1/\sigma^2)$$

จะได้ว่า $$1/\sigma^2 \sim Gamma(a_{pos}, b_{pos})$$

โดยที่ $a_{pos} = \frac{1}{2}(N + \nu_\varepsilon)$ และ $b_{pos} = \frac{1}{2}\left(\sum e_n^2 + \nu_\varepsilon s_\varepsilon^2\right)$

<u>ขั้นที่ 4</u>: สุ่มตัวอย่างพารามิเตอร์เมทริกซ์ความแปรปรวนร่วมในระดับที่ l V_l

เนื่องจาก $$p(V_l \mid y,...) \propto p(\beta_l \mid V_l)p(V_l^{-1})$$

จะได้ว่า $$V_l^{-1} \sim Wishart_{nrl}(S_{pos} = (\sum_{i=1}^{n_l} \beta_{li}\beta_{li}^T + S_{Pl})^{-1}, \nu_{pos} = n_l + \nu_{Pl})$$

เมื่อ n_l เป็นจำนวนหน่วยตัวอย่างในระดับที่ l

ตอนที่ 4 แนวคิดทฤษฎีที่เกี่ยวกับความคลาดเคลื่อนจากการวัด

การวัดในทางจิตวิทยาและการศึกษานั้น มีธรรมชาติที่สำคัญ 4 อย่างได้แก่ 1) เป็นการวัดหรือสังเกตทางอ้อม (indirect) 2) การวัดหรือสังเกตแต่ละครั้งเป็นการรวบรวมข้อมูลเพียงบางส่วนของพฤติกรรม 3) ผลที่ได้จากการวัดเป็นคุณลักษณะในเชิงสัมพัทธ์ และ 4) การวัดมีความคลาดเคลื่อน (measurement errors) เกิดขึ้นเสมอ (ศิริชัย กาญจนวาสี, 2550) จากข้อความในข้างต้นจะเห็นว่าในงานวิจัยทางจิตวิทยา ทางการศึกษา รวมทั้งทางสังคมศาสตร์นั้นการวัดค่าตัวแปรต่างๆที่สนใจนั้นมักมีความคลาดเคลื่อนในการวัดเกิดขึ้นได้เสมอ สาเหตุของการเกิดความคลาดเคลื่อนในการวัด เช่น เกิดจากเครื่องมือที่ใช้ในการวัดไม่มีความน่าเชื่อถือ การนิยามตัวแปรที่ไม่ดี เกิดจากตัวผู้ตอบเอง หรือเกิดจากขั้นตอนในการจัดระทำข้อมูล เป็นต้น จากตัวอย่างจะพบว่าความคลาดเคลื่อนในการวัดนั้นสามารถแบ่งออกได้เป็น 2 ลักษณะคือความคลาดเคลื่อนในการวัดแบบเป็นระบบ (systematic measurement errors) และความคลาดเคลื่อนในการวัดแบบสุ่ม (random measurement errors) ซึ่งต่อไปนี้หากไม่กล่าวเป็นอื่นความคลาดเคลื่อนในการวัดที่กล่าวถึงในรายงานฉบับนี้จะเป็นความคลาดเคลื่อนในการวัดแบบสุ่มเท่านั้น ความคลาดเคลื่อนในการวัดแบบสุ่มนั้นยังสามารถแบ่งออกได้อีกเป็น 2 ประเภท (Goldstein, Kounail, and Robinson, 2008) คือ 1) กรณีตัวแปรสังเกตได้เป็นตัวแปรต่อเนื่องและความคลาดเคลื่อนในการวัดมีการแจกแจงความน่าจะเป็นแบบต่อเนื่อง และ 2) กรณีที่เป็นความคลาดเคลื่อนในการจัดกลุ่มผิด (misclassification errors) ซึ่งสามารถเกิดขึ้นได้ในเหตุการณ์ที่ในการจัดหน่วยตัวอย่างเข้ากลุ่ม (categories) มีความน่าจะเป็นที่จะจัดกลุ่มหน่วยเข้าผิดกลุ่ม โดยในรายงานฉบับนี้จะสนใจพิจารณาเฉพาะปัญหาความคลาดเคลื่อนจากการวัดของตัวแปรสังเกตได้ที่เป็นตัวแปรต่อเนื่องเท่านั้น

4.1 <u>ทฤษฎีการวัดแบบดั้งเดิม</u> (classical test theory)

แนวคิดของทฤษฎีการวัดแบบดั้งเดิม (classical test theory) เชื่อว่าหากสามารถสังเกตค่าซ้ำดังนี้ $x_1^o, x_2^o, ..., x_t^o$ ความคลาดเคลื่อนในการวัดประเภทแรกนั้นสามารถอธิบายได้โดยโมเดลดังต่อไปนี้

$$x_i^o = x + e_i \qquad (2.100)$$

เมื่อ x_i^o คือตัวแปรสังเกตได้ (observed variable) x คือตัวแปรที่แท้จริง (true variable) และ e_i คือความคลาดเคลื่อนในการวัด โดยที่ $E(e_i) = 0, Var(x) = \sigma_x^2, Var(e_i) = \sigma_{ex}^2$ และ $Cov(x, e_i) = 0$ ดังนั้นจะได้ว่าความแปรปรวนของตัวแปรสังเกตได้มีค่าเท่ากับ

$$\sigma_{x^o}^2 = \sigma_x^2 + \sigma_{ex}^2 \qquad (2.101)$$

ค่าสถิติที่ใช้บ่งบอกถึงปริมาณความคลาดเคลื่อนในการวัดตามแนวคิดของทฤษฎีการวัดแบบดั้งเดิมคือ "ความเที่ยง (reliability)" ซึ่งเป็นค่าที่ใช้วัดสัดส่วนของความผันแปรในตัวแปรจริงที่ตัวแปรสังเกตได้สามารถอธิบายได้ หรืออาจกล่าวให้ง่ายขึ้นว่า ความเที่ยงคือค่าที่ใช้วัดความใกล้ชิดกันระหว่างตัวแปรจริงกับตัวแปรสังเกตได้นั่นเอง จากการนิยามความเที่ยงในข้างต้นสามารถเขียนให้อยู่ในรูปสมการทางคณิตศาสตร์ได้ดังนี้ (ศิริชัย กาญจนวาสี, 2550; Goldstein, Kounali, and Robinson, 2008)

$$\rho_{x^o x}^2 = \frac{\sigma_x^2}{\sigma_{x^o}^2} = 1 - \frac{\sigma_{ex}^2}{\sigma_{x^o}^2} \qquad (2.102)$$

เมื่อ $\rho_{xx^T}^2$ เป็นค่าความเที่ยง (reliability) โดยที่ $0 \leq \rho_{xx^T}^2 \leq 1$

4.2 <u>ผลกระทบจากความคลาดเคลื่อนในการวัด</u>

ค่าความเที่ยงที่ต่ำนั้นแสดงว่าตัวแปรสังเกตได้ตัวนั้นมีความคลาดเคลื่อนไปจากตัวแปรจริงมากหรือมีความคลาดเคลื่อนในการวัดสูง ซึ่งจะทำให้ผลการวิเคราะห์โดยใช้ตัวแปรดังกล่าวนั้นมีความผิดพลาดไปจากความเป็นจริง (ศิริชัย กาญจนวาสี, 2550; ไพฑูรย์ ไกรพรศักดิ์, 2548; Kutner M.H., Nachtsheim C.J., Neter J. and Li W., 2005; Browne, Goldstein, Woodhouse and Yang, 2002; Raudenbush and Bryk, 2002) และเพื่อให้สามารถเข้าใจได้ง่ายจึงขอยกตัวอย่างในกรณีโมเดลการวิเคราะห์การถดถอย (regression analysis) เพื่อให้เห็นผลกระทบของการนำตัวแปรที่มีความคลาดเคลื่อนในการวัดไปใช้ในการวิเคราะห์ ซึ่งตัวอย่างดังกล่าวยังสามารถใช้อธิบายปรากฏการณ์ที่เกิดความคลาดเคลื่อนในการวัดในโมเดลเชิงเส้นพหุระดับได้ด้วย สมมติว่านักวิจัยต้องการใช้การวิเคราะห์การถดถอยในการอธิบายความสัมพันธ์ระหว่างตัวแปรตาม และตัวแปรอิสระ จะได้ว่าโมเดลการถดถอยที่นักวิจัยใช้จะมีสมการดังนี้

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \qquad (2.103)$$

เมื่อ $\underline{Y}$ เป็นเวกเตอร์ของตัวแปรตาม X เป็นเมทริกซ์ของตัวแปรอิสระ $\underline{\beta}$ เป็นเวกเตอร์ของพารามิเตอร์สัมประสิทธิ์ความถดถอย และ $\underline{\varepsilon}$ เป็นเวกเตอร์ความคลาดเคลื่อนของโมเดล

ความคลาดเคลื่อนจากการวัดนั้นสามารถเกิดได้ทั้งตัวแปรตาม และตัวแปรอิสระ ซึ่งผลกระทบที่เกิดขึ้นจะมีความแตกต่างกัน สมมติว่าความคลาดเคลื่อนจากการวัดเกิดขึ้นในตัวแปรตามและกำหนดให้โมเดลที่ใช้ในการอธิบายธรรมชาติของความคลาดเคลื่อนจากการวัดเป็นโมเดลการวัดแบบดั้งเดิมจะได้ว่า โมเดลการวัดค่าตัวแปรตามเป็นดังสมการ

$$\underline{Y}^o = \underline{Y} + \underline{\varepsilon}_y \tag{2.104}$$

เมื่อ $\underline{Y}^o$ คือเวกเตอร์ของตัวแปรตามที่สังเกตค่าได้ และ $\underline{\varepsilon}_y$ คือเวกเตอร์ของความคลาดเคลื่อนในการวัดตัวแปรตาม จากความคลาดเคลื่อนที่เกิดขึ้นในตัวแปรตามดังกล่าว โมเดลการถดถอยที่แท้จริงที่ควรใช้ในการวิเคราะห์ข้อมูลควรมีรูปแบบดังสมการ

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} - \underline{\varepsilon}_y \tag{2.105}$$

หรือเขียนได้เป็น $$\underline{Y} = X\underline{\beta} + \underline{u} \qquad \text{เมื่อ } \underline{u} = \underline{\varepsilon} - \underline{\varepsilon}_y \tag{2.106}$$

ถ้ากำหนดข้อตกลงเบื้องต้นให้ความคลาดเคลื่อนในโมเดล $\underline{\varepsilon}$ มีคุณสมบัติตามข้อตกลงดั้งเดิมของการวิเคราะห์การถดถอยทุกประการ และความคลาดเคลื่อนในการวัด $\underline{\varepsilon}_y$ มีคุณสมบัติ $E(\underline{\varepsilon}_y) = 0$, $Cov(\underline{\varepsilon}_y) = \sigma^2_{\varepsilon y} \cdot I_n$, $Cov(\underline{\varepsilon}_y, \underline{\varepsilon}) = 0$ และ $Cov(\underline{\varepsilon}_y, X) = 0$ หากใช้การประมาณค่าพารามิเตอร์โดยที่ไม่พิจารณาเทอมของความคลาดเคลื่อนจากการวัดด้วย เช่น การใช้ตัวประมาณกำลังสองน้อยสุด (ordinary least squares: OLS) หรือตัวประมาณภาวะความควรจะเป็นสูงสุดจะได้ว่าค่าประมาณสัมประสิทธิ์ความถดถอย คือ

$$\begin{aligned}\underline{\hat{\beta}} &= (X^T X)^{-1}(X^T \underline{Y}) = (X^T X)^{-1}(X^T(\underline{Y} + \underline{\varepsilon}_y)) \\ &= (X^T X)^{-1} X^T \underline{Y} + (X^T X)^{-1} X^T \underline{\varepsilon}_y \end{aligned} \tag{2.107}$$

จากข้อตกลงเบื้องต้นทำให้ได้ว่าค่าคาดหวังของตัวประมาณสัมประสิทธิ์ความถดถอยมีค่าเท่ากับ

$$E(\underline{\hat{\beta}}) = \underline{\beta} \tag{2.108}$$

และความแปรปรวนของตัวประมาณมีค่าเท่ากับ

$$Var(\underline{\hat{\beta}}) = (\sigma^2_{\varepsilon} + \sigma^2_{\varepsilon y}) \cdot (X^T X)^{-1} \tag{2.109}$$

จากสมการที่ (2.108) และ (2.109) ทำให้สามารถสรุปได้ว่าเมื่อเกิดปัญหาความคลาดเคลื่อนในการวัดค่าตัวแปรตาม ค่าประมาณของสัมประสิทธิ์ความถดถอยยังสามารถประมาณค่าได้โดยที่ตัวประมาณยังมีคุณสมบัติความไม่ลำเอียง (unbiased) อยู่ แต่อย่างไรก็ตามค่าคลาดเคลื่อนมาตรฐานของตัวประมาณสัมประสิทธิ์ความถดถอยจะมีค่ามากขึ้นกว่าที่ควรจะเป็น ซึ่งจะทำให้ในการอนุมานเชิงสถิติเกี่ยวกับพารามิเตอร์ดังกล่าวจะมีความคลาดเคลื่อน กล่าวคือ ในการประมาณค่าแบบช่วงจะทำให้ช่วงความเชื่อมั่นของพารามิเตอร์สัมประสิทธิ์ความถดถอยดังกล่าวมีความยาวช่วงที่กว้างมากกว่าปกติ และในการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ในโมเดลนั้นจะมีโอกาสที่จะปฏิเสธสมมติฐานหลักเพิ่มน้อยลง (หรือทำให้อำนาจการทดสอบของตัวสถิติทดสอบมีน้อยลง)

ในกรณีที่เกิดความคลาดเคลื่อนจากการวัดในตัวแปรอิสระและกำหนดให้อธิบายความคลาดเคลื่อนจากการวัดดังกล่าวด้วยโมเดลการวัดแบบดั้งเดิมดังนี้

$$X^{o} = X + E_{x} \qquad (2.110)$$

เมื่อ X^{o} เป็นเมทริกซ์ของตัวแปรอิสระที่สังเกตค่าได้ และ E_{x} เป็นเมทริกซ์ของความคลาดเคลื่อนในการวัดของตัวแปรอิสระ โมเดลการถดถอยที่แท้จริงคือ

$$\underline{Y} = (X + \underline{E}_{x})\underline{\beta} + \underline{\varepsilon} = X\underline{\beta} + (E_{x}\underline{\beta} + \underline{\varepsilon}) = X\underline{\beta} + \underline{v} \qquad (2.111)$$

สมมติให้ความคลาดเคลื่อนในการวัดค่าตัวแปรอิสระมีคุณสมบัติ $E(E_{x}) = 0$ และ $Cov(E_{x}) = \Sigma = \sigma_{x}^{2} \cdot I$ จะได้ว่าตัวประมาณกำลังสองน้อยสุดของสัมประสิทธิ์ความถดถอยจะเป็น

$$\begin{aligned}\hat{\underline{\beta}} &= (X^{T}X)^{-1}(X^{T}\underline{Y}) = ((X + E_{x})^{T}(X + E_{x}))^{-1}((X + E_{x})^{T}\underline{Y}) \\ &= (X^{T}X + E_{X}^{T}X + X^{T}E_{X} + E_{X}^{T}E_{X})^{-1}(X\underline{Y} + E_{X}^{T}\underline{Y}) \qquad (2.112)\end{aligned}$$

จะเห็นว่า $E(\hat{\underline{\beta}}) \neq \underline{\beta}$ ดังนั้นแม้ว่าเราจะสมมติให้ความคลาดเคลื่อนในการวัดค่าตัวแปรอิสระ มีคุณสมบัติเช่นเดียวกับความคลาดเคลื่อนในการวัดค่าตัวแปรตามดังที่ได้กล่าวไว้ในข้างต้นก็ตาม แต่จะไม่สามารถได้สมมติว่า E_{x} เป็นอิสระกับ X ดังนั้นจึงเป็นการละเมิดข้อตกลงเบื้องต้นของการโมเดลการถดถอย ทำให้ตัวประมาณเกิดความลำเอียง (biased) และมีความไม่คงเส้นคงวา (inconsistency)(มณฑิรา ดวงสาพล, 2550; วุฒิพงษ์ เดโชดมพันธ์, 2546)

จากตัวอย่างที่ได้แสดงไว้ในข้างต้นจะเห็นได้ว่า ปัญหาความคลาดเคลื่อนในการวัดค่าที่เกิดขึ้นในตัวแปรอิสระนั้น เป็นปัญหาที่มีผลกระทบที่รุนแรงมากกว่าปัญหาความคลาดเคลื่อนในการวัดค่าตัวแปรตาม การใช้ตัวแปรอิสระที่ค่าความเที่ยงต่ำในการวิเคราะห์จะทำให้การประมาณค่าพารามิเตอร์ในโมเดล รวมทั้งค่าความคลาดเคลื่อนมาตรฐานมีความผิดพลาด ส่งผลให้การอนุมานเกี่ยวกับค่าพารามิเตอร์ในโมเดลนั้นมีความไม่น่าเชื่อถือหากไม่ได้มีการปรับแก้ค่าประมาณให้มีความถูกต้อง (มณฑิรา ดวงสาพล, 2550; วุฒิพงษ์ เดโชดมพันธ์, 2546; Goldstein, Kounail, and Robinson, 2008; George and Goldstein, 2007; Fox, 2005, 2003, 2001; Paris, Q., 2004; Browne, Goldstein, Woodhouse and Yang, 2002; Nounou et al., 2001; Browne, 1998)

4.3 <u>แนวทางการแก้ไขปัญหาความคลาดเคลื่อนจากการวัด</u>

การแก้ไขหรือการลดทอนปัญหาความคลาดเคลื่อนในการวัดที่เกิดขึ้นนั้น อาจสามารถแบ่งออกได้เป็น 3 แนวทาง แนวทางแรกคือ หากผู้วิจัยพิจารณาผลของการทดสอบเครื่องมือปรากฏว่ามีความเที่ยงอยู่ในระดับที่ต่ำ ผู้วิจัยสามารถแก้ไขได้โดยการปรับแก้เครื่องมือเพื่อให้มีค่าความเที่ยงสูงขึ้น แล้วจึงนำเครื่องมือที่ปรับแก้จนมีค่าความเที่ยงเป็นที่น่าพอใจแล้วไปใช้ในการวิเคราะห์ต่อไป อย่างไรก็ตามแนวทางนี้ถึงแม้จะสามารถกระทำได้ง่ายในทางปฏิบัติแต่ก็มีข้อเสียคือ ผู้วิจัยจะต้องเสียเวลารวมถึงทรัพยากรทั้งในการแก้ไขปรับเครื่องมือและในการนำเครื่องมือไปทดลองใช้ แนวทางที่สองคือ ใช้การปรับค่าตัวแปรที่มีความคลาดเคลื่อนในการวัดเพื่อลดทอนความคลาดเคลื่อนในการวัดของตัวแปรสังเกตได้ แนวทางนี้เป็นแนวทางที่ง่ายอีกแนวทางหนึ่ง ผู้วิจัยอาจใช้ค่าความเที่ยงมาช่วยในการปรับค่าตัวแปรสังเกตได้ดังนี้

$$x_{adji} = x_i^o \cdot \rho^2_{x^o x} + \overline{x^o} \cdot (1 - \rho^2_{x^o x}) \qquad (2.113)$$

สูตรดังกล่าวพบว่าเป็นการประยุกต์แนวคิดของการหาค่าเฉลี่ยแบบถ่วงน้ำหนักมาช่วยในการปรับคะแนนที่ได้จากตัวแปรสังเกตได้ให้มีความสอดคล้องกับคะแนนจากตัวแปรที่แท้จริงมากยิ่งขึ้น อย่างไรก็ตามการปรับค่าดังกล่าวถึงเป็นการแปลงค่าตัวแปรดังนั้นจึงควรพิจารณาว่าตัวแปรที่สร้างใหม่ยังมีคุณสมบัติต่างๆ เช่น ตำแหน่ง ความแปรปรวน ลักษณะความสัมพันธ์กับตัวแปรตาม ฯลฯ เหมือนเดิมหรือไม่ ในกรณีของสูตรปรับแก้ตามสมการ (2.113) จะพบว่าตำแหน่งและ

ลักษณะความสัมพันธ์ของตัวแปรปรับค่ายังมีลักษณะเหมือนเดิม กล่าวคือ $E\left(x_{adji}\right) = x = E\left(x_i^o\right)$ และ $Cov(x_{adji}, y) = Cov(x_i^o, y)$ แต่เมื่อพิจารณาความแปรปรวนของตัวแปรปรับค่าจะพบว่ามีค่าเท่ากับ

$$Var\left(x_{adji}\right) = \sigma_{x^o}^2 \left[\frac{\rho^4 (n+1) - (2\rho^2 - 1)}{n}\right] \tag{2.114}$$

จากสมการ (2.114) จะเห็นว่าไม่สามารถที่จะใช้ความแปรปรวนของตัวแปรที่ปรับค่าแล้วในการวิเคราะห์ได้ ดังนั้นในการนำตัวแปรที่ปรับค่าแล้วไปทำการวิเคราะห์ต่อไป เช่น การเปรียบเทียบค่าเฉลี่ย การวิเคราะห์การถดถอย หากไม่ได้มีการปรับค่าความแปรปรวนในสูตรของการวิเคราะห์แล้วจะทำให้ผลการวิเคราะห์ที่ได้มีความผิดพลาดไม่ตรงกับความเป็นจริง การปรับค่าตัวแปรตามแนวทางดังกล่าวจึงควรใช้แนวคิดของการปรับโมเมนต์ (moment reconstruction) เพื่อปรับโมเมนต์ต่างๆของตัวแปรให้มีคุณสมบัติคงเดิม รายละเอียดที่กล่าวมาในข้างต้นนำเสนอเพื่อชี้ให้เห็นว่าแนวทางการปรับความคลาดเคลื่อนในการวัดโดยใช้การปรับจากตัวแปรสังเกตได้โดยตรงนั้นเป็นแนวทางที่ไม่เหมาะสมนัก หากผู้วิจัยต้องการที่จะนำตัวแปรดังกล่าวไปทำการวิเคราะห์เพื่อตอบวัตถุประสงค์ของการวิจัยต่อไป

แนวทางสุดท้ายที่คือการใช้วิธีการทางสถิติเป็นเครื่องมือในการลดทอนความคลาดเคลื่อนในการวัด เพื่อให้เห็นภาพจึงขอยกตัวอย่างการแก้ไขปัญหาความคลาดเคลื่อนจากการวัดวิธีการหนึ่งซึ่งเป็นการแก้ไขปัญหาในโมเดลการถดถอยเชิงเส้นอย่างง่ายดังนี้ พิจารณาโมเดลการถดถอยเชิงเส้นอย่างง่าย

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{2.115}$$

ในกรณีที่มีความคลาดเคลื่อนจากการวัดเกิดขึ้นในตัวแปรอิสระ เราจะสามารถปรับค่าประมาณของความแปรปรวนในตัวแปรอิสระได้ดังนี้ $\sigma_x^2 = \rho_{XX^T}^2 \sigma_{x^o}^2$ และ $Cov(x, y) = Cov(x^o, y) = c_{xy}$

ดังนั้นจะได้ว่าตัวประมาณสัมประสิทธิ์ความถดถอยที่แท้จริงจะมีสูตรเป็น

$$\frac{c_{xy}}{\rho_{XX^T}^2 \sigma_{x^o}^2} = \frac{b_{obs}}{\rho_{XX^T}^2} \tag{2.116}$$

เมื่อ b_{obs} เป็นค่าประมาณสัมประสิทธิ์ความถดถอยที่ได้จากข้อมูลค่าสังเกต โดยยังไม่ได้มีการปรับความคลาดเคลื่อนจากการวัด

จากตัวอย่างข้างต้นจะเห็นว่าแนวทางในการแก้ไขปัญหาแนวทางนี้เป็นการออกแบบให้เทคนิคการวิเคราะห์มีการคำนึงถึงความคลาดเคลื่อนในการวัดด้วย การแก้ไขปัญหาความคลาดเคลื่อนในการวัดที่เกิดขึ้นในโมเดลเชิงเส้นระดับเดียว (single-level linear model) นั้นได้มีการศึกษาทั้งผลกระทบของปัญหาและพัฒนาวิธีการแก้ไขปัญหาในโมเดลดังกล่าวไว้จำนวนมาก เช่น วิธีผลรวมกำลังสองน้อยสุด (total least squares method (TLS)) วิธีสร้างจากโมเมนต์ (moment reconstruction method (MR)) วิธีกำลังสองน้อยสุดร่วม (joint least squares method (JLS)) วิธีกำลังสองน้อยสุดเชิงส่วน (partial least square (PLS)) และวิธีการถดถอยแบบองค์ประกอบหลัก (principle component regression (PCR)) ซึ่งรายละเอียดของวิธีการต่างๆ นั้นจะไม่นำเสนอในโครงร่างฉบับนี้เนื่องจากไม่ได้อยู่ในประเด็นที่สนใจจะพัฒนา ผู้สนใจสามารถศึกษารายละเอียดจากเอกสารอ้างอิงที่ระบุไว้ (มณฑิรา ดวงสาพล, 2550; วุฒิพงษ์ เดโชดมพันธ์, 2546; Paris, Q., 2004; Chen, Kreider, Merwin, and Stern, 2001; Nounou et al., 2001; Huffel and Vandewalle, 1991; Fuller, 1987; Degracie and Fuller, 1972; Joleskog, 1970)

ในงานวิจัยนี้สนใจศึกษาการแก้ไขปัญหาความคลาดเคลื่อนจากการวัดในโมเดลเชิงเส้นพหุระดับโดยเน้นเพื่อให้ใช้ในงานวิจัยทางสังคมศาสตร์ ซึ่งเมื่อพิจารณาการแก้ไขปัญหาความคลาดเคลื่อนจากการวัดในงานวิจัยทางสังคมศาสตร์แล้วจะพบว่าโดยปรกติแล้วการวิจัยทางสังคมศาตร์ได้ให้ความสำคัญกับการลดทอนหรือหลีกเลี่ยงปัญหาเกี่ยวกับความคลาดเคลื่อนจากการวัดพอสมควร กล่าวคือได้มีการคำนึงตั้งแต่การสร้างแบบวัด ซึ่งแบบวัดที่จะสามารถวัดตัวแปรแฝงได้ดีจะต้องมีการศึกษาทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้องเพื่อเป็นข้อมูลในการนิยามตัวแปรและพัฒนากรอบแนวคิดในการวิจัย จากนั้นผู้วิจัยจะนำนิยามของตัวแปรต่างๆที่ได้กำหนดไว้มาสร้างแบบวัด โดยให้มีตัวแปรบ่งชี้หลายตัวในการวัดตัวแปรแฝง การที่มีตัวแปรบ่งชี้หลายตัวทำให้ผู้วิจัยสามารถประมาณความเที่ยง (reliability) ในการวัดตัวแปรได้ นอกจากนี้ในงานวิจัยบางชิ้นยังเลือกวิธีการวิเคราะห์ที่ช่วยลดทอนความคลาดเคลื่อนจากการวัด เช่นการใช้โมเดลลิสเรล (LISREL) ในการวิเคราะห์ (Dunson, Palomo, & Bollen, 2005; Lee & Song, 2004; Joleskog, 1970) โมเดลดังกล่าว โมเดลลิสเรลมีแนวคิดในการลดทอนความคลาดเคลื่อนจากการวัดโดยการ

ใช้โมเดลการวัดมาช่วยแก้ไขปัญหา โมเดลการวัดที่ใช้ในโมเดลลิสเรลคือโมเดลการวิเคราะห์องค์ประกอบ ดังนั้นในส่วนต่อไปผู้วิจัยจะนำเสนอรายละเอียดโดยย่อของโมเดลการวิเคราะห์องค์ประกอบเพื่อเป็นพื้นฐานในการพัฒนาวิธีการประมาณต่อไป

<u>โมเดลการวิเคราะห์องค์ประกอบ (Factor Analysis Model)</u>

วิธีการหนึ่งที่ผู้วิจัยอาจใช้ในการลดทอนความคลาดเคลื่อนจากการวัดคือการใช้เทคนิคการวิเคราะห์องค์ประกอบเพื่อตรวจสอบความตรงและลดทอนความคลาดเคลื่อนในการวัดตัวแปรแฝง โดยการใช้ค่าคะแนนองค์ประกอบในการวิเคราะห์เพื่อตอบคำถามวิจัย โมเดลการวิเคราะห์องค์ประกอบสามารถแบ่งออกได้เป็น 2 ประเภทใหญ่ ประเภทแรกคือ โมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ (exploratory factor analysis) ซึ่งมีวัตถุประสงค์หลักคือ เพื่อค้นหาองค์ประกอบหรือตัวแปรแฝงจำนวนหนึ่งที่มีความสัมพันธ์กับตัวแปรสังเกตได้ โดยที่ตัวแปรแฝงดังกล่าวนั้นเป็นอิสระซึ่งกันและกัน และเพื่อศึกษาความสัมพันธ์ระหว่างองค์ประกอบกับตัวแปรสังเกตได้ ประเภทที่สองคือ โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) มีวัตถุประสงค์ที่แตกต่างออกไป โดยโมเดลดังกล่าวนักวิจัยใช้ในการตรวจสอบความตรงของทฤษฎีที่ใช้ในการวัดองค์ประกอบหรือตัวแปรแฝงที่ต้องการ กล่าวคือนักวิจัยสามารถกำหนดค่าพารามิเตอร์ในโมเดลก่อนการวิเคราะห์ได้ นอกจากนี้ยังสามารถทดสอบสมมติฐานของค่าพารามิเตอร์ต่างๆในโมเดลได้แก่ พารามิเตอร์ดังกล่าวประกอบไปด้วย ค่าน้ำหนักองค์ประกอบ (factor loading) ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างองค์ประกอบ (factor correlation) และความแปรปรวนของส่วนเฉพาะ (unique variance) ทดสอบสมมติฐานเกี่ยวกับความสอดคล้องของโมเดลที่นำเสนอกับข้อมูลเชิงประจักษ์ ซึ่งนักวิจัยสามารถเปรียบเทียบความเหมาะสมระหว่างโมเดลการวัดที่แตกต่างกันจากข้อมูลเชิงประจักษ์ชุดเดียวกันได้ ในส่วนนี้จะนำเสนอรายละเอียดเกี่ยวกับโมเดลการวิเคราะห์องค์ประกอบทั้งสองประเภทก่อน จากนั้นจะนำเสนอรายละเอียดเกี่ยวกับการประมาณค่าและการทดสอบสมมติฐานของพารามิเตอร์ในโมเดลทั้งสองเพื่อเป็นพื้นฐานความรู้เกี่ยวกับหลักการของวิธีการประมาณค่าพารามิเตอร์ที่ใช้ในงานวิจัยนี้

<u>โมเดลการวิเคราะห์</u>

โมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจเป็นโมเดลเชิงเส้นระหว่างเวกเตอร์สุ่มของตัวแปรสังเกตได้ที่ได้รับอิทธิพลจากองค์ประกอบหรือตัวแปรแฝง สามารถเขียนในรูปสมการทางคณิตศาสตร์ได้ดังนี้

$$\underline{y} = \underline{\mu} + \Lambda \cdot \underline{\xi} + \underline{\varepsilon} \tag{2.117}$$

เมื่อ $\underline{y}$ เป็นเวกเตอร์สุ่มของตัวแปรสังเกตได้ขนาด $p \times 1$

$\underline{\mu}$ เป็นเวกเตอร์สุ่มของพารามิเตอร์ค่าเฉลี่ยของแต่ละตัวแปรสังเกตได้ขนาด $p \times 1$

$\underline{\xi}$ เป็นเวกเตอร์สุ่มขององค์ประกอบที่เป็นตัวแปรแฝงขนาด $q \times 1$

Λ เป็นเมทริกซ์ของน้ำหนักองค์ประกอบขนาด $p \times q$

และ $\underline{\varepsilon}$ เป็นเวกเตอร์ของความคลาดเคลื่อนจากการวัดขนาด $p \times 1$ (อาจเรียกว่าเวกเตอร์ขององค์ประกอบเฉพาะ (unique factor))

ข้อสมมติเบื้องต้นของโมเดลได้แก่

1) การแจกแจงความน่าจะเป็นของเวกเตอร์สุ่ม $\underline{\xi}$ กำหนดให้มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับเวกเตอร์ $\underline{0}$ และมีเมทริกซ์ความแปรปรวนร่วมเท่ากับ I

2) การแจกแจงความน่าจะเป็นของเวกเตอร์สุ่ม $\underline{\varepsilon}$ กำหนดให้มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับเวกเตอร์ $\underline{0}$ และมีเมทริกซ์ความแปรปรวนร่วมเท่ากับ Ψ_ε โดยที่ Ψ_ε เป็นเมทริกซ์ทแยงมุม (diagonal matrix)

3) เวกเตอร์สุ่ม $\underline{\xi}$ เป็นอิสระซึ่งกันและกันกับเวกเตอร์ $\underline{\varepsilon}$

จากข้อสมมติเบื้องต้นจะได้ว่าเวกเตอร์ของตัวแปรสังเกตได้นั้นจะมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ $\Lambda \cdot \underline{\xi}$ และมีเมทริกซ์ความแปรปรวนร่วมของเวกเตอร์ของตัวแปรสังเกตได้คือ

$$Cov(\underline{y}) = \Sigma = \Lambda \cdot \Lambda^T + \Psi_\varepsilon \tag{2.118}$$

และจะได้ว่า $Cov(\underline{x}, \underline{\xi}) = \Lambda$ จะมีค่าเท่ากับสหสัมพันธ์ระหว่างตัวแปรสังเกตได้และองค์ประกอบ จากสมการเมทริกซ์ความแปรปรวนร่วมในข้างต้นจะเรียก $\Lambda \cdot \Lambda^T$ ว่าค่าความร่วม (communality) มีความหมายเป็นร้อยละของความผันแปร (variation) ในตัวแปรสังเกตได้แต่ละตัวที่สามารถอธิบายได้โดยองค์ประกอบหรือตัวแปรแฝงทั้งหมด และ Ψ_ε จะเรียกว่า ส่วนเฉพาะ (uniqueness) ซึ่งเป็นความผันแปรส่วนที่เหลือในตัวแปรสังเกตได้ที่ไม่สามารถอธิบายได้ด้วยองค์ประกอบ

โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันเป็นโมเดลที่ขยายมาจากโมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ ซึ่งรูปแบบของโมเดลเหมือนกับโมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ ดังในสมการที่ (2.117) นอกจากนี้การกำหนดข้อสมมติเบื้องต้นของโมเดลยังเหมือนกับข้อสมมติของโมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจยกเว้นข้อสมมติของพารามิเตอร์ $\underline{\xi}$ ที่กำหนดให้มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ $\underline{0}$ และมีเมทริกซ์ความแปรปรวนร่วมเท่ากับ Φ กล่าวคือ $\underline{\xi} \sim N(\underline{0}, \Phi)$ การกำหนดข้อสมมติดังกล่าวทำให้การวิเคราะห์องค์ประกอบเชิงยืนยันยอมให้องค์ประกอบหรือตัวแปรแฝงที่ต้องการศึกษานั้นมีความสัมพันธ์กันได้ จากการกำหนดข้อสมมติดังกล่าวจะทำให้ได้ว่าเมทริกซ์ความแปรปรวนร่วมของตัวแปรสังเกตได้จะเท่ากับ

$$Cov(\underline{y}) = \Sigma = \Lambda\Phi\Lambda^T + \Psi_\varepsilon \tag{2.119}$$

สิ่งที่แตกต่างของวิเคราะห์องค์ประกอบเชิงยืนยันคือการที่ผู้วิจัยสามารถกำหนดค่าของพารามิเตอร์ในโมเดลได้แก่ Λ, Φ และ Ψ_ε ก่อนการวิเคราะห์ได้ ซึ่งเปิดโอกาสให้ผู้วิจัยสามารถระบุที่จะเลือกประมาณค่าน้ำหนักองค์ประกอบบางตัวซึ่งเป็นการระบุว่าตัวแปรสังเกตได้กลุ่มใดใช้ในการวัดองค์ประกอบใด รวมไปถึงการกำหนดความสัมพันธ์ระหว่างองค์ประกอบ และยังสามารถประเมิน (assessment) และทดสอบความสอดคล้องกันระหว่างโมเดลที่ผู้วิจัยได้กำหนดขึ้นกับข้อมูลเชิงประจักษ์ (goodness of fit test) โดยการกำหนดค่าดังกล่าวควรที่จะต้องกำหนดโดยอาศัยพื้นฐานจากทฤษฎีหรืองานวิจัยต่างๆที่เกี่ยวข้อง

ปัญหาความระบุได้ (identification problem)

ก่อนการประมาณค่าพารามิเตอร์นั้นผู้วิจัยจำเป็นที่จะต้องคำนึงความระบุได้ (identification) ของโมเดลและค่าพารามิเตอร์ที่ประมาณด้วย ปัญหาความระบุได้ (identification problem) เป็นปัญหาจากการที่ค่าประมาณพารามิเตอร์ในโมเดลนั้นไม่ได้เป็นหนึ่งเดียว (unique) เราสามารถนิยามความระบุได้ของโมเดลดังนี้ กำหนดให้ $\underline{\theta}_1$ และ $\underline{\theta}_2$ เป็นเวกเตอร์ของพารามิเตอร์ในปริภูมิพารามิเตอร์ ถ้า $\underline{\theta}_1 \neq \underline{\theta}_2$ ทุกคู่เปรียบเทียบของเวกเตอร์ของพารามิเตอร์แล้ว $\Sigma(\underline{\theta}_1) \neq \Sigma(\underline{\theta}_2)$ จะได้ว่าโมเดลการวิเคราะห์ของผู้วิจัยเป็นโมเดลที่มีคุณสมบัติสามารถระบุได้ (whole model identification) กล่าวในอีกเชิงหนึ่งคือ ถ้าความสัมพันธ์ระหว่าง $\Sigma(\underline{\theta})$ กับพารามิเตอร์ $\underline{\theta}$ เป็นฟังก์ชันแบบหนึ่งต่อหนึ่ง จะกล่าวว่าโมเดลที่มีลักษณะดังกล่าวเป็นโมเดลที่

สามารถระบุได้ (จากนิยามดังกล่าวจะได้ว่า หากโมเดลการวิเคราะห์นั้นไม่สามารถระบุได้ อาจเป็นไปได้ว่าพารามิเตอร์บางตัวอาจมีคุณสมบัติสามารถระบุได้ (identified parameters) พารามิเตอร์ที่สามารถระบุได้ดังกล่าวนั้นจะเป็นค่าประมาณที่คงเส้นคงวาในทางสถิติ (consistent estimate) ในทางปฏิบัติหากโมเดลในการวิเคราะห์นั้นไม่สามารถระบุได้ผู้วิจัยสามารถแก้ไขปัญหาได้โดยการเลือกกำหนดค่าพารามิเตอร์บางตัวในโมเดลให้เป็นค่าคงที่ อย่างไรก็ตามผู้วิจัยต้องมีความระมัดระวังในการกำหนดเพราะการกำหนดดังกล่าวจะส่งผลต่อการแปลผลการวิเคราะห์ของโมเดล

ความระบุได้นั้นขึ้นอยู่กับการกำหนดโมเดลของผู้วิจัย สมมติให้มีจำนวนพารามิเตอร์ที่ต้องประมาณค่าจากโมเดลจำนวน t พารามิเตอร์ พิจารณาสมการ

$$\sigma_{ij} = f(\underline{\theta}) \qquad \text{เมื่อ } i \leq j \tag{2.120}$$

จากสมการดังกล่าวจะเห็นว่าในการประมาณค่าพารามิเตอร์ที่ไม่ทราบค่านั้นจำเป็นที่จะต้องใช้ข้อมูลจากค่าความแปรปรวนหรือความแปรปรวนร่วม (σ_{ij}) ซึ่ง $\sigma_{ij} \in \Sigma$ จากสมการจะเห็นว่าจะมีจำนวนสมการทั้งหมดเท่ากับ $\frac{1}{2}p(p+1)$ สมการ ซึ่งจะเห็นได้ว่า เงื่อนไขจำเป็น (necessary condition) ที่จะทำให้ผลเฉลยของสมการดังกล่าวมีคำตอบเดียวคือจำนวนพารามิเตอร์ที่ต้องการประมาณค่าต้องมีจำนวนไม่เกินจำนวนสมการกล่าวคือ $t \leq \frac{1}{2}p(p+1)$ ในกรณีที่จำนวนพารามิเตอร์ที่ต้องการประมาณนั้นมีจำนวนเท่ากับสมการ จะเรียกโมเดลดังกล่าวว่า โมเดลระบุได้พอดี (just-identified model) ในกรณีที่ผู้วิจัยประมาณพารามิเตอร์โดยใช้เซตย่อยของสมการในข้างต้น จะเรียกโมเดลดังกล่าวว่า โมเดลระบุเกิน (over identified model) ส่วนในกรณีที่จำนวนพารามิเตอร์ที่ต้องการประมาณมีมากกว่าจำนวนสมการในข้างต้นจะเรียกโมเดลดังกล่าวว่า โมเดลระบุไม่พอเพียง (underidentified model)

การตรวจสอบความระบุได้ของโมเดลนั้นสามารถกระทำได้หลายวิธีการดังนี้ 1) การพิจารณาจากเมทริกซ์สารสนเทศ (information matrix) กล่าวคือในกรณีที่ผู้วิจัยใช้วิธีกำลังสองน้อยสุดทั่วไป (generalized least squares) หรือวิธีภาวะความควรจะเป็นสูงสุด (maximum likelihood) ในการประมาณค่าพารามิเตอร์ ผู้วิจัยสามารถคำนวณเมทริกซ์สารสนเทศได้ โดยถ้าโมเดลที่ใช้ในการวิเคราะห์เป็นโมเดลที่สามารถระบุได้แล้วจะได้ว่าเมทริกซ์สารสนเทศเป็นเมทริกซ์

ที่เป็นบวกแน่นอน (positive definite) และถ้าเมทริกซ์สารสนเทศเป็นซิงกูลาร์เมทริกซ์ (singular matrix) จะได้ว่าโมเดลที่ใช้ในการวิเคราะห์เป็นโมเดลที่ไม่สามารถระบุได้ 2) การพิจารณาโดยการเลือกกำหนดค่าของพารามิเตอร์ $\underline{\theta}$ ที่สมเหตุสมผลตามโมเดลที่ต้องการวิเคราะห์มาจำนวนหนึ่ง จากนั้นใช้ค่าพารามิเตอร์ที่กำหนดขึ้นในการประมาณเมทริกซ์ความแปรปรวนร่วม Σ จากนั้นนำเมทริกซ์ความแปรปรวนร่วม Σ ดังกล่าวใช้เป็นข้อมูลในการประมาณค่าพารามิเตอร์ $\underline{\theta}$ ถ้าผลการประมาณค่าพารามิเตอร์ดังกล่าวได้ผลการประมาณที่ไม่แตกต่างจากค่าพารามิเตอร์ที่กำหนดไว้ในการสร้าง Σ ในตอนแรก แล้วส่วนใหญ่โมเดลดังกล่าวจะเป็นโมเดลที่ระบุได้

การประมาณค่าพารามิเตอร์ในโมเดล

1) *วิธีการประมาณค่าแบบดั้งเดิม*

โมเดลการวิเคราะห์องค์ประกอบทั้งสองนั้นจะเห็นว่า เมทริกซ์ความแปรปรวนร่วมของโมเดลทั้งสองล้วนเป็นฟังก์ชันที่ขึ้นกับค่าพารามิเตอร์ในโมเดล ดังนั้นจึงสามารถใช้สารสนเทศที่อยู่ภายในเมทริกซ์ความแปรปรวนร่วมหรือเมทริกซ์สหสัมพันธ์ในการประมาณค่าพารามิเตอร์ของโมเดล ซึ่งสามารถเขียนความสัมพันธ์ระหว่างเมทริกซ์ทั้งสองได้ดังนี้

$$\Sigma = D_\sigma P D_\sigma \tag{2.121}$$

โดยที่ เมทริกซ์ D_σ เป็นเมทริกซ์ทแยงมุมที่มีสมาชิกในแนวทแยงมุมเป็นส่วนเบี่ยงเบนมาตรฐานของตัวแปรสังเกตได้ที่ i และ P เป็นเมทริกซ์สัมพันธ์ (correlation matrix)

อย่างไรก็ตามผู้วิจัยไม่สามารถทราบค่าเมทริกซ์ความแปรปรวนร่วมหรือเมทริกซ์สหสัมพันธ์ของประชากรได้ แต่สามารถประมาณได้จากเมทริกซ์ความแปรปรวนร่วมของตัวแปรสังเกตได้ที่ได้จากตัวอย่างซึ่งสามารถคำนวณได้จาก

$$S = (n-1)^{-1}\sum_{i=1}^{n}(\underline{x}_i - \bar{x})(\underline{x}_i - \bar{x})^T \tag{2.122}$$

เมื่อ $\underline{x}_i = (x_{11}, x_{21}, ..., x_{p1})^T$ และ $\bar{x}$ คือเวกเตอร์ค่าเฉลี่ยของตัวแปรสังเกตได้

ตัวประมาณ S มีคุณสมบัติเป็นตัวประมาณที่ไม่เอนเอียง (unbiased estimator) ของ Σ และจะได้ว่า $(n-1)\cdot S$ มีการแจกแจงแบบ Wishart ที่มีพารามิเตอร์ p, Σ และ $n-1$ เขียนแทนได้ด้วย $(n-1)\cdot S \sim W_p(\Sigma, n-1)$

ในการศึกษาโดยทั่วไปผู้วิจัยมักไม่ค่อยสนใจที่จะพิจารณาค่าเฉลี่ยและหน่วยการวัดของตัวแปร เนื่องจากตัวแปรส่วนใหญ่ล้วนเป็นตัวแปรทางจิตวิทยาที่เป็นนามธรรม ดังนั้นในการวิเคราะห์ผู้วิจัยจึงมักใช้เมทริกซ์สหสัมพันธ์แทน เนื่องจากเหตุผลในข้างต้นดังนั้นปัญหาในการประมาณค่าพารามิเตอร์จึงจะพิจารณาหาค่าประมาณพารามิเตอร์เพื่อสร้างโครงสร้างของความแปรปรวนและความแปรปรวนร่วมใช้ในการอธิบาย เมทริกซ์ความแปรปรวนร่วมหรือเมทริกซ์สหสัมพันธ์ที่ได้จากตัวอย่าง วิธีการทางสถิติที่ใช้ในการประมาณค่าพารามิเตอร์ในโมเดลข้างต้นนั้นมีหลายวิธีการแต่อย่างไรก็ตามทุกวิธีการล้วนมีวัตถุประสงค์ที่คล้ายกันคือ หาค่าประมาณพารามิเตอร์ที่ทำให้ฟังก์ชันวัตถุประสงค์ (objective function) เขียนแทนด้วย $F = F(S,\Sigma)$ มีค่าต่ำสุด ฟังก์ชันวัตถุประสงค์ที่นิยมใช้กันในปัจจุบันเช่น ฟังก์ชันกำลังสองทั่วไปน้อยสุด (generalized least squares function) หรือฟังก์ชันภาวะความควรจะเป็น (maximum likelihood function) เป็นต้น ในกรณีที่ผู้วิจัยต้องการประมาณค่าพารามิเตอร์โดยใช้ฟังก์ชันกำลังสองน้อยสุดจะได้ว่าฟังก์ชันวัตถุประสงค์คือ

$$F_{GLS} = \frac{1}{2} tr\left(I - S^{-1}\Sigma\right)^2 \qquad (2.123)$$

ส่วนในกรณีที่ผู้วิจัยต้องการประมาณค่าพารามิเตอร์โดยใช้ฟังก์ชันภาวะความควรจะเป็นจะได้ว่าฟังก์ชันวัตถุประสงค์คือ

$$F_{ML} = tr\left(\Sigma^{-1}S\right) - \log\left|\Sigma^{-1}S\right| - p \qquad (2.124)$$

กำหนดให้ $\underline{\hat{\theta}}$ เป็นเวกเตอร์ของค่าประมาณพารามิเตอร์ จะได้ว่าเวกเตอร์ $\underline{\hat{\theta}}$ จะเป็นค่าประมาณพารามิเตอร์ของวิธีกำลังสองน้อยสุดทั่วไปหรือวิธีภาวะความควรจะเป็นสูงสุด อย่างไรก็ตามมีปัญหาน้อยมากที่ผู้วิจัยจะสามารถเขียนสูตรของการประมาณค่าพารามิเตอร์ของโมเดลดังกล่าวให้อยู่ในรูปปิดได้ (closed form) ดังนั้นวิธีการประมาณค่าพารามิเตอร์ในโมเดลดังกล่าวจึงจะเป็นวิธีการทวนซ้ำ (iterative procedure) เพื่อหาค่าที่ดีที่สุด ผู้ที่สนใจสามารถศึกษาได้จากเอกสารอ้างอิงที่ได้ระบุไว้ (Jöreskog, 1977,1978) ในกรณีที่การแจกแจงความน่าจะเป็นของตัวแปรสังเกตได้มีการแจกแจงแบบปกติหลายตัวแปรและขนาดตัวอย่างใหญ่เพียงพอ จะได้ว่าวิธีกำลังสองน้อยสุดทั่วไป และวิธีภาวะความควรจะเป็นสูงสุดนี้ จะเป็นตัวประมาณที่มีประสิทธิภาพ

(efficient estimator) โดยที่มีเงื่อนไขว่าเมทริกซ์ความแปรปรวนร่วมหรือเมทริกซ์สหสัมพันธ์จะต้องมีคุณสมบัติเป็นบวกแน่นอน (positive definite matrix) และมีคุณสมบัติคงเส้นคงวา (consistent)

กรณีที่ผู้วิจัยเลือกใช้วิธีภาวะความควรจะเป็นสูงสุดในการประมาณค่าพารามิเตอร์ ผู้วิจัยสามารถทดสอบความเหมาะสมของโมเดล (assessment of fit) ที่ผู้วิจัยต้องการศึกษาได้ (proposed model) ซึ่งใช้วิธีการสร้างตัวทดสอบโดยใช้วิธีอัตราส่วนของภาวะความควรจะเป็น (likelihood ratio test) ซึ่งเป็นอัตราส่วนของฟังก์ชันภาวะความควรจะเป็นระหว่างฟังก์ชันภาวะความควรจะเป็นเมื่อสมมติว่าโมเดลของผู้วิจัยเป็นจริง และฟังก์ชันภาวะความควรจะเป็นเมื่อสมมติว่าโมเดลของผู้วิจัยไม่เป็นจริงดังนี้

$$LR = \frac{f\left(S \mid \Sigma(\hat{\theta})\right)}{f\left(S \mid (n-1)n^{-1}S\right)} \tag{2.125}$$

เมื่อ $\Sigma(\hat{\theta})$ เป็นเมทริกซ์ความแปรปรวนร่วมที่สร้างจาก proposed model และ $(n-1)n^{-1}S$ เป็นตัวประมาณภาวะความควรจะเป็นสูงสุดของเมทริกซ์ความแปรปรวนร่วมที่ได้จากตัวอย่าง เนื่องจากถ้าเวกเตอร์ของตัวแปรสังเกตได้มีการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรที่มีค่าเฉลี่ย μ_0 และ เมทริกซ์ความแปรปรวนร่วมเป็น Σ_0 จะได้ว่าการแจกแจงความน่าจะเป็นของเมทริกซ์ความแปรปรวนร่วมของตัวอย่าง S มีการแจกแจงแบบ Wishart ที่มีฟังก์ชันความหนาแน่นเป็น

$$f(S \mid \Sigma_0) = \frac{C \exp\left\{-(n-1) \cdot tr(S\Sigma_0^{-1})/2\right\}}{\left|\Sigma_0\right|^{(n-1)/2}} \tag{2.126}$$

จากสมการ (2.125) และ (2.126) จะได้ว่า $-2\log(LR)$ คือ

$$-2\log(LR) = -2\log\left[\frac{\left|\Sigma(\hat{\theta})\right|^{-(n-1)/2} \cdot \exp\left\{-(n-1)tr(S\Sigma^{-1}(\hat{\theta})/2\right\}}{\left|(n-1)n^{-1}S\right|^{-(n-1)/2} \cdot \exp\left\{-(n-1)tr(S\left\{(n-1)n^{-1}S\right\}^{-1}/2\right\}}\right] \xrightarrow{p} nF_{ML}(\hat{\theta}) \tag{2.127}$$

จากทฤษฎีภาวะความควรจะเป็นสูงสุดจะได้ว่า $nF_{ML}(\hat{\theta}) \xrightarrow{L} \chi^2_{p^*-q}$ ดังนั้นจะปฏิเสธ $\Sigma(\theta)$ เมื่อ $nF_{ML}(\hat{\theta})$ มีค่ามากกว่า $\chi^2_{p^*-q}$ ที่ระดับนัยสำคัญ α และจะไม่สามารถปฏิเสธ $\Sigma(\theta)$ เมื่อ

$nF_{ML}(\hat{\theta})$ มีค่าน้อยกว่า $\chi^2_{p^*-q}$ อย่างไรก็ตามเมื่อไม่สามารถปฏิเสธ $\Sigma(\theta)$ นั้นไม่ได้หมายความว่า $\Sigma(\theta)$ เป็นโมเดลที่ถูกต้องในทางทฤษฎี

2) *วิธีการประมาณค่าพารามิเตอร์แบบเบส์*

การประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบโดยใช้วิธีการแบบเบส์นั้น มีหลักการที่แตกต่างจากวิธีแบบดั้งเดิม กล่าวคือจะใช้ข้อมูลของตัวแปรสังเกตได้เป็นข้อมูลสำหรับการหาค่าประมาณพารามิเตอร์ในโมเดล เพื่อความสะดวกในการนำเสนอรายละเอียดของวิธีการประมาณค่าพารามิเตอร์ผู้วิจัยจะนำเสนอรายละเอียดเบื้องต้นของโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอีกครั้งดังนี้ สำหรับหน่วยตัวอย่างที่ $i = 1,2,\dots,n$

$$\underline{y}_i = \Lambda\underline{\omega}_i + \underline{\epsilon}_i \tag{2.128}$$

โดยที่ $\underline{y}_i$ คือเวกเตอร์ของตัวแปรสังเกตได้ขนาด $p \times 1$, Λ คือเมทริกซ์ของน้ำหนักองค์ประกอบขนาด $p \times q$, $\underline{\omega}_i$ คือเวกเตอร์ของตัวแปรแฝงขนาด $q \times 1$ และ $\underline{\epsilon}_i$ คือเวกเตอร์ของความคลาดเคลื่อนจากการวัดขนาด $p \times 1$ และข้อสมมติสำหรับโมเดลในสมการที่ (2.128) เป็นดังนี้ 1) $\underline{\epsilon}_i \sim N(\underline{0}, \Psi_\epsilon)$ เมื่อ Ψ_ϵ คือเมทริกซ์ทแยงมุมของความแปรปรวนของความคลาดเคลื่อนจากการวัดขนาด $p \times p$ และ 2) $\underline{\omega}_i \sim N(\underline{0}, \Phi)$ เมื่อ Φ คือเมทริกซ์ของความแปรปรวนและความแปรปรวนร่วมของตัวแปรแฝงขนาด $q \times q$

กำหนดให้ $Y = \left(\underline{y}_1, \underline{y}_2, \dots, \underline{y}_n\right)$ เป็นเมทริกซ์ของตัวแปรสังเกตได้ของทุกหน่วยตัวอย่าง และ $\Omega = (\underline{\omega}_1, \underline{\omega}_2, \dots, \underline{\omega}_n)$ เป็นเมทริกซ์ของตัวแปรแฝงของทุกหน่วยตัวอย่าง และ $\underline{\theta} = (\Lambda, \Psi_\epsilon, \Phi)$ เป็นเวกเตอร์ของพารามิเตอร์ในโมเดล (2.128) ในการประมาณค่าพารามิเตอร์ตามโมเดล (2.128) ด้วยวิธีการแบบเบส์ด้วยอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ (gibbs-sampling algorithm) ประกอบไปด้วยขั้นตอนย่อยจำนวน 4 ขั้นตอน (Dunson, Palomo, & Bollen, 2005) ดังนี้ สำหรับรอบที่ $m+1$ ของการทวนซ้ำ

<u>ขั้นที่ 1:</u> สุ่มตัวอย่างตัวแปรแฝง $\Omega^{(m+1)}$ จาก $p\left(\Omega \middle| \Psi_\epsilon^{(\mathrm{m})}, \Lambda^{(\mathrm{m})}, \Phi^{(m)}, Y\right)$

โดยที่ $p\left(\Omega \middle| \Psi_\epsilon\ , \Lambda\ , \Phi\ , Y\right) = \prod_{i=1}^{n} p\left(\underline{\omega}_i | \Psi_\epsilon\ , \Lambda\ , \Phi\ , Y\right)$

และ

$$p\left(\underline{\omega}_i | \Psi_\epsilon\ , \Lambda\ , \Phi\ , Y\right) = N\left((\Phi^{-1} + \Lambda^{\mathrm{T}}\Psi_\epsilon^{-1}\Lambda)^{-1}\Lambda^{\mathrm{T}}\Psi_\epsilon^{-1}\underline{y}_i, (\Phi^{-1} + \Lambda^{\mathrm{T}}\Psi_\epsilon^{-1}\Lambda)^{-1}\right)$$

<u>ขั้นที่ 2:</u> สุ่มตัวอย่างพารามิเตอร์ $\Psi_\epsilon^{(\mathrm{m+1})}$ จาก $p\left(\Psi_\epsilon \left|\Omega^{(m+1)}, \Lambda^{(\mathrm{m})}, \Phi^{(m)}, Y\right.\right)$

โดยที่ $p\left(\Psi_\epsilon \left|\Omega \quad, \Lambda \quad, \Phi \quad, Y\right.\right) = Gamma\left(\frac{n}{2} + \alpha_{0\epsilon k}, \beta_{\epsilon k}\right)$

$$\beta_{\epsilon k} = \beta_{0\epsilon k} + 0.5\left(Y_k^T Y_k^T - a_k^Y A_k^{-1} a_k + \Lambda_{ok}^{\mathrm{T}} H_{0yk}^{-1} \Lambda_{0k}\right)$$
$$a_k = A_k\left(H_{ko}^{-1} \Lambda_{0k} + \Omega \mathrm{Y_k}\right)$$
$$A_k = \left(H_{ko}^{-1} + \Omega\Omega^T\right)^{-1}$$

<u>ขั้นที่ 3:</u> สุ่มตัวอย่างพารามิเตอร์ $\Lambda^{(m+1)}$ จาก $p\left(\Lambda \left|\Omega^{(m+1)}, \Psi_\epsilon^{(\mathrm{m+1})}, \Phi^{(m)}, Y\right.\right)$

โดยที่ $p\left(\Lambda \left|\Omega \quad, \Psi_\epsilon \quad, \Phi \quad, Y\right.\right) = \prod_{k=1}^{p} p\left(\Lambda_k \middle| \Omega \quad, \Psi_\epsilon \quad, \Phi \quad, Y\right)$

เมื่อ Λ_k เป็นเวกเตอร์ของสมาชิกในเมทริกซ์น้ำหนักองค์ประกอบในแถวที่ k

และ $p\left(\Lambda_k \middle| \Omega \quad, \Psi_\epsilon \quad, \Phi \quad, Y\right) \sim N(a_k, \psi_{\epsilon k} A_k)$

<u>ขั้นที่ 4:</u> สุ่มตัวอย่างพารามิเตอร์ $\Phi^{(m+1)}$ จาก $p\left(\Phi \left|\Omega^{(m+1)}, \Lambda^{(\mathrm{m+1})}, \Psi_\epsilon^{(\mathrm{m})+1}, Y\right.\right)$

โดยที่ $p\left(\Phi \left|\Omega \quad, \Lambda \quad, \Psi_\epsilon \quad, Y\right.\right) \sim IW_q(\Omega\Omega^{\mathrm{T}} + R_0^{-1}, n + \rho_0)$

4.4 <u>การแก้ไขปัญหาความคลาดเคลื่อนจากการวัดในโมเดลเชิงเส้นพหุระดับ</u>

จากที่กล่าวไว้ในหัวข้อที่แล้ว ปัญหาความคลาดเคลื่อนจากการวัดในโมเดลแบบระดับเดียวนั้นได้มีการศึกษาทั้งผลกระทบและวิธีการแก้ปัญหา ในทางทฤษฎีปัญหาความคลาดเคลื่อนจากการวัดในโมเดลพหุระดับจะทำให้ตัวประมาณมีความเอนเอียง ไม่มีประสิทธิภาพส่งผลต่อประสิทธิภาพในการประมาณ อย่างไรก็ตามในปัจจุบันยังไม่ได้มีผู้ทำการศึกษาผลกระทบของปัญหาดังกล่าวในโมเดลพหุระดับอย่างชัดเจน เนื่องจากโมเดลพหุระดับนั้นเป็นโมเดลทางสถิติที่มีความซับซ้อน ดังนั้นการแก้ปัญหาความคลาดเคลื่อนจากการวัดในโมเดลดังกล่าวจึงกระทำได้ยาก และวิธีการที่พัฒนาขึ้นนั้นไม่สามารถใช้ได้ในกรณีทั่วไป ดังจะเห็นได้จาก วิธีการของ Woodhouse ที่ได้ทำการพัฒนาวิธีการแก้ปัญหาโดยใช้ตัวประมาณโมเมนต์ (moment type estimator) วิธีการดังกล่าวสามารถใช้แก้ไขปัญหาความคลาดเคลื่อนจากการวัดในโมเดลเชิงเส้นพหุระดับได้จริง แต่ไม่สามารถใช้ได้ในกรณีที่สัมประสิทธิ์ความถดถอยในโมเดลเป็นแบบสุ่ม นอกจากนี้ในปี พ.ศ. 2551 สิวะโชติ ศรีสุทธิยากร และศิริชัย กาญจนวาสี ได้พัฒนาวิธีการประมาณค่าพารามิเตอร์ในโมเดลการถดถอยเชิงเส้นพหุระดับในกรณีที่มีความคลาดเคลื่อนในการวัดตัวแปรอิสระในระดับที่

1 เรียกว่าวิธี adjusted IGLS ซึ่งเป็นการผสมผสานเทคนิควิธีการประมาณ 2 วิธีการเข้าด้วยกัน ได้แก่วิธี IGLS และวิธี JLS ซึ่งในการประมาณนั้นได้ใช้วิธี IGLS เป็นพื้นฐานในการประมาณ และใช้เทคนิคของวิธี JLS ปรับสูตรการประมาณในวิธี IGLS เพื่อแยกแยะความผันแปรของความคลาดเคลื่อนในการวัดของตัวแปรอิสระในระดับที่ 1 ออกมา และทำการหาค่าประมาณพารามิเตอร์ที่ทำให้ ผลบวกของผลรวมกำลังสองของโมเดล และผลบวกของผลรวมกำลังสองของความคลาดเคลื่อนในการวัดมีค่าต่ำที่สุด ผลการวิจัยทำให้สามารถพัฒนาวิธีการประมาณค่าพารามิเตอร์ในโมเดลการถดถอยเชิงเส้นพหุระดับ ที่เรียกว่าโมเดลสัมประสิทธิ์จุดตัดแกนสุ่ม (random intercept model) ได้สำเร็จ และจากการเปรียบเทียบผลที่ได้จากวิธีการประมาณที่พัฒนาขึ้นใหม่พบว่าวิธี adjusted IGLS จะให้ค่าประมาณพารามิเตอร์ที่ถูกต้องแม่นยำกว่าวิธี IGLS แบบดั้งเดิม โดยเฉพาะเมื่อความเที่ยงของตัวแปรอิสระมีค่าต่ำลง อย่างไรก็ตามการพัฒนาวิธีการประมาณตามแนวคิดดังกล่าวนั้นมีข้อเสียคือ เนื่องจากโมเดลเชิงเส้นพหุระดับนั้นเป็นโมเดลที่มีความซับซ้อน การคำนวณค่าพารามิเตอร์ตามแนวคิดดังกล่าวนั้นไม่สามารถพิสูจน์จนได้สูตรการประมาณค่าให้อยู่ในรูปปิดได้ (closed form) ดังนั้นจึงจำเป็นที่จะต้องพึ่งวิธีเชิงตัวเลข (numerical methods) มาช่วยในการประมาณค่าพารามิเตอร์ อีกทั้งวิธีการประมาณยังเกี่ยวข้องกับการสร้างเมทริกซ์ขนาดใหญ่ และการประมาณค่าแบบทวนซ้ำ ทำให้อัลกอริทึมของวิธีการประมาณดังกล่าวเป็นอัลกอริทึมที่มีขนาดใหญ่และใช้ทรัพยากรมากในการปฏิบัติ นอกจากนี้การพัฒนาวิธีการดังกล่าวเพื่อให้สามารถใช้ได้ในกรณีทั่วไปตามแนวทางเดิมที่ใช้และกระทำให้มีประสิทธิภาพสามารถทำได้ยาก (สิวะโชติ ศรีสุทธิยากร และ ศิริชัย กาญจนวาสี, 2551; Woodhouse G., 1996)

<u>*โมเดลสมการโครงสร้างพหุระดับ (multi-level SEM)*</u>

เป็นโมเดลทางสถิติที่ออกแบบมาเพื่อใช้ในการวิเคราะห์ข้อมูลแบบพหุระดับโดยใช้แนวทางของการวิเคราะห์แบบโมเดลสมการโครงสร้าง (structural equation model:SEM) กล่าวคือเป็นการขยายขอบเขตของโมเดลสมการโครงสร้างเพื่อให้สามารถวิเคราะห์กับข้อมูลแบบพหุระดับได้นั่นเอง Muthen และ Asporouhov (2010) ได้เสนอโมเดลสำหรับการวิเคราะห์ข้อมูล

แบบสองระดับ (two-levels model) ซึ่งโมเดลหลักทั่วไปแล้วประกอบไปด้วย โมเดลการวัด (measurement model) และโมเดลสมการโครงสร้าง (structural model) ดังสมการต่อไปนี้

<u>โมเดลระดับที่ 1</u>

โมเดลการวัด (measurement model)

$$Y_{ij} = \nu_j + \Lambda_j \eta_{ij} + K_j X_{ij} + \varepsilon_{ij} \quad (2.129)$$

โมเดลสมการโครงสร้าง (structural model)

$$\eta_{ij} = \alpha_j + B_j \eta_{ij} + \Gamma_j X_{ij} + \zeta_{ij} \quad (2.130)$$

<u>โมเดลระดับที่ 2</u>

โมเดลสมการโครงสร้าง (structural model)

$$\eta_j = \mu + \beta \eta_j + \gamma X_j + \zeta_j \quad (2.131)$$

จากโมเดลในข้างต้นจะเห็นว่าโมเดลสมการโครงสร้างพหุระดับมีความสามารถในการแก้ไขปัญหาความคลาดเคลื่อนจากการวัดที่เกิดขึ้นในข้อมูลพหุระดับได้อย่างเหมาะสม เนื่องจากได้มีการคำนึงถึงความผันแปรที่เกิดขึ้นจากความแตกต่างระหว่างหน่วย นอกจากนี้เมื่อพิจารณากรอบแนวคิดทั่วไปของโมเดลการวิเคราะห์ในโปรแกรม Mplus จะเห็นว่าโปรแกรม Mplus มีความสามารถในการวิเคราะห์โมเดลตัวแปรแฝงทั้งในกรณีที่ตัวแปรสังเกตเป็นตัวแปรจัดประเภท (categorical variables) และเป็นตัวแปรต่อเนื่อง (continuous variable) โดยได้มีการวิเคราะห์ในกรณีที่ตัวแปรสังเกตได้เป็นตัวแปรต่อเนื่องสามารถใช้โมเดลในสมการที่ (2.129) เป็นโมเดลการวัด แต่ในกรณีที่ตัวแปรสังเกตได้เป็นตัวแปรจัดประเภท โปรแกรม Mplus จะใช้แนวคิดของการระบุค่า threshold (threshold specifications) เพื่อสร้างตัวแปรบ่งชี้แบบต่อเนื่องจากตัวแปรสังเกตได้แบบจัดประเภท จากนั้นจึงใช้โมเดลในสมการที่ (2.129) วิเคราะห์ต่อไป (Muthen, 2004) (Muthen, 2004)

การประมาณค่าพารามิเตอร์ที่ใช้ในโปรแกรม Mplus พบว่าโปรแกรม Mplus มีตัวประมาณหลายตัวประมาณ เช่น

1. ตัวประมาณภาวะความควรจะเป็นสูงสุด (maximum likelihood estimator)

หลักการคือหาค่าประมาณค่าพารามิเตอร์ในโมเดลที่ทำให้ฟังก์ชันภาวะความควรจะเป็นมีค่าสูงสุด (maximized likelihood function) โดยที่ข้อสมมติที่สำคัญคือการแจกแจงความน่าจะเป็นของตัวแปรสังเกตได้ต้องมีการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปร (multivariate normal distribution) ฟังก์ชันภาวะความควรจะเป็นในโปรแกรม Mplus เขียนได้ดังสมการต่อไปนี้

$$F_{ML}(\pi) = \frac{1}{2}\sum_{g=1}^{G}\{n_g[\ln|\Sigma_g| + trace(\Sigma_g^{-1}T_g) - \ln|S_g| - (p+g)]\}/n \quad (2.132)$$

เมื่อ ν_{gi} คือเวกเตอร์ของตัวแปรสังเกตกลุ่มที่ g หน่วยตัวอย่างที่ i, $\mu_g = E(\nu_{gi})$ คือเวกเตอร์ของค่าเฉลี่ยของตัวแปรสังเกตได้, Σ_g คือเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวแปรสังเกตได้ของประชากร, S_g คือเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวแปรสังเกตได้ของตัวอย่างและ $T_g = S_g + (\overline{\upsilon}_g - \mu_g)(\overline{\upsilon}_g - \mu_g)^T$

2. ตัวประมาณกำลังสองน้อยสุดแบบถ่วงน้ำหนัก (weigted least-squares estimation)

วิธีการนี้ไม่จำเป็นที่ข้อสมมติเกี่ยวกับการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรของตัวแปรสังเกตได้ต้องเป็นจริง เนื่องจากตัวประมาณไม่ได้ขึ้นกับการแจกแจงความน่าจะเป็นของข้อมูลตัวแปรสังเกตได้ หลักการคือการหาค่าประมาณค่าพารามิเตอร์ที่ทำให้ฟังก์ชันกำลังสองแบบถ่วงน้ำหนักมีค่าน้อยที่สุด (minimized weighted squares function) ฟังก์ชันกำลังสองแบบถ่วงน้ำหนักเขียนได้ดังสมการต่อไปนี้

$$F_{WLS}(\pi) = \frac{1}{2}\sum_{g=1}^{G}(s_g - \sigma_g)^T W_g (s_g - \sigma_g) \quad (2.133)$$

ในกรณีที่ตัวแปรสังเกตได้เป็นตัวแปรแบบต่อเนื่องแต่ไม่ได้มีการแจกแจงแบบปกติหลายตัวแปร การประมาณค่าด้วยวิธีการนี้ให้ค่าประมาณแบบที่ไม่ขึ้นกับการแจกแจงความน่าจะเป็นของข้อมูล (distribution free method: ADF) ส่วนในกรณีที่มีตัวแปรสังเกตได้แบบจัดประเภทผู้ที่สนใจสามารถอ่านศึกษารายละเอียดเพิ่มเติมได้จากงานวิจัยของ Muthen (1983, 1984) นอกจากนี้การประมาณค่าโดยใช้ตัวประมาณกำลังสองน้อยสุดแบบถ่วงน้ำหนักนี้ยังประกอบไปด้วยขั้นตอนการประมาณค่าพารามิเตอร์จำนวน 3 ขั้นตอน 1) การประมาณค่าเฉลี่ยและความแปรปรวนความแปรปรวนร่วมของตัวแปรสังเกตได้ 2) การประมาณเมทริกซ์ถ่วงน้ำหนัก และ 3) การประมาณ

ค่าพารามิเตอร์โดยหาค่าประมาณค่าพารามิเตอร์ที่ทำให้ฟังก์ชันกำลังสองแบบถ่วงน้ำหนักมีค่าน้อยที่สุด

วิธีการประมาณค่าด้วย adjusted MCMC (adjusted MCMC algorithm)

การแก้ไขปัญหาด้วยสถิติแบบเบส์นั้นอาจสามารถแยกออกได้เป็น 2 แนวทาง แนวทางแรกคือ การพัฒนาวิธีการแก้ปัญหาโดยการใช้โมเดลการวัดแบบดั้งเดิม พัฒนาขึ้นโดย William J. Browne และคณะ ซึ่งอาศัยหลักการของวิธีประมาณ MCMC เดิมแต่มีการขยายอัลกอริทึมในการประมาณเพื่อรวมความคลาดเคลื่อนจากการวัดไว้ในกระบวนการวิเคราะห์ดังกล่าวเรียกชื่ออัลกอริทึมดังกล่าวว่า “adjusted MCMC” แนวทางดังกล่าวมีข้อสมมติเกี่ยวกับโมเดลที่สำคัญคือ ความคลาดเคลื่อนจากการวัดนั้นจะมีธรรมชาติของความคลาดเคลื่อนตามโมเดลการวัดแบบดั้งเดิม เป็นอิสระซึ่งกันและกัน และมีข้อจำกัดคือใช้ได้กับโมเดลที่มีตัวแปรอิสระเกิดความคลาดเคลื่อนจากการวัดเพียง 1 ตัวเท่านั้น ต่อมา Harvey Goldstein และคณะได้ทำการพัฒนาขยายอัลกอริทึมของ Browne โดยการยอมให้ความคลาดเคลื่อนจากการวัดนั้นมีความสัมพันธ์กันได้ นอกจากนี้ยังขยายอัลกอริทึมให้มีการรวมความคลาดเคลื่อนจากการวัดในตัวแปรตามไว้ในกระบวนการวิเคราะห์อีกด้วย โดยสารสนเทศที่สำคัญจำเป็นจำต้องทราบในการประมาณค่าพารามิเตอร์ตามแนวทางดังกล่าวคือค่าความเที่ยงซึ่งจะนำมาเป็นข้อมูลที่ใช้ในการปรับค่าความคลาดเคลื่อนจากการวัดตัวแปรในโมเดล แนวทางที่สองคือ การพัฒนาวิธีการแก้ปัญหาโดยการใช้โมเดลการวัดแบบตอบสนองข้อสอบ พัฒนาขึ้นในปี ค.ศ. 2001 โดย Jean-Paul Fox และคณะ อย่างไรก็ตามการพัฒนาวิธีการแก้ปัญหาตามแนวทางนี้อยู่นอกเหนือขอบเขตการวิจัย ดังนั้นจึงจะไม่ขอกล่าวรายละเอียดของวิธีการดังกล่าวในที่นี้ ผู้ที่สนใจสามารถศึกษาได้จากเอกสารที่เกี่ยวข้องตามที่ได้ระบุไว้ (Fox. J. –P., 2005; Fox. J. –P., and Glas, 2003)

วิธี adjusted MCMC ที่พัฒนาขึ้นสำหรับโมเดลเชิงเส้นพหุระดับแบบ 2 ระดับ โดย William J. Browne และคณะนั้นมีข้อกำหนดเบื้องต้นในโมเดลเพิ่มเติมข้อกำหนดเบื้องต้นของโมเดลเชิงเส้นพหุระดับทั่วไปคือ 1) ความคลาดเคลื่อนจากการวัดเป็นอิสระซึ่งกันและกันกับตัวแปรอิสระ 2) ต้องทราบความแปรปรวนของความคลาดเคลื่อนจากการวัด และ 3) ตัวแปรอิสระที่

แท้จริงนั้นสมมติให้มีการแจกแจงแบบปกติ รายละเอียดของโมเดลและวิธีการในการประมาณค่าพารามิเตอร์เป็นดังต่อไปนี้

<u>โมเดลระดับที่ 1</u>

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij}^{o} + \varepsilon_{ij} \tag{2.134}$$

$$x_{ij}^{o} = x_{ij} + e_{ij} \tag{2.135}$$

<u>โมเดลระดับที่ 2</u>

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{2.136}$$

$$\beta_{1j} = \gamma_{01} + u_{1j} \tag{2.137}$$

<u>โมเดลรวม</u>

$$y_{ij} = \gamma_{00} + \gamma_{01} x_{ij}^{o} + u_{0j} + x_{ij}^{o} u_{1j} + \varepsilon_{ij} \tag{2.138}$$

เมื่อ $\varepsilon_{ij} \sim Normal(0, \sigma_{\varepsilon}^{2})$, $\underline{u}_{j} = \left(u_{oj}, u_{1j}\right)^{T} \sim MVN(0, \Omega_{u})$, $x_{1ij}^{0} \sim Normal(x_{1ij}, \sigma_{m}^{2})$ และ $x_{1ij} \sim Normal(\theta, \phi^{2})$ โดยที่ $i = 1,2,...,n_{j}$ และ $j = 1,2,...,J$

จากการกำหนดโมเดลจะเห็นว่าโมเดลดังกล่าวมีรายละเอียดที่เพิ่มเติมจากโมเดลเชิงเส้นพหุระดับแบบปกติเนื่องจากมีการนำความคลาดเคลื่อนในการวัดเข้ามาในโมเดล พารามิเตอร์ในโมเดลประกอบไปด้วยพารามิเตอร์ที่เป็นอิทธิพลคงที่ $\underline{\gamma} = (\gamma_{00}, \gamma_{01})^{T}$ เศษเหลือในระดับที่ 2 $\underline{u}_{j} = \left(u_{oj}, u_{1j}\right)^{T}$ เศษเหลือในระดับที่ 1 ตัวแปรอิสระที่สังเกตค่าได้ x_{1ij}^{o} เป็นตัวแปรที่สังเกตค่าโดยมีความคลาดเคลื่อนในการวัดและสมมติให้ความคลาดเคลื่อนในการวัดเป็นอิสระกันด้วยและมีความแปรปรวนของความคลาดเคลื่อนในการวัดเท่ากับ σ_{m}^{2} ตัวแปรอิสระที่แท้จริงที่ไม่มีความคลาดเคลื่อนในการวัดคือ x_{1ij} เพื่อความสะดวกจึงกำหนดสัญลักษณ์ให้ $x_{0ij} = x_{0ij}^{o} = 1, \forall i, j$ และกำหนดให้ $X_{ij} = \left(x_{0ij}, x_{1ij}\right)^{T}$ แทนคะแนนของตัวแปรที่แท้จริงทั้งหมด

การกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดลนั้นกำหนดไว้ดังต่อไปนี้ การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนในระดับที่ 1 กำหนดให้มีการแจกแจงแบบสเกลอินเวอร์สไคสแควร์ (scale inverse χ^{2} prior) ที่มีพารามิเตอร์ ν_{ε} และ s_{ε}^{2} การแจกแจงความน่าจะเป็นของส่วนประกอบความแปรปรวนในระดับที่ 2 ซึ่งอยู่ในรูป

ของเมทริกซ์ความแปรปรวนร่วมนั้น กำหนดให้มีการแจกแจงความน่าจะเป็นแบบอินเวอร์ส Wishart (inverse Wishart prior) ที่มีพารามิเตอร์ ν_2 และ S_2 การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ที่เป็นอิทธิพลคงที่ กำหนดให้มีการแจกแจงแบบปกติ (normal prior) ที่มีค่าเฉลี่ยเท่ากับ μ_p และความแปรปรวนเท่ากับ S_p การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ θ กำหนดให้มีการแจกแจงแบบสม่ำเสมอ (uniform prior) และการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ ϕ^2 กำหนดให้มีการแจกแจงแบบสเกลอินเวอร์สไคสแควร์ที่มีพารามิเตอร์ ν_ϕ และ s_ϕ^2 (Browne, Goldstein, Woodhouse, and Yang, 2001)

อัลกอริทึมที่ใช้ในการประมาณค่าพารามิเตอร์นั้นจะใช้อัลกอริทึม Gibbs sampling ซึ่งประกอบไปด้วย 7 ขั้นตอนในการสุ่มตัวอย่างพารามิเตอร์ ดังนี้

ขั้นที่ 1: สุ่มตัวอย่างพารามิเตอร์ที่เป็นอิทธิพลคงที่จาก $\gamma_{00} \mid \underline{y}, X, \underline{u}, \Omega_u, \sigma_\varepsilon^2 \sim MVN(\underline{\hat{\gamma}}, \hat{D})$

เมื่อ $$\hat{D} = \left[\frac{\sum_{i.,j} (X_{ij})^T (y_{ij} - X_{ij}\underline{u}_j)}{\sigma_\varepsilon^2} + S_p^{-1} \right]^{-1}$$

และ $$\underline{\hat{\gamma}} = \hat{D} \cdot \left[\frac{\sum_{i.,j} (X_{ij})^T (y_{ij} - X_{ij}\underline{u}_j)}{\sigma_\varepsilon^2} + S_p^{-1}\mu_p \right]^{-1}$$

ขั้นที่ 2: สุ่มตัวอย่างเศษเหลือในระดับที่ 2 จาก $u_j \mid \underline{y}, X, \underline{\gamma}, \Omega_u, \sigma_\varepsilon^2 \sim MVN(\hat{u}_j, \hat{D}_j)$

เมื่อ $$\hat{D}_j = \left[\frac{\sum_{i=1}^{n_j} (X_{ij})^T X_{ij}}{\sigma_\varepsilon^2} + \Omega_u^{-1} \right]^{-1}$$ และ $$\hat{u}_j = \frac{\hat{D}_j}{\sigma_\varepsilon^2} \cdot \sum_{i=1}^{n_j} (X_{ij})^T (y_{ij} - X_{ij}\underline{\gamma})$$

ขั้นที่ 3: สุ่มตัวอย่างพารามิเตอร์ความแปรปรวนในระดับที่ 1

จาก $1/\sigma_\varepsilon^2 \mid \underline{y}, X, \underline{\gamma}, \underline{u}, \Omega_u \sim Gamma(a_\varepsilon, b_\varepsilon)$

เมื่อ $$a_\varepsilon = \frac{N + \nu_\varepsilon}{2}$$ และ $$b_\varepsilon = \frac{1}{2}\left(\nu_\varepsilon s_\varepsilon^2 + \sum_{i.j} e_{ij}^2 \right)$$

ขั้นที่ 4: สุ่มตัวอย่างพารามิเตอร์เมทริกซ์ความแปรปรวนร่วมในระดับที่ 2 จาก

$$\Omega_u \mid \underline{y}, X, \underline{\gamma}, \underline{u}, \sigma_\varepsilon^2 \sim Wishart_{n2}(S_{pos} = \left(\sum_{j=1}^{J} u_j (u_j)^T + S_P\right)^{-1}, \nu_{pos} = J + \nu_P)$$

เมื่อ n_2 คือจำนวนของตัวแปรสุ่มในระดับที่ 2 ของโมเดล

<u>ขั้นที่ 5</u>: สุ่มตัวอย่างตัวแปรอิสระที่แท้จริงในระดับที่ 1

จาก $x_{1ij} \mid \underline{y}, x_{1ij}^o, \underline{\gamma}, \underline{u}, \sigma_\varepsilon^2, \Omega_u \sigma_m^2, \theta, \phi^2 \sim N(\hat{x}_{1ij}, \hat{V}_{1ij})$

เมื่อ $\hat{V}_{1ij} = \left[\frac{(\gamma_{01} + u_{1j})^2}{\sigma_\varepsilon^2} + \frac{1}{\sigma_m^2} + \frac{1}{\phi^2}\right]^{-1}$

และ $\hat{x}_{1ij} = \hat{V}_{1ij} \cdot \left[\frac{(\gamma_{01} + u_{1j})(y_{ij} - \gamma_{00} - u_{0j})}{\sigma_\varepsilon^2} + \frac{x_{1ij}^o}{\sigma_m^2} + \frac{\theta}{\phi^2}\right]^{-1}$

<u>ขั้นที่ 6</u>: สุ่มตัวอย่างพารามิเตอร์ค่าเฉลี่ยของตัวแปรอิสระที่แท้จริงในระดับที่ 1

จาก $\theta \mid x_{1ij}, \phi^2 \sim N(\hat{\theta}, \hat{V}_\theta)$

เมื่อ $\hat{V}_\theta = \frac{\phi^2}{N}$ และ $\hat{\theta} = \frac{\sum_{i,j} x_{1ij}}{N}$

<u>ขั้นที่ 7</u>: สุ่มตัวอย่างพารามิเตอร์ความแปรปรวนของตัวแปรอิสระที่แท้จริงในระดับที่ 1

จาก $1/\phi^2 \mid x_{1ij}, \theta \sim Gamma(a_\phi, b_\phi)$

เมื่อ $a_\phi = \frac{N + \nu_\phi}{2}$ และ $b_\phi = \frac{1}{2}\left(\nu_\phi s_\phi^2 + \sum_{i.j}(x_{1ij} - \theta)^2\right)$

จากอัลกอริทึมในข้างต้นจะเห็นว่าประกอบไปด้วย 7 ขั้นตอนโดยที่ 4 ขั้นตอนแรกนั้นเป็นส่วนที่ใช้ประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับเช่นเดียวกับอัลกอริทึม MCMC แบบดั้งเดิม หากพิจารณาจากเห็นว่าทั้ง 4 ขั้นตอนแรกนี้จะสมมติว่าได้ทราบค่าของตัวแปรอิสระที่แท้จริงแล้ว อีก 3 ขั้นตอนสุดท้ายเป็นการปรับค่าของตัวแปรอิสระให้ใกล้เคียงค่าของตัวแปรอิสระที่แท้จริงโดยใช้ความแปรปรวนของความคลาดเคลื่อนจากการวัด (σ_m^2) เป็นข้อมูลในการปรับค่าตัวแปรอิสระดังกล่าว วิธีการที่จะทราบค่า σ_m^2 ได้นั้นมีหลายวิธี วิธีการหนึ่งคือการประมาณ σ_m^2 โดยการใช้ค่าความเที่ยงกล่าวคือหากทราบค่าความเที่ยงของตัวแปรแล้วเราสามารถประมาณ σ_m^2 ได้จากสมการที่ (75) ในโครงร่างการวิจัยฉบับนี้ เมื่อกระทำตามอัลกอริทึมขั้นตอนที่ 5 ถึง 7 เรียบร้อยแล้วจะนำค่าของตัวแปรอิสระที่ปรับแล้วนั้นไปใช้ในการประมาณขั้นตอนที่ 1 ถึง 4 ต่อไป การ

ประมาณค่าในข้างต้นจะเป็นการประมาณค่าแบบทวนซ้ำจนกระทั่งลูกโซ่ของพารามิเตอร์ที่สร้างขึ้นนั้นลู่เข้าสู่การแจกแจงความหน้าจะเป็นภายหลังซึ่งจะต้องมีการตรวจสอบการลู่เข้าตามที่ได้กล่าวไว้แล้วในส่วนทฤษฎีที่เกี่ยวข้องตอนที่ 2

จากแนวคิดของวิธี adjusted MCMC ในข้างต้น Goldstein และคณะ (Goldstein, Kounail, and Robinson, 2008; George and Goldstein, 2007) ได้พัฒนาขยายแนวคิดของอัลกอริทึมดังกล่าวให้มีความเป็นกรณีทั่วไปมากขึ้น กล่าวคือ 1) สามารถใช้กับโมเดลเชิงเส้นแบบพหุระดับที่มีความคลาดเคลื่อนจากการวัดทั้งในตัวแปรตามและตัวแปรอิสระ 2) ไม่จำเป็นที่ตัวแปรทุกตัวจะต้องมีความคลาดเคลื่อนจากการวัด และ 3) ความคลาดเคลื่อนจากการวัดสามารถสัมพันธ์กันได้ สมมติว่าในโมเดลเชิงเส้นพหุระดับที่ใช้มีตัวแปรอิสระจำนวน p ตัวที่มีความคลาดเคลื่อนจากการวัด และมีตัวแปรอิสระจำนวน q ตัวที่ไม่มีความคลาดเคลื่อนจากการวัด เราจะสามารถเขียนโมเดลเชิงเส้นพหุระดับแบบ 2 ระดับในรูปโมเดลรวม (combined model) ได้ดังนี้

$$y_{ij} = \lfloor X_{1ij}(\beta_1 + Z_{1ij} \cdot U_{1j}) \rfloor + \lfloor X_{2ij}(\beta_2 + Z_{2ij} \cdot U_{2j}) \rfloor + e_{ij} \qquad (2.139)$$

โดยที่ X_1 คือเมทริกซ์ของตัวแปรอิสระแฝงที่มีความคลาดเคลื่อนจากการวัดขนาด $\left(\sum_j n_j \times p\right)$ X_2 คือเมทริกซ์ของตัวแปรอิสระที่ไม่มีความคลาดเคลื่อนจากการวัดขนาด $\left(\sum_j n_j \times q\right)$ Z_1 และ Z_2 คือ design matrix ของอิทธิพลสุ่มในระดับที่ 2 ขนาด $(p \times 1)$ และ $(q \times 1)$ ตามลำดับ

นอกจากนี้ยังกำหนดให้การแจกแจงความน่าจะเป็นของตัวแปรสังเกตได้และการแจกแจงความน่าจะเป็นก่อนหน้าของตัวแปรอิสระแฝง มีการแจกแจงแบบปกติหลายตัวแปร ดังนี้

$$X_1^o \sim MVN(X_1, \Omega_m) \text{ และ } X_1 \sim MVN(\theta, \Omega_\Phi) \qquad (2.140)$$

โดยที่ X_1^o คือเมทริกซ์ของตัวแปรอิสระสังเกตได้ Ω_m คือเมทริกซ์ความแปรปรวนร่วมของความคลาดเคลื่อนจากการวัด (สมมติให้เมทริกซ์ดังกล่าวคงที่ทุกหน่วยตัวอย่างในระดับที่ 1) θ คือเวกตเตอร์ค่าเฉลี่ยของตัวแปรแฝงที่แท้จริง และ Ω_Φ คือเมทริกซ์ความแปรปรวนร่วมของตัวแปรอิสระแฝง (สมมติว่าทราบค่า)

อัลกอริทึมในการประมาณค่าพารามิเตอร์ในโมเดลนั้นจะใช้อัลกอริทึมการสุ่มตัวอย่างแบบกิบบ์เช่นเดียวกับวิธี MCMC ที่ใช้ในการประมาณพารามิเตอร์อิทธิพลคงที่และอิทธิพลสุ่มของ

โมเดลเชิงเส้นพหุระดับแบบดั้งเดิม ซึ่งจะมีทั้งหมด 4 ขั้นตอนดังรายละเอียดที่ได้กล่าวไว้ในข้างต้น แต่จะมีการเพิ่มขั้นตอนในการประมาณค่าตัวแปรอิสระที่แท้จริงซึ่งเป็นกระบวนการเพื่อลดทอนความคลาดเคลื่อนจากการวัดให้มีน้อยลงอีก 4 ขั้นตอน รายละเอียดของขั้นตอนที่เพิ่มขึ้นมามีดังต่อไปนี้

1) ขั้นตอนการสุ่มตัวอย่างพารามิเตอร์ θ

กำหนดให้การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ θ มีการแจกแจงแบบสม่ำเสมอ ดังนั้นจะได้ว่าการแจกแจงความน่าจะเป็นของพารามิเตอร์ θ เมื่อกำหนด X_1, Ω_Φ คือ

$$\theta \mid X_1, \Omega_\Phi \sim MVN(\hat{\theta}, \hat{V}_\theta) \qquad \text{เมื่อ} \qquad \hat{\theta} = \overline{X}_1^0 \text{ และ } \hat{V}_\theta = \frac{\Omega_\Phi}{\sum_j n_j} \tag{2.141}$$

2) ขั้นตอนการสุ่มตัวอย่างพารามิเตอร์ Ω_Φ

กำหนดให้การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ Ω_Φ มีการแจกแจงแบบอินเวอรส์Wishart ในกรณีนี้จะถือว่าผู้วิจัยมีสารสนเทศเกี่ยวกับพารามิเตอร์ Ω_Φ น้อยมากจึงเลือกกำหนดพารามิเตอร์องศาความเป็นอิสระของการแจกแจงเป็น $\sum_j n_j - 3$ ซึ่งจะเทียบเท่ากับการแจกแจงความน่าจะเป็นแบบสม่ำเสมอ ดังนั้นจะได้ว่าการแจกแจงความน่าจะเป็นของพารามิเตอร์ Ω_Φ เมื่อกำหนด X_1, θ คือ

$$\Omega_\Phi^{-1} \mid X_1, \theta \sim Wishart(\sum_j n_j - 3, \left[(X_1 - \hat{\theta})^T (X_1 - \hat{\theta})\right]^{-1}) \tag{2.142}$$

3) ขั้นตอนการสุ่มตัวอย่างพารามิเตอร์ Ω_m

ผู้วิจัยสามารถเลือกสุ่มตัวอย่าง Ω_m จากการแจกแจงความน่าจะเป็นก่อนหน้าซึ่งสามารถกำหนดได้หลายวิธี วิธีการแรกคือการกำหนดให้การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ Ω_m ให้มีการแจกแจงความน่าจะเป็นแบบอินเวอร์ส Wishart กล่าวคือ $\Omega_m^{-1} \sim Wishart(\delta_p, \delta_p S_m)$ โดยที่ในกรณีที่ผู้วิจัยมีสารสนเทศเกี่ยวกับความคลาดเคลื่อนจากการวัดน้อยผู้วิจัยอาจกำหนดให้พารามิเตอร์องศาความเป็นอิสระ δ_p มีค่าเท่ากับลำดับ (order) ของเมทริกซ์ Ω_m และ S_m คือเมทริกซ์ความแปรปรวนร่วมของความคลาดเคลื่อนจากการวัดที่กำหนดจากหลักฐานที่มีอยู่หรือจากทฤษฎี

4) ขั้นตอนการสุ่มตัวอย่าง X_1

เนื่องจากการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของ X_1 เมื่อกำหนด $y, X_1^o, \beta, U, \sigma_e^2, \Omega_\Phi$ และ Ω_m สามารถหาได้จาก

$$p(X_1 \mid y, X_1^o, \beta, U, \sigma_e^2, \Omega_\Phi, \Omega_m) = p(y \mid X_1, \beta, U, \sigma_e^2) p(X_1^o \mid X_1, \Omega_m) p(X_1 \mid \Omega_m) \quad (2.143)$$

จากสมการ (2.138) จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขดังกล่าวมีการแจกแจงแบบปกติหลายตัวแปรที่มีเวกเตอร์ค่าเฉลี่ยเท่ากับ $\hat{X}_{1ij}$ และเมทริกซ์ความแปรปรวนร่วมเท่ากับ $\hat{V}_{1ij}$ กล่าวคือ

$$X_1 \mid y, X_1^o, \beta, U, \sigma_e^2, \Omega_\Phi, \Omega_m \sim MVN(\hat{X}_{1ij}, \hat{V}_{1ij}) \quad (2.144)$$

เมื่อ $$\hat{V}_{1ij} = \left[\frac{(\beta_1 + Z_1 \cdot U_{1j})(\beta_1 + Z_1 \cdot U_{1j})^T}{\sigma_e^2} + \Omega_m^{-1} + \Omega_\Phi^{-1} \right]^{-1}$$

และ $$\hat{X}_{1ij} = \hat{V}_{1ij} \left[\left[\frac{(\beta_1 + Z_1 \cdot U_{1j})(y_{ij} - X_{2ij}(\beta_2 + Z_2 \cdot U_{2j}))}{\sigma_e^2} + X_{1ij}^o \Omega_m^{-1} + \theta \Omega_\Phi^{-1} \right] \right] \quad (2.145)$$

ในปี ค.ศ. 2007 Ferrao, Leckie และ Goldstein ได้ทำการวิจัยโดยการเปรียบเทียบผลการวิเคราะห์ข้อมูลพหุระดับระหว่างการใช้วิธีการประมาณค่าด้วยวิธี MCMC แบบดั้งเดิมและวิธี adjusted MCMC โดยใช้ข้อมูลเชิงประจักษ์จากฐานข้อมูลนักเรียนระดับชาติในประเทศอังกฤษ (National Pupil Database: NPD) ประกอบไปด้วยข้อมูลตัวอย่างของนักเรียนจำนวน 8901 คน จากโรงเรียนทั้งหมด 244 โรงเรียน โมเดลที่ใช้คือโมเดลสัมประสิทธิ์ความถดถอยสุ่มแบบสองระดับเพื่อวิเคราะห์ประสิทธิผลของโรงเรียน (school effecitiveness) รายละเอียดของโมเดลเป็นดังนี้

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{1ij} + \beta_{2j} x_{2ij} + \beta_{3j} x_{3ij} + \beta_{4j} x_{4ij} + \beta_{5j} x_{5ij} + \varepsilon_{ij} \quad (2.146)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2.147)$$

$$\beta_{1j} = \gamma_{01} + u_{1j} \quad (2.148)$$

โดยที่ $\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim MVN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \\ \tau_{01} & \tau_{11} \end{pmatrix} \right]$ และ $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$

เมื่อ y_{ij} คือค่ามาตรฐานของคะแนนผลสัมฤทธิ์วิชาคณิตศาสตร์ของนักเรียนเมื่ออายุ 11 ปี

x_{1ij} คือค่ามาตรฐานของคะแนนผลสัมฤทธิ์วิชาคณิตศาสตร์ของนักเรียนเมื่ออายุ 7 ปี

x_{2ij} คือค่ามาตรฐานของคะแนนทักษะการเขียนของนักเรียนเมื่ออายุ 7 ปี

x_{3ij} คือค่ามาตรฐานของคะแนนทักษะการอ่านของนักเรียนเมื่ออายุ 7 ปี

x_{4ij} คือตัวแปรบ่งชี้การมีสิทธิ์ที่จะได้รับยกเว้นค่าเล่าเรียน

x_{5ij} คือตัวแปรบ่งชี้ความต้องการการศึกษาแบบพิเศษ

จากผลการวิเคราะห์ด้วยวิธี MCMC แบบดั้งเดิมและวิธี adjusted MCMC ในงานวิจัยดังกล่าวจะพบว่าเมื่อพิจารณาค่าประมาณพารามิเตอร์การวิเคราะห์ด้วยวิธี MCMC แบบดั้งเดิมที่ไม่ได้มีการคำนึงถึงความคลาดเคลื่อนจากการวัดในตัวแปรอิสระในระดับที่ 1 จะประมาณค่าพารามิเตอร์อิทธิพลคงที่และความแปรปรวนในระดับที่ 1 แตกต่างจากวิธี adjusted MCMC ในขณะที่พารามิเตอร์ในส่วนอื่นๆนั้นไม่ได้ค่าประมาณจากวิธีการประมาณค่าพารามิเตอร์ทั้งสองไม่ได้มีความแตกต่างกันอย่างเห็นได้ชัดเจน ผลการวิเคราะห์นี้มีรูปแบบที่สอดคล้องกับงานวิจัยที่ได้ทำการวิจัยในทำนองเดียวกันนี้ได้แก่งานวิจัยของ Goldstein และคณะ (2007) และงานวิจัยของ Ferrao และ Goldstein (2007) ดังนั้นจึงสามารถสรุปได้ว่าผลกระทบของความคลาดเคลื่อนจากการวัดของตัวแปรอิสระในระดับที่ 1 นั้นจะส่งผลกระทบต่อค่าประมาณพารามิเตอร์ในส่วนของอิทธิพลคงที่ และค่าคลาดเคลื่อนมาตรฐานโดยเฉพาะในส่วนของอิทธิพลคงที่และความแปรปรวนในระดับที่ 1 ในขณะที่ค่าประมาณของส่วนประกอบความแปรปรวนในระดับที่ 2 จะไม่ได้รับผลกระทบจากความคลาดเคลื่อนจากการวัดดังกล่าว (Ferrao, Leckie & Goldstein,2007; Goldstein, 2007)

จากทฤษฎีที่กล่าวในข้างต้นพบว่า วิธี adjusted MCMC ถึงแม้จะสามารถแก้ไขปัญหาความคลาดเคลื่อนจากการวัดได้อย่างมีประสิทธิภาพแต่มีข้อด้อยสำหรับการใช้ในงานวิจัยทางสังคมศาสตร์กล่าวคือ เป็นวิธีการที่มีข้อสมมติว่าตัวแปรสังเกตได้ที่ใช้ในการวัดตัวแปรแฝงต้องมีความสำคัญที่เท่าเทียมกัน ดังนั้นการใช้แนวทางของโมเดลสมการโครงสร้างพหุระดับจึงเป็นวิธีที่มีประสิทธิภาพมากกว่าสำหรับการแก้ไขปัญหาดังกล่าว อย่างไรก็ตามในงานวิจัยที่มีจำนวนหน่วยจำนวนมากจะพบว่าข้อสมมติในการวัดเกี่ยวกับความไม่แปรเปลี่ยนในการวัด (measurement invariance) มักถูกละเมิด การวิเคราะห์โดยใช้โมเดลการวัดที่พารามิเตอร์ในโมเดลเป็นพารามิเตอร์แบบคงที่จะให้ผลการวิเคราะห์ที่คลาดเคลื่อน ในสถานการณ์ดังกล่าวโมเดลการ

วิเคราะห์ที่เหมาะสมคือการยอมให้พารามิเตอร์ในโมเดลการวัดเป็นพารามิเตอร์แบบสุ่ม ซึ่งการใช้สถิติแบบดั้งเดิมเพื่อประมาณค่าพารามิเตอร์ในโมเดลดังกล่าวมักไม่สามารถกระทำได้เนื่องจากติดข้อจำกัดในเชิงเทคนิคของวิธีการประมาณเพราะจำนวนมิติของการอินทิเกรตจะมีจำนวนมากเกินไป การแก้ปัญหาดังกล่าวสามารถกระทำได้โดยการเปลี่ยนไปใช้วิธีการประมาณค่าพารามิเตอร์แบบเบส์ซึ่งสามารถประมาณค่าพารามิเตอร์ได้ถึงแม้ว่าจำนวนพารามิเตอร์ในโมเดลจะมีจำนวนมากก็ตาม

ตอนที่ 5: กรอบแนวคิดการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องผู้วิจัยได้นำมาพัฒนากรอบแนวคิดการวิจัยดังแสดงในรูปที่ 2.6 ซึ่งจะเห็นว่ากรอบแนวคิดของการวิจัยดังกล่าวสามารถแบ่งออกได้เป็นสามส่วนตามวัตถุประสงค์ของการวิจัยได้แก่ ส่วนของอัลกอริทึมการประมาณค่าพารามิเตอร์แบบเบส์ ส่วนของการศึกษาด้วยการจำลองแบบมอนติคาร์โล และส่วนของการศึกษาด้วยข้อมูลทุติยภูมิ

กรอบแนวคิดการวิจัยสำหรับส่วนของการพัฒนาอัลกอริทึมการประมาณค่าพารามิเตอร์แบบเบส์ ผู้วิจัยสนใจพัฒนาวิธีการประมาณค่าพารามิเตอร์ภายใต้โมเดลในสมการที่ (2.149) ถึง (2.150) ดังต่อไปนี้

<u>โมเดลการวัด (measurement model)</u>

$$\underline{y}_{ij} = \underline{\mu} \quad + \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij} \tag{2.149}$$

<u>โมเดลสมการโครงสร้าง (structural model)</u>

$$\eta_{ij} = \beta_j^T \underline{\Xi}_{ij} + \delta_{ij} = \underline{\gamma}^T \underline{\Xi}_{ij} + \underline{u}_j^T \underline{\Xi}_{ij} + \delta_{ij} \tag{2.150}$$

เมื่อ $\underline{\Xi}_{ij} = \begin{pmatrix} 1 & \underline{\xi}_{ij} \end{pmatrix}^T$ และ $\beta_j = \underline{\gamma} + \underline{u}_j$,

$\underline{\epsilon}_{ij} \sim N(\underline{0}, \Psi_{\epsilon j})$ โดยที่ $\Psi_{\epsilon j} = diag(\psi_{\epsilon k j})$,

$\underline{\omega}_{ij} = \left(\eta_{ij}, \underline{\xi}_{ij}\right)^T \sim N\left(\underline{\theta}_{\omega j}, \Sigma_{\omega j}\right)$

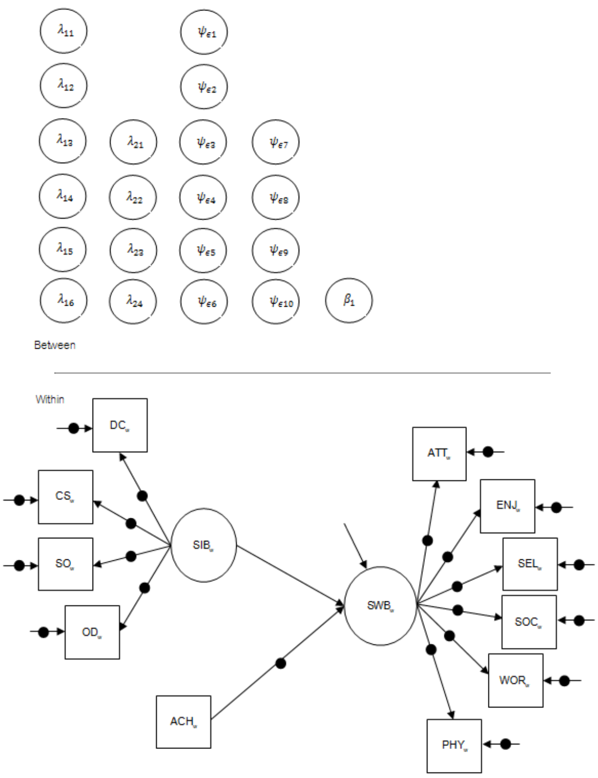
โดยที่ $\underline{\theta}_{\omega j} = \begin{pmatrix} \underline{\beta_j^T} \underline{\nu}_j \\ \underline{\nu}_j \end{pmatrix}$, $\Sigma_{\omega j} = \begin{bmatrix} \underline{\beta_j} \Phi \underline{\beta_j^T} + \sigma_\delta^2 & \underline{\beta_j^T} \Phi \\ \underline{\beta_j} \Phi & \Phi \end{bmatrix}$

และ $\delta_{ij} \sim N(\underline{0}, \sigma_\delta^2)$

การศึกษาด้วยข้อมูลจำลองโดยวิธีการจำลองแบบมอนติคาร์โล ผู้วิจัยมุ่งที่จะศึกษาเพื่อตรวจสอบความถูกต้องและเปรียบเทียบความสามารถของการประมาณค่าพารามิเตอร์ในมุมมองของความแม่นยำและประสิทธิภาพ จากทฤษฎีและงานวิจัยที่เกี่ยวข้องที่ได้ศึกษาปัจจัยที่ส่งผลต่อความสามารถในการประมาณค่าพารามิเตอร์ในโมเดลที่ใช้ในการศึกษาได้แก่ ค่าความเที่ยง (มณฑิรา ดวงสาพล, 2550; วุฒิพงษ์ เดโชดมพันธ์, 2546; Srisuttiyakorn, and Kanjanawasee, 2009; Browne, Ferrao, Leckie, and Goldstein, 2008; Goldstein, Kounali, and Robinson, 2008; Goldstein, Woodhouse, and Yang, 2001) จำนวนตัวอย่าง (สิวะโชติ ศรีสุทธิยากร, 2550; Raudenbush, and Bryk, 2002; Mass, and Hox, 2001) จำนวนตัวแปรแฝงหรือจำนวนตัวแปรที่มีความคลาดเคลื่อนจากการวัดในโมเดล (มณฑิรา ดวงสาพล, 2550; Lee, and Song, 2004; Paris, 2004; Nounou, Bakshi, Goil, and Shen, 2001) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดล (Albert, 2009; Gelman, 2006; Browne, 1998; Gelman, Carlin, Stern, and Rubin, 1995) และระดับของความผันแปรระหว่างกลุ่ม (สิวะโชติ ศรีสุทธิยากร, 2550; Mass and Hox, 2001)

กรอบแนวคิดการวิจัยสำหรับส่วนที่สาม ซึ่งเป็นส่วนของการศึกษาด้วยข้อมูลเชิงประจักษ์ ผู้วิจัยใช้กรอบแนวคิดจากวิทยานิพนธ์ระดับดุษฎีบัณฑิต สาขาวิชาวิธีวิทยาวิจัยทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เรื่อง “อิทธิพลของพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลและสุขภาวะครูที่มีต่อสุขภาวะของนักเรียน : โมเดลการปรับและการส่งผ่านพหุระดับ” (ถมรัตน์ ศิริภาพ, 2554) โดยทำการเลือกตัวแปรมาเพียงบางส่วนของกรอบแนวคิดใหญ่ด้วยเหตุผลเพื่อให้โมเดลการวิเคราะห์มีความสอดคล้องกับโมเดลที่ใช้ในการศึกษาในส่วนแรก ตัวแปรที่เลือกมาเพื่อใช้ในการศึกษาได้แก่ 1) ตัวแปรตามแฝงสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) วัดได้จากตัวแปรสังเกตได้จำนวน 6 ตัวแปรได้แก่ เจตคติและอารมณ์เชิงบวก (AFF) ความเพลิดเพลิน (ENJ) อัตมโนทัศน์เชิงวิชาการ (SEL) ปัญหาทางสังคม (SOC) ความวิตกกังวล (WOR) และปัญหาสุขภาพกาย (PHY) 2) ตัวแปรอิสระแฝงพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) วัด

ได้จากตัวแปรสังเกตได้จำนวน 4 ตัวแปรได้แก่ การใช้อำนาจ-การร่วมมือ (DC) การร่วมมือ-การคล้อยตาม (CS) การคล้อยตาม-การต่อต้าน (SO) และการต่อต้าน-การใช้อำนาจ (OD) และ 3) ตัวแปรอิสระผลสัมฤทธิ์ทางการเรียนในระดับนักเรียน (ACH_w)พิจารณาจากเกรดเฉลี่ยสะสมของนักเรียน



รูปที่ 2.6 กรอบแนวคิดการวิจัย

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยเรื่อง “วิธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด: การศึกษาสถานการณ์จำลองแบบมอนติคาร์โลจากข้อมูลจริง” มีวัตถุประสงค์ 3 ประการได้แก่ 1) เพื่อพัฒนาวีธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด 2) เพื่อศึกษาและตรวจสอบความสามารถของวิธีการประมาณค่าแบบเบส์ที่พัฒนาขึ้นและเปรียบเทียบความสามารถของการประมาณค่าพารามิเตอร์แบบภาวะความควรจะเป็นสูงสุดแบบ restricted ที่ใช้ในโปรแกรม Mplus และ 3) เพื่อทดลองใช้วิธีการประมาณค่าแบบเบส์ที่พัฒนาขึ้นเพื่อประมาณค่าพารามิเตอร์โดยใช้ข้อมูลจริง

รายละเอียดของวิธีการดำเนินงานวิจัยแบ่งออกเป็น 2 ส่วน ส่วนแรกเป็นการวิจัยเชิงทดลอง ศึกษาโดยใช้สถานการณ์จำลองแบบมอนติคาร์โล (Monte Carlo simulation) เพื่อตรวจสอบและศึกษาความสามารถของวิธีประมาณพารามิเตอร์แบบเบส์ที่ผู้วิจัยพัฒนาขึ้น และเปรียบเทียบความสามารถในการประมาณค่าพารามิเตอร์กับวิธีการประมาณค่าพารามิเตอร์จากโปรแกรม Mplus โดยใช้เกณฑ์การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (MSE) ส่วนที่

สองเป็นการทดลองใช้วิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่ผู้วิจัยได้พัฒนาขึ้นกับการวิเคราะห์ข้อมูลจริงและเปรียบเทียบผลการวิเคราะห์ที่ได้กับการวิเคราะห์ด้วยโปรแกรม Mplus

ส่วนที่ 1: การศึกษาโดยใช้การจำลองด้วยเทคนิคมอนติคาร์โล

งานวิจัยในส่วนนี้ผู้วิจัยได้ทำการตรวจสอบความถูกต้องของวิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้นโดยเปรียบเทียบกับการประมาณค่าด้วยวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted ที่ประมาณด้วยโปรแกรม Mplus สถานการณ์จำลองประกอบไปด้วยการกำหนดให้

<u>โมเดลการวัด (measurement model)</u>

$$\underline{y}_{ij} = \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij} \tag{3.1}$$

<u>โมเดลสมการโครงสร้าง (structural model)</u>

$$\eta_{ij} = \underline{\gamma}^T \underline{\xi}_{ij} + \underline{u}_j^T \underline{\xi}_{ij} + \delta_{ij} \tag{3.2}$$

ขั้นตอนการดำเนินงานวิจัยเป็นดังต่อไปนี้

1) กำหนดขนาดของประชากร $N = 150000$ โดยแบ่งเป็น

 1.1) จำนวนหน่วยตัวอย่างในระดับที่ 1 จำนวน 3000 หน่วย

 1.2) จำนวนหน่วยในระดับที่ 2 จำนวน 5000 หน่วย

2) สร้างตัวแปรอิสระแฝงจำนวน 1 ตัวจากการแจกแจงความน่าจะเป็นแบบปกติดังนี้ $\underline{\xi}_{ij} \sim N(0, \Phi = 1)$

3) สร้างความคลาดเคลื่อนสุ่มของโมเดลสมการโครงสร้างดังนี้

 3.1) สร้างความคลาดเคลื่อนสุ่มของโมเดลระดับที่ 1 จากการแจกแจงความน่าจะเป็นแบบปกติดังนี้ $\delta_{ij} \sim N\left(0, \sigma_\delta^2 = 0.8\right)$

 3.2) สร้างความคลาดเคลื่อนสุ่มของโมเดลระดับที่ 2 จากการแจกแจงความน่าจะเป็นแบบปกติดังนี้ $\underline{u}_j \ = u_j \sim N(0,0.2)$

4) กำหนดค่าสัมประสิทธิ์ความถดถอยหรือค่าพารามิเตอร์อิทธิพลคงที่ $\gamma = 1$

5) สร้างตัวแปรตาม η_{ij} จากโมเดลในสมการที่ (3.2)

6) สร้างค่าพารามิเตอร์ความเที่ยงรวมแบบสุ่ม (random composite reliability) ของตัวแปรตามและตัวแปรอิสระแฝงจากการแจกแจงความน่าจะเป็นแบบปกติดังนี้

$$\rho_{mj}^2 \sim N\left(\overline{\rho}^2, \sigma_\rho^2\right)$$

เมื่อ ρ_{mj}^2 คือค่าความเที่ยงรวมของตัวแปรแฝงที่ m ในหน่วยที่ j

โดยที่ ค่าเฉลี่ยของพารามิเตอร์ความเที่ยงแบบสุ่มแบ่งออกเป็น 4 กรณีได้แก่ $\overline{\rho^2}$= 0.3, 0.5, 0.7 และ 0.9 และความแปรปรวนของพารามิเตอร์ความเที่ยงแบบสุ่มของแต่ละกรณีกำหนดให้มีค่าเท่ากับ σ_ρ^2=0.0009, 0.0025, 0.0049 และ 0.0081 ตามลำดับ

7) สร้างค่าพารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่มจากการแจกแจงความน่าจะเป็นแบบปกติดังนี้

$$\Lambda_{kj} \sim N(E[\Lambda_k], cov[\Lambda_k])$$

โดยที่ Λ_{kj} คือเวกเตอร์ของสมาชิกใน $E[\Lambda_k] = (1,0.8,0.6)^T$ และ

$$cov[\Lambda_k] = \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0.0064 & 0 \\ 0 & 0 & 0.0036 \end{pmatrix}$$

8) สร้างค่าพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่ม โดยกำหนดให้พารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดของตัวแปรสังเกตได้ในหน่วยที่ j มีค่าเท่ากันกล่าวคือ $\psi_{\epsilon kj} = \psi_{\epsilon k' j}\ \forall k \neq k'$ และกำหนดค่าโดยการคำนวณกลับจากสูตรของความเที่ยงรวมสำหรับกลุ่มที่ j

$$\rho_{mj}^2 = \frac{\left(\sum_k \lambda_{kmj}\right)^2}{\left(\sum_k \lambda_{kmj}\right)^2 + \sum_k \psi_{\epsilon kj}}$$

9) สร้างค่าตัวแปรสังเกตได้ $\underline{y}_{ij}$ จากโมเดลในสมการที่ (3.1)

10) เมื่อได้ประชากรตามขนาดที่ต้องการแล้ว ให้สุ่มตัวอย่างหน่วยตัวอย่างในระดับที่ 2 โดยให้มีขนาดเท่ากับ 15, 30 และ 50 หน่วย และภายในแต่ละหน่วยให้สุ่มหน่วยตัวอย่างในระดับที่ 1 จำนวน 30 หน่วยตัวอย่าง โดยการสุ่มตัวอย่างในแต่ละขนาดให้สุ่มซ้ำจำนวน 100 ชุดตัวอย่าง

11) ทำการประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณแบบเบส์ที่พัฒนาขึ้นโดยใช้โปรแกรม R และวิธีภาวะความควรจะเป็นสูงสุดโดยใช้โปรแกรม Mplus

12) ทำการทดลองซ้ำ 100 ครั้งโดยใช้ข้อมูล 100 ชุดที่สุ่มมาในขั้นตอนที่ (10)

13) ผู้วิจัยแบ่งการศึกษาและเปรียบเทียบค่าประมาณพารามิเตอร์ในโมเดลออกเป็น 6 เซต ดังนี้

1. เซตของพารามิเตอร์น้ำหนักองค์ประกอบ
2. เซตของพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด
3. เซตของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมของตัวแปรแฝง
4. เซตของพารามิเตอร์อิทธิพลคงที่
5. เซตของพารามิเตอร์ความแปรปรวนในระดับที่หนึ่ง
6. เซตของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมในระดับที่สอง

ในแต่ละเซตของพารามิเตอร์ผู้วิจัยเลือกใช้เกณฑ์ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (mean square error: MSE) เพื่อใช้ในการศึกษาค่าประมาณของพารามิเตอร์ดังกล่าว ซึ่งสามารถคำนวณได้ดังสมการที่ (3.3)

$$MSE_l = \frac{1}{rep}\sum_{r=1}^{rep}\left(\hat{\theta}_{lr} - \theta_l\right)^2 \qquad (3.3)$$

เมื่อ $\hat{\theta}_{lr}$ เป็นของค่าประมาณพารามิเตอร์ตัวที่ l จากชุดข้อมูลจำลองที่ r, θ_l เป็นค่าพารามิเตอร์ที่แท้จริงตัวที่ l และ rep คือจำนวนรอบของการจำลองโดยในการศึกษานี้จะกระทำซ้ำจำนวน 100 รอบ และเพื่อความสะดวกในการศึกษาสำหรับเซตของพารามิเตอร์ที่มีจำนวนพารามิเตอร์หลายตัวผู้วิจัยจึงใช้ค่าเฉลี่ยของ MSE (average of mean square error: $\overline{MSE}$) ซึ่งสามารถคำนวณได้ดังนี้

$$\overline{MSE} = \frac{1}{L}\sum_{l=1}^{L} MSE_l \qquad (3.4)$$

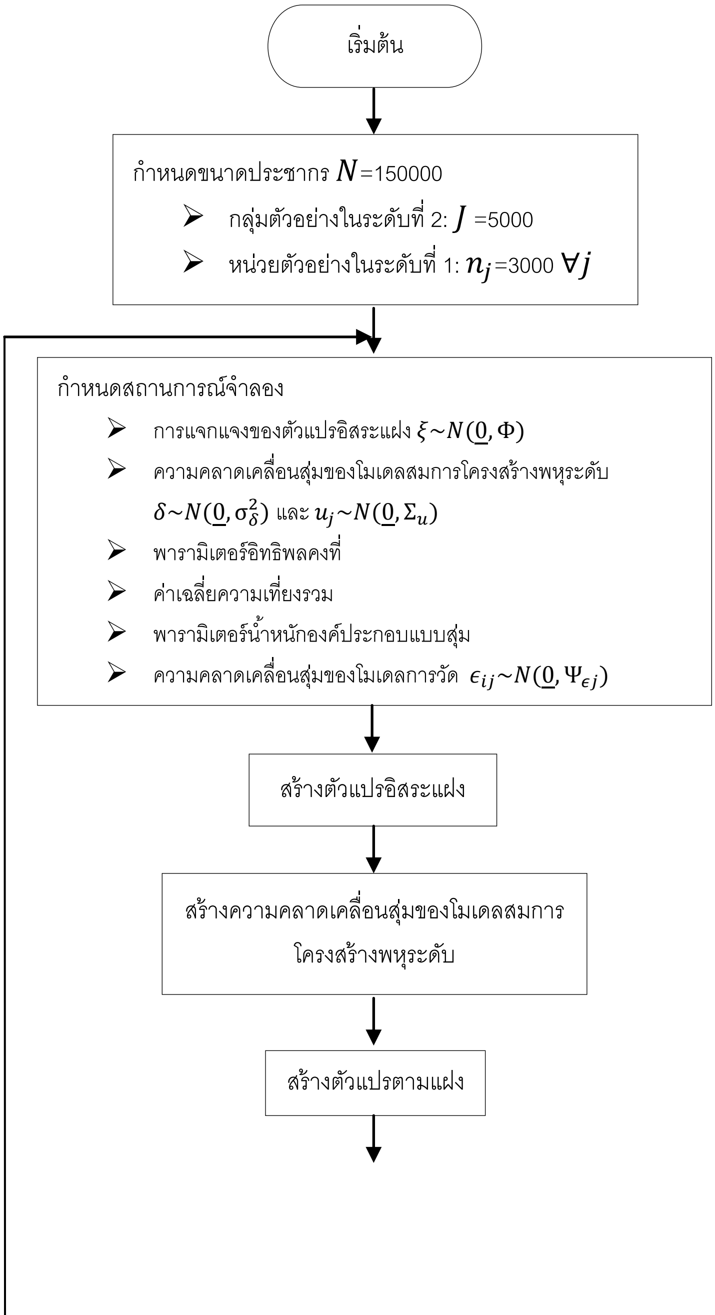
ในการเปรียบเทียบประสิทธิภาพเชิงสัมพัทธ์ผู้วิจัยใช้ตัวสถิติอัตราส่วนของค่าคลาดเคลื่อนกำลังสองระหว่างวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted เทียบกับวิธีการประมาณค่าแบบเบส์ เขียนแทนด้วย RAMSE ซึ่งสามารถคำนวณได้ดังนี้

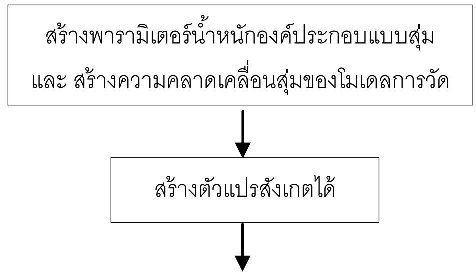
$$RAMSE \quad = \frac{\overline{MSE}_{MLR}}{\overline{MSE}_{MLR}} \qquad (3.5)$$

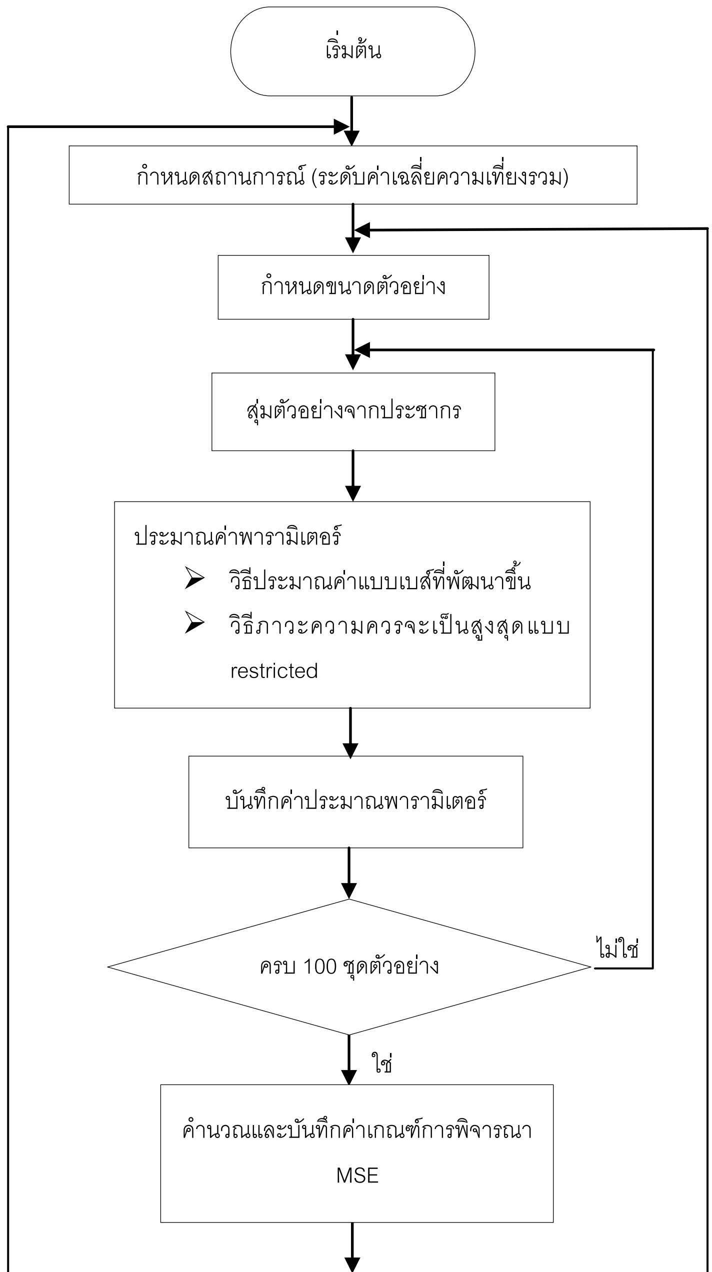
โดยที่ $\overline{MSE}_{Bayes}$ คือค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองของวิธีการประมาณค่าแบบเบส์ และ $\overline{MSE}_{MLR}$ คือค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองของวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted

ค่า RDMSE จะเป็นตัวสถิติที่ใช้วัดว่าวิธีการประมาณค่าแบบเบส์จะให้ค่าประมาณที่ดีกว่าวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted ที่ได้จากโปรแกรม Mplus กี่เท่า

จากขั้นตอนการดำเนินงานในข้างต้นสามารถเขียนให้อยู่ในรูปผังการทำงานได้ดังรูปที่ 3.1 และ 3.2 ดังนี้




สร้างพารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่ม และ สร้างความคลาดเคลื่อนสุ่มของโมเดลการวัด
สร้างตัวแปรสังเกตได้


เริ่มต้น
กำหนดสถานการณ์ (ระดับค่าเฉลี่ยความเที่ยงรวม)
กำหนดขนาดตัวอย่าง
สุ่มตัวอย่างจากประชากร
ประมาณค่าพารามิเตอร์
วิธีประมาณค่าแบบเบส์ที่พัฒนาขึ้น
วิธีภาวะความควรจะเป็นสูงสุดแบบ restricted
บันทึกค่าประมาณพารามิเตอร์
ครบ 100 ชุดตัวอย่าง
ไม่ใช่
ใช่
คำนวณและบันทึกค่าเกณฑ์การพิจารณา MSE

<u>ส่วนที่ 2</u>: การศึกษาโดยใช้ข้อมูลทุติยภูมิ

การศึกษาจากข้อมูลทุติยภูมิจะใช้ข้อมูลจากวิทยานิพนธ์ระดับดุษฎีบัณฑิต สาขาวิชาวิธีวิทยาวิจัยทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เรื่อง "อิทธิพลของพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลและสุขภาวะครูที่มีต่อสุขภาวะของนักเรียน : โมเดลการปรับและการส่งผ่านพหุระดับ" (ถมรัตน์ ศิริภาพ, 2554)

2.1 <u>ประชากรและหน่วย</u>

<u>ประชากร</u>

ประชากรในการวิจัยครั้งนี้จำแนกได้เป็น 2 กลุ่มหลัก ได้แก่ กลุ่มประชากรนักเรียนที่เป็นนักเรียนมัธยมศึกษาตอนปลายจำนวน 1,057,389 คน และกลุ่มประชากรครู ที่เป็นครูผู้สอนในระดับมัธยมศึกษาตอนปลาย จำนวน 106,840 คน โดยกลุ่มประชากรทั้งสองเป็นนักเรียนและครูของสถานศึกษา สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ภาคการศึกษาต้น ปีการศึกษา 2554 จากโรงเรียนมัธยมศึกษาตอนปลายทั้งสิ้น 2,527 โรงใน 29,613 ห้องเรียน (สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2554: ออนไลน์ อ้างถึงใน ถมรัตน์ ศิริภาพ, 2554)

<u>หน่วย</u>

การเก็บรวบรวมข้อมูลตัวอย่างดำเนินการโดยวิธีการสุ่มแบบหลายขั้นตอน (stratified random sampling) โดยใช้ภูมิภาคและจังหวัดของประเทศเป็นเกณฑ์ในการแบ่งชั้น ดังนี้ขั้นที่ 1 การสุ่มจังหวัด สุ่ม 6 ภูมิภาคละ 3 จังหวัด โดยวิธีการสุ่มอย่างง่าย (simple random sampling) ให้ได้ 18 จังหวัด ขั้นที่ 2 การสุ่มโรงเรียน สุ่มโรงเรียนที่ได้จากจังหวัดในขั้นที่ 1 จังหวัดละ 2 โรง โดยวิธีการสุ่มอย่างง่าย ทำให้ได้จำนวนโรงเรียน 36 โรง ขั้นที่ 3 การสุ่มห้องเรียน สุ่มห้องเรียนที่ได้จากโรงเรียนในขั้นที่ 2 โรงเรียนละ 2 ห้อง โดยวิธีการสุ่มอย่างง่าย ทั้งนี้ ห้องเรียนในขั้นตอนที่ 3 จะเป็นห้องเรียนที่อยู่ต่างระดับชั้นการศึกษาและต่างสายการเรียน รวมจำนวนห้องเรียนทั้งสิ้น 72 ห้อง นักเรียนและครูที่อยู่ในห้องเรียนที่สุ่มได้ในขั้นนี้จะกลายมาเป็นหน่วยในงานวิจัย (ถมรัตน์ ศิริภาพ, 2554)

จากกระบวนการในข้างต้นทำให้มีจำนวนหน่วยตัวอย่างนักเรียนรวมทั้งสิ้น 2706 คน ติดอยู่ในหน่วยครูจำนวน 71 คนจาก 71 ห้องเรียนใน 36 โรงเรียนทั่วทุกภาคของประเทศ โดยผู้วิจัยเลือกใช้ตัวแปรจากงานวิจัยดังกล่าวจำนวน 3 ตัวแปรดังนี้

1) ตัวแปรสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) วัดได้จากตัวแปรสังเกตได้จำนวน 6 ตัวแปรได้แก่ เจตคติและอารมณ์เชิงบวก (AFF) ความเพลิดเพลิน (ENJ) อัตมโนทัศน์เชิงวิชาการ (SEL) ปัญหาทางสังคม (SOC) ความวิตกกังวล (WOR) และปัญหาสุขภาพกาย (PHY)
2) ตัวแปรพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) ได้จากตัวแปรสังเกตได้จำนวน 4 ตัวแปรได้แก่ การใช้อำนาจ-การร่วมมือ (DC) การร่วมมือ-การคล้อยตาม (CS) การคล้อยตาม-การต่อต้าน (SO) และการต่อต้าน-การใช้อำนาจ (OD)
3) ตัวแปรผลสัมฤทธิ์ทางการเรียนในระดับนักเรียน (ACH_w)

<u>2.2 ขั้นตอนดำเนินงานวิจัย</u>

1) ขอความอนุเคราะห์ข้อมูลจากคุณ ถมรัตน์ ศิริภาพ, 2554

2) จัดกระทำข้อมูลเพื่อให้สามารถนำมาวิเคราะห์ในโปรแกรมการประมาณค่าพารามิเตอร์แบบเบส์ที่ผู้วิจัยได้พัฒนาขึ้น และโปรแกรม Mplus
3) ประมาณค่าพารามิเตอร์ในโมเดลโดยใช้โปรแกรมการประมาณค่าพารามิเตอร์ที่ผู้วิจัยได้พัฒนาขึ้น และโปรแกรม Mplus
4) แปลผลและเปรียบเทียบผลการวิเคราะห์ที่ได้

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาและศึกษาคุณภาพของวิธีการประมาณค่าพารามิเตอร์แบบเบส์เพื่อแก้ไขปัญหาความคลาดเคลื่อนจากการวัดสำหรับข้อมูลพหุระดับ การศึกษาคุณภาพของวิธีการประมาณจะใช้การศึกษาด้วยวิธีการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation) ภายใต้เงื่อนไขการจำลองข้อมูลต่างๆ รวมถึงการศึกษากับข้อมูลจริง โดยการวิเคราะห์เปรียบเทียบระหว่างวิธีการที่พัฒนาขึ้นและโปรแกรม Mplus ผู้วิจัยได้แบ่งการนำเสนอผลการวิเคราะห์ออกเป็น 3 ตอนดังนี้

ตอนที่ 1 รายละเอียดของวิธีการประมาณค่าพารามิเตอร์ที่พัฒนาขึ้น

ตอนที่ 2 การวิเคราะห์ข้อมูลกรณีการจำลองข้อมูลด้วยวิธีมอนติคาร์โล

ตอนที่ 3 การวิเคราะห์ข้อมูลด้วยข้อมูลจริง

ตอนที่ 1: รายละเอียดของวิธีการประมาณค่าพารามิเตอร์ที่พัฒนาขึ้น

1.1 โมเดล (model)

จากวัตถุประสงค์ของการวิจัยที่สนใจพัฒนาวิธีการประมาณค่าพารามิเตอร์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด ผู้วิจัยจึงทำการกำหนดโมเดลการวัดเพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรสังเกตได้กับตัวแปรแฝง และเพื่อประมาณค่าตัวแปรแฝงที่ต้องการทำซึ่งจะทำให้สามารถลดทอนระดับของความคลาดเคลื่อนจากการวัดที่เกิดขึ้นทั้งตัวแปรของการวิจัยได้ แต่เนื่องจากความเป็นพหุระดับของข้อมูลโมเดลการวัดที่ใช้จึงควรที่จะสามารถอธิบายแหล่งของความผันแปรที่เกิดจากโครงสร้างพหุระดับของข้อมูลได้ด้วย ในงานวิจัยนี้สนใจศึกษาในกรณีที่ข้อมูลมีโครงสร้างแบบสองระดับ โมเดลการวัดที่เลือกใช้คือ โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันแบบสองระดับ (two-levels confirmatory factor analysis model) กำหนดให้ J เป็นจำนวนของหน่วยตัวอย่างในระดับที่ 2 (หน่วย) และ n_j เป็นจำนวนของหน่วยตัวอย่างในระดับที่ 1 จะได้ว่าโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันแบบสองระดับสามารถเขียนเป็นโมเดลทั่วไปได้ดังสมการ (4.1)

$$\underline{y}_{ij} = \underline{\mu}_j + \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij} \tag{4.1}$$

โดยที่ $i = 1,2, \dots n_j$, $j = 1,2, \dots, J$, $\underline{y}_{ij}$ คือ เวกเตอร์ของตัวแปรสังเกตได้ของหน่วยตัวอย่างที่ i ในหน่วยที่ j ขนาด $p \times 1$, $\underline{\mu}_j$ คือ เวกเตอร์ของพารามิเตอร์จุดตัดแกนของหน่วยที่ j ขนาด $p \times 1$, Λ_j คือเมทริกซ์ของน้ำหนักองค์ประกอบของหน่วยที่ j ขนาด $p \times q$, $\underline{\omega}_{ij} = \left(\eta_{ij}, \underline{\xi}_{ij}\right)^T$ คือเวกเตอร์ของตัวแปรแฝงของหน่วยตัวอย่างที่ i ในหน่วยที่ j ขนาด $q \times 1$ และ $\underline{\epsilon}_{ij}$ คือเวกเตอร์ของความคลาดเคลื่อนจากการวัดของหน่วยตัวอย่างที่ i ในหน่วยที่ j ขนาด $p \times 1$

จากโมเดลในสมการที่ (4.1) กำหนดข้อสมมติให้ พารามิเตอร์จุดตัดแกน $\underline{\mu}_j$ มีการแจกแจงความน่าจะเป็นแบบปกติที่มีค่าเฉลี่ย $\underline{\kappa}$ ขนาด $p \times 1$ และเมทริกซ์ความแปรปรวนร่วมคือ Σ_b ขนาด $p \times p$, ตัวแปรแฝง $\underline{\omega}_{ij}$ มีการแจกแจงความน่าจะเป็นแบบปกติที่มีเวกเตอร์ค่าเฉลี่ย $\underline{\nu}_j$ ขนาด $q \times 1$ และเมทริกซ์ความแปรปรวนร่วม $\Sigma_{\omega j}$ ขนาด $q \times q$ กล่าวคือ $\underline{\omega}_{ij} \sim Normal(\underline{\theta}_j, \Sigma_{\omega j})$ และความคลาดเคลื่อนจากการวัด $\underline{\epsilon}_{ij}$ กำหนดให้มีการแจกแจงความน่าจะเป็นแบบปกติที่มีเวกเตอร์ค่าเฉลี่ย $\underline{0}$ ขนาด $p \times 1$ และเมทริกซ์ความแปรปรวนร่วม $\Psi_{\epsilon j}$ ขนาด $p \times p$ กล่าวคือ $\underline{\epsilon}_{ij} \sim Normal\left(\underline{0}, \Psi_{\epsilon j}\right)$ นอกจากนี้ยังกำหนดให้ความคลาดเคลื่อนจากการวัด $\underline{\epsilon}_{ij}$ เป็นอิสระซึ่งกันและกันกับ $\underline{\omega}_{ij}$

โมเดลแสดงความสัมพันธ์ระหว่างตัวแปรที่เลือกใช้ศึกษาในงานวิจัยนี้คือโมเดลสัมประสิทธิ์ความถดถอยแบบสุ่มสองระดับ (two-levels random coefficients model) ซึ่งมีรายละเอียดดังสมการที่ (4.2 ก) และ (4.2 ข)

<u>โมเดลระดับที่หนึ่ง</u>

$$\eta_{ij} = \underline{\beta}_j^T \underline{\Xi}_{ij} + \delta_{ij} \qquad (4.2\ ก)$$

<u>โมเดลระดับที่สอง</u>

$$\underline{\beta}_j = \underline{\gamma} + \underline{u}_j \qquad (4.2\ ข)$$

โดยที่ η_{ij} คือตัวแปรตามแฝงในระดับที่ 1 ของหน่วยตัวอย่างที่ i ในหน่วยที่ j , $\underline{\xi}_{ij}$ คือเวกเตอร์ของตัวแปรอิสระแฝงในระดับที่ 1 ของหน่วยตัวอย่างที่ i ในหน่วยที่ j ขนาด $q_2 \times 1$ โดยที่กำหนดให้ $\underline{\xi}_{ij} \sim Normal(\nu_j, \Phi)$, $\underline{\beta}_j$ คือเวกเตอร์ของสัมประสิทธิ์ความถดถอยในระดับที่ 1 ของหน่วยที่ j ขนาด $q_2 \times 1$, δ_{ij} คือความคลาดเคลื่อนของโมเดลระดับที่ 1 ของหน่วยตัวอย่างที่ i ในหน่วยที่ j, $\underline{\gamma}$ คือเวกเตอร์ของสัมประสิทธิ์ความถดถอยในระดับที่ 2 ขนาด $q_2 \times 1$ และ $\underline{u}_j$ คือเวกเตอร์ของความคลาดเคลื่อนของโมเดลระดับที่ 2 ของหน่วยที่ j ขนาด $q_2 \times 1$

จากโมเดลในสมการที่ (4.2) กำหนดข้อสมมติของโมเดลให้ δ_{ij} มีการแจกแจงความน่าจะเป็นแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และมีความแปรปรวนเท่ากับ σ_δ^2 กล่าวคือ $\delta_{ij} \sim Normal(0, \sigma_\delta^2)$ และ $\underline{u}_j$ มีการแจกแจงความน่าจะเป็นแบบปกติที่มีค่าเฉลี่ย $\underline{0}$ ขนาด $q_2 \times 1$, เมทริกซ์ความแปรปรวนร่วม Σ_u ขนาด $q_2 \times q_2$ กล่าวคือ $\underline{u}_j \sim Normal(0, \Sigma_u)$, δ_{ij} เป็นอิสระซึ่งกันและกันกับ $\underline{\xi}_{ij}$ และจากการกำหนดข้อสมมติในเบื้องต้นทำให้ได้ว่า $\underline{\beta}_j \sim Normal\left(\underline{\gamma}, \Sigma_u\right)$

อย่างไรก็ตามเมื่อพิจารณาโมเดลตามสมการที่ (4.1) จะพบว่าเป็นโมเดลที่ไม่สามารถระบุได้ กล่าวคือในการวิเคราะห์ไม่สามารถให้พารามิเตอร์ทุกตัวในโมเดลเป็นพารามิเตอร์แบบสุ่มได้ทั้งหมดเพราะจะทำให้ไม่สามารถหาผลเฉลยของค่าประมาณพารามิเตอร์ได้เพียงค่าเดียว การแก้ไขปัญหาดังกล่าวสามารถกระทำได้โดยการกำหนดข้อจำกัด (restriction) ให้กับพารามิเตอร์บางตัวในโมเดล การกำหนดข้อจำกัดดังกล่าวสามารถกระทำได้หลายวิธี ในงานวิจัยนี้ผู้วิจัยได้กำหนดให้พารามิเตอร์จุดตัดแกนไม่มีความแตกต่างกันระหว่างกลุ่ม ดังนั้นจะได้ว่าโมเดลการวัดที่ใช้ในการอธิบายปรากฏการณ์ดังกล่าวคือ

$$\underline{y}_{ij} = \underline{\mu} \quad + \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij} \tag{4.3}$$

จากโมเดลในสมการที่ (4.3) จะได้ว่า

$$E\left(\underline{y}_{ij} | \Lambda_j, \underline{\nu}_j\right) = \underline{\mu} \quad + \Lambda_j \underline{\nu}_j \tag{4.4}$$

และ

$$Cov\left(\underline{y}_{ij} | \Lambda_j, \Psi_{\epsilon j}\right) = \Lambda_j \Phi \Lambda_j^T + \Psi_{\epsilon j} \tag{4.5}$$

นอกจากนี้จากข้อสมมติของสมการที่ (4.1) $\underline{\omega}_{ij} \sim N(\underline{\theta}_j, \Sigma_{\omega j})$ จะได้ว่า

$$\underline{\theta}_j = E(\underline{\omega}_{ij}) = \begin{bmatrix} E\left(\eta_{ij} | \underline{\beta}_j, \underline{\nu}_j\right) \\ E\left(\underline{\xi}_{ij}\right) \end{bmatrix} = \begin{bmatrix} \beta_{0j} + \underline{\beta}_j^T \underline{\nu}_j \\ \underline{\nu}_j \end{bmatrix} \tag{4.6}$$

และ

$$\Sigma_{\omega j} = Cov(\underline{\omega}_{ij}) = \begin{bmatrix} Var(\eta_{ij} | \underline{\beta}_j, \Phi) & Cov\left(\eta_{ij}, \underline{\xi}_{ij}\right) \\ Cov\left(\underline{\xi}_{ij}, \eta_{ij}\right) & Var\left(\underline{\xi}_{ij}\right) \end{bmatrix} = \begin{bmatrix} \underline{\beta}_j \Phi \underline{\beta}_j^T + \sigma_\delta^2 & \underline{\beta}_j^T \Phi \\ \underline{\beta}_j \Phi & \Phi \end{bmatrix} \tag{4.7}$$

***1.2 อัลกอริทึมสำหรับประมาณค่าพารามิเตอร์** (estimation algorithm)*

ขั้นตอนแรกของการประมาณค่าพารามิเตอร์แบบเบส์คือ การกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดล ผู้วิจัยได้กำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดลดังต่อไปนี้

1) $\underline{\mu} \quad \sim Normal(\kappa, \Sigma_\mu)$
2) $\underline{\nu}_j \sim Normal(0, \Delta)$ โดยที่ Δ คือเมทริกซ์ความแปรปรวนร่วมก่อนหน้าขนาด $q_2 \times q_2$
3) $\underline{\Lambda}_{kj} \sim Normal\left(\underline{g}_{kj}, H_{kj}\right)$ โดยที่ $\underline{\Lambda}_{kj}$ คือเวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์น้ำหนักองค์ประกอบ Λ_j ที่เป็น free parameters
4) $\psi_{\epsilon kj}^{-1} \sim Gamma(\alpha_{kj}, \beta_{kj})$ โดยที่ $\psi_{\epsilon kj} \in \Psi_{\epsilon j}$
5) $\left(\sigma_\delta^2\right)^{-1} \sim Gamma(\alpha_\delta, \beta_\delta)$
6) $\Phi^{-1} \sim Wishart(\rho_\phi, R_\phi)$
7) $\Sigma_u^{-1} \sim Wishart(\rho_u, R_u)$
8) $\underline{\gamma} \sim Normal\left(\underline{\gamma_0}, \Sigma_{\gamma 0}\right)$

การอนุมานเชิงสถิติแบบเบส์นั้นจะอาศัยการแจกแจงความน่าจะเป็นภายหลังเป็นเครื่องมือหลักในการอนุมาน การแจกแจงความน่าจะเป็นภายหลังสามารถคำนวณได้จากทฤษฎี

ของเบส์ในทฤษฎีความน่าจะเป็น อย่างไรก็ตามในกรณีที่โมเดลที่มีความซับซ้อนจะทำให้จำนวนพารามิเตอร์มีจำนวนมากขึ้นไปด้วย ทำให้การหาพจน์อินทิกรัลไม่สามารถกระทำได้โดยวิธีการเชิงวิเคราะห์ ผู้วิจัยจึงเลือกใช้วิธีการจำลอง (simulation based method) เพื่อสุ่มตัวอย่างพารามิเตอร์และใช้ตัวอย่างของพารามิเตอร์ที่ได้ในการประมาณการแจกแจงความน่าจะเป็นภายหลังที่ต้องการ ในกรณีนี้พบว่าพารามิเตอร์ในโมเดลมีจำนวนมากกว่าหนึ่งตัว ดังนั้นการแจกแจงความน่าจะเป็นภายหลังที่ต้องการจะเป็นการแจกแจงความน่าจะเป็นภายหลังร่วม (joint posterior distribution) ซึ่งรูปแบบของการแจกแจงความน่าจะเป็นภายหลังร่วมของพารามิเตอร์ของโมเดลในข้างต้นได้แก่ $\left\{\underline{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\underline{v}_j\}, \Phi, \{\Psi_{\epsilon j}\}, \underline{\gamma}, \{\underline{u}_j\}, \Sigma_u, \sigma_\delta^2\right\}$ จะพบว่าไม่ได้อยู่ในรูปแบบมาตรฐานที่จะสามารถสุ่มตัวอย่างโดยอาศัยวิธีการจำลองโดยตรง (direct simulation) ได้ วิธีการที่สามารถแก้ไขปัญหาข้างต้นคือการใช้เทคนิคลูกโซ่มาร์คอฟมอนติคาร์โล ผู้วิจัยเลือกใช้อัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ (Gibbs sampling algorithm) เป็นเครื่องมือในการประมาณการแจกแจงความน่าจะเป็นภายหลังที่ต้องการ อัลกอริทึมการประมาณประกอบด้วยขั้นตอนย่อยจำนวน 10 ขั้นตอนดังต่อไปนี้

1. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\mu}$ จาก

   $$\underline{\mu} \mid \left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_\mu, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\} \sim Normal\left(\hat{\kappa}, \hat{\Sigma}_\mu\right)$$

   โดยที่ $\hat{\Sigma}_\mu = \left[\Sigma_\mu^{-1} + \sum_{j=1}^{J} n_j \Psi_{\epsilon j}^{-1}\right]^{-1}$

   และ $\hat{\kappa} = \hat{\Sigma}_\mu \left[\Sigma_\mu^{-1}\kappa + \sum_{j=1}^{J} \Psi_{\epsilon j}^{-1} \sum_{i=1}^{n_j} \left(\underline{y}_{ij} - \Lambda_j \underline{\omega}_{ij}\right)\right]$

2. สำหรับหน่วยที่ j และตัวแปรสังเกตได้ที่ k สุ่มตัวอย่างพารามิเตอร์ $\underline{\Lambda}_{kj}$ จาก

   $$\underline{\Lambda}_{kj} \mid \left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj} \sim Normal(\underline{\hat{g}}_{kj}, \widehat{H}_{kj})$$

   เมื่อ $\underline{\omega}_{-ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ_j และเป็น free parameters, $\underline{\widetilde{\omega}}_{ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ_j และเป็น fix parameters, $\underline{\tilde{y}}_{ijk}$ เป็นค่าตัวแปรสังเกตได้ที่ปรับค่าที่เนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบในแถวที่ k พารามิเตอร์เมทริกซ์ Λ_j ซึ่ง $\underline{\tilde{y}}_{ijk} = y_{ijk} - \underline{1}^T \underline{\widetilde{\omega}}_{ijk} - \mu_k$

โดยที่ $\widehat{H}_{kj} = \left[H_{kj}^{-1} + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\omega}_{-ijk}^{T}\right]^{-1}$

และ $\underline{\hat{g}}_{kj} = \widehat{H}_{kj} \left[H_{kj}^{-1} \underline{g}_{kj} + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\tilde{y}}_{ijk}\right]$

3. สำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j สุ่มตัวอย่างตัวแปรแฝง $\underline{\omega}_{ij}$ จาก

$$\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}\ \ , \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij} \sim Normal\left(\underline{\hat{\theta}}_{ij}, \hat{\Sigma}_{\omega j}\right)$$

โดยที่ $\hat{\Sigma}_{\omega j} = \left[\Sigma_{\omega j}^{-1} + \Lambda_j^T \Psi_{\epsilon j}^{-1} \Lambda_j\right]^{-1}$

และ $\underline{\hat{\theta}}_{ij} = \hat{\Sigma}_{\omega j} \left[\Sigma_{\omega j}^{-1} \underline{\theta}_j + \Lambda_j^T \Psi_{\epsilon j}^{-1} \left(\underline{y}_{ij} - \underline{\mu}\ \ \right)\right]$

4. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ค่าเฉลี่ยของตัวแปรแฝง $\underline{\nu}_j$ จาก

$$\underline{\nu}_j | \left\{\underline{\xi}_{ij}\right\}, \Phi, \Delta \sim Normal(\underline{\hat{\nu}}_j, \hat{\Delta})$$

โดยที่ $\hat{\Delta} = \left[\Delta^{-1} + n_j \Phi^{-1}\right]^{-1}$ และ $\underline{\hat{\nu}}_j = \hat{\Delta} \left[\Phi^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ij}\right]$

5. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Φ^{-1} จาก

$$\Phi^{-1} | \left\{\underline{\xi}_{ij}\right\}, \underline{\nu}_j, \rho_\phi, R_\phi \sim Wishart\left(\hat{\rho}_\phi, \hat{R}_\phi\right)$$

โดยที่ $\hat{\rho}_\phi = n + \rho_\phi$

และ $\hat{R}_\phi = \left[R_\phi^{-1} + \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(\underline{\xi}_{ij} - \underline{\nu}_j\right)\left(\underline{\xi}_{ij} - \underline{\nu}_j\right)^T\right]^{-1}$

6. สำหรับตัวแปรสังเกตได้ที่ k และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\psi_{\epsilon kj}^{-1}$ จาก

$$\psi_{\epsilon kj}^{-1} | \mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \left\{\underline{y}_{ij}\right\}, \alpha_{kj}, \beta_{kj} \sim Gamma\left(\hat{\alpha}_{\epsilon kj}, \hat{\beta}_{\epsilon kj}\right)$$

โดยที่ $\hat{\alpha}_{\epsilon kj} = \frac{n_j}{2} + \alpha_{kj}$

และ $\hat{\beta}_{\epsilon kj} = \beta_{kj} + \frac{1}{2} \sum_{i=1}^{n_j} \left(y_{ij} - \mu_k - \underline{\Lambda}_k^T \underline{\omega}_{ij}\right)^T \left(y_{ij} - \mu_k - \underline{\Lambda}_k^T \underline{\omega}_{ij}\right)$

7. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\gamma}$ จาก

$$\underline{\gamma} | \{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{\gamma}}\ \ , \hat{\Sigma}_\gamma\right)$$

โดยที่ $\hat{\Sigma}_\gamma = \left[\Sigma_u^{-1} + \frac{1}{\sigma_\delta^2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \underline{\Xi}_{ij} \underline{\Xi}_{ij}^T\right]^{-1}$

และ $\underline{\hat{\gamma}}\ \ = \hat{\Sigma}_\gamma \left[\Sigma_u^{-1} \underline{\gamma} + \frac{1}{\sigma_\delta^2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \underline{\Xi}_{ij} (\eta_{ij} - \underline{\Xi}_{ij}^T \underline{u}_j)\right]$ เมื่อ $\underline{\Xi}_{ij} = \begin{pmatrix}1 & \underline{\xi}_{ij}\end{pmatrix}^T$

8. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{u}_j$ จาก

$$\underline{u}_j | \{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{u}}_j, \hat{\Sigma}_{\mathrm{u}j}\right)$$

โดยที่ $\hat{\Sigma}_{u\mathrm{j}} = \left[\frac{1}{\sigma_\delta^2}\sum_{i=1}^{n_j} \underline{\Xi}_{ij}\underline{\Xi}_{ij}^T + \Sigma_u^{-1}\right]^{-1}$

และ $\underline{\hat{u}}_j = \frac{1}{\sigma_\delta^2}\hat{\Sigma}_{u\mathrm{j}}\left[\sum_{i=1}^{n_j} \underline{\Xi}_{ij}^T\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma}\right)\right]$

9. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Σ_u^{-1} จาก

$$\Sigma_u^{-1}|\underline{u}_j, \rho_u, R_u \sim Wishart\left(\hat{\rho}_u, \hat{R}_u\right)$$

โดยที่ $\hat{\rho}_u = J + \rho_u$ และ $\hat{R}_u = \left[R_u^{-1} + \sum_{j=1}^{J}\underline{u}_j^T\underline{u}_j\right]^{-1}$

10. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $(\sigma_\delta^2)^{-1}$ จาก

$$(\sigma_\delta^2)^{-1}|\{\underline{\omega}_{ij}\}, \left\{\underline{\beta_j}\right\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta \sim Gamma\left(\hat{\alpha}_\delta, \hat{\beta}_\delta\right)$$

โดยที่ $\hat{\alpha}_\delta = \frac{\sum_{j=1}^{J} n_j}{2} + \alpha_\delta$ และ $\hat{\beta}_\delta = \beta_\delta + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma} - \underline{\Xi}_{ij}^T\underline{u}_j\right)^2$

อัลกอริทึมในข้างต้นเป็นวิธีการประมาณแบบทวนซ้ำ (iterative method) กล่าวคือวิธีการข้างต้นจะสร้างข้อมูลของพารามิเตอร์ในโมเดลจากการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขทีละขั้นตอน การแจกแจงความน่าจะเป็นของตัวอย่างพารามิเตอร์ที่สร้างขึ้นจะลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ในโมเดลที่ต้องการหากมีการจำลองด้วยจำนวนรอบของการจำลองที่เพียงพอ นอกจากนี้อัลกอริทึมในข้างต้นยังสามารถประยุกต์ใช้กับโมเดลที่มีความซับซ้อนมากขึ้นโดยการขยายอัลกอริทึมในส่วนของการประมาณค่าพารามิเตอร์ในโมเดลพหุระดับได้โดยง่าย

1.3 รายละเอียดของการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขที่ใช้ในอัลกอริทึม

กำหนดให้ $\Theta = \left\{\underline{\mu}, \{\Lambda_j\}, \{\underline{\nu}_j\}, \Phi, \{\Psi_{\epsilon j}\}, \underline{\gamma}, \{\underline{u}_j\}, \Sigma_u, \sigma_\delta^2\right\}$ เป็นเซตของพารามิเตอร์ที่ไม่ทราบค่าในโมเดล และ $\Omega = \left\{\{\underline{\omega}_{ij}\}, \{\underline{u}_j\}\right\}$ เป็นเซตของตัวแปรแฝงในโมเดล จากการวิเคราะห์การแจกแจงความน่าจะเป็นภายหลัง *(*posterior distribution analysis) จะพบว่าการแจกแจงความน่าจะเป็นภายหลังที่ต้องการในกรณีนี้คือ $p\left(\Omega, \Theta|\left\{\underline{y}_{ij}\right\}\right)$ จากแนวคิดของการสุ่มตัวอย่างแบบกิบส์ทำให้ผู้วิจัยสามารถแบ่งอัลกอริทึมสำหรับประมาณค่าพารามิเตอร์ในโมเดลออกได้เป็นสองส่วนใหญ่ได้แก่ 1) $p\left(\Omega|\Theta, \left\{\underline{y}_{ij}\right\}\right)$ และ 2) $p\left(\Theta|\Omega, \left\{\underline{y}_{ij}\right\}\right)$

พิจารณาการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ในโมเดลเมื่อกำหนดตัวแปรแฝงและข้อมูลค่าสังเกต $p\left(\Theta|\Omega, \left\{\underline{y}_{ij}\right\}\right)$ กำหนดให้ $\Theta = \left(\Theta_y, \Theta_\omega\right)^T$ โดยที่ Θ_y คือเวกเตอร์

ของพารามิเตอร์ในโมเดลการวัดได้แก่ $\underline{\mu}, \{\Lambda_j\}, \{\Psi_{\epsilon j}\}$ และ Θ_ω คือ เวกเตอร์ของพารามิเตอร์ในโมเดลสมการโครงสร้างได้แก่ $\underline{\nu}_j, \Phi, \underline{\gamma}, \sigma_\delta^2, \Sigma_u$ ดังนั้นการแจกแจงความน่าจะเป็นแบบมีเงื่อนไข

$$p\left(\Theta|\Omega, \left\{\underline{y}_{ij}\right\}\right) = p\left(\Theta_y, \Theta_\omega|\Omega, \left\{\underline{y}_{ij}\right\}\right) \propto p(\Theta_y, \Theta_\omega)p\left(\Omega, \left\{\underline{y}_{ij}\right\}|\Theta_y, \Theta_\omega\right) \quad (4.8)$$

ผู้วิจัยสมมติให้โมเดลกำหนดให้พารามิเตอร์ในโมเดลการวัดและโมเดลสมการโครงสร้างเป็นอิสระซึ่งกันและกัน ดังนั้นจะได้ว่า $p(\Theta) = p(\Theta_y) \times p(\Theta_\omega)$ นอกจากนี้จากสมการที่ (4.8) จะได้ว่าฟังก์ชันภาวะความควรจะเป็นสามารถเขียนได้เป็น

$$p\left(\Omega, \left\{\underline{y}_{ij}\right\}|\Theta_y, \Theta_\omega\right) = p(\Omega|\Theta_\omega)p\left(\left\{\underline{y}_{ij}\right\}|\Omega, \Theta_y\right) \quad (4.9)$$

ดังนั้นจะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ในโมเดลสามารถเขียนได้ดังสมการที่ (4.10)

$$p\left(\Theta_y, \Theta_\omega|\Omega, \left\{\underline{y}_{ij}\right\}\right) \propto \left[p(\Theta_y)p\left(\left\{\underline{y}_{ij}\right\}|\Omega, \Theta_y\right)\right] \times [p(\Theta_\omega)p(\Omega|\Theta_\omega)] \quad (4.10)$$

จากสมการที่ (4.10) จะเห็นว่าในการประมาณค่าพารามิเตอร์ในโมเดลการวัดและโมเดลสมการโครงสร้างสามารถกระทำแยกจากกันได้เพราะเป็นอิสระซึ่งกันและกัน โดยที่

$p(\Theta_y)p\left(\left\{\underline{y}_{ij}\right\}|\Omega, \Theta_y\right) = p\left(\underline{\mu}, \{\Lambda_j\}, \{\Psi_{\epsilon j}\}|\left\{\underline{y}_{ij}\right\}, \{\underline{\omega}_{ij}\}\right)$

และ $p(\Theta_\omega)p(\Omega|\Theta_\omega) = p\left(\{\eta_{ij}\}|\underline{\gamma}, \sigma_\delta^2, \left\{\underline{\xi}_{ij}\right\}, \{\underline{u}_j\}\right)p\left(\underline{\gamma}, \sigma_\delta^2\right)p(\{\underline{u}_j\}|\Sigma_u)p(\Sigma_u)p\left(\left\{\underline{\xi}_{ij}\right\}|\underline{\nu}_j, \Phi\right)p\left(\underline{\nu}_j, \Phi\right)$

จากการพิจารณาในข้างต้นและการใช้อัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ (Gibbs sampling algorithm) จะได้ว่าสามารถเขียนเป็นรายละเอียดของการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ในโมเดลในแต่ละขั้นตอนได้ดังนี้

1. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์* $\underline{\mu}$ *เมื่อกำหนด* $\left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_\mu, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}$

พิจารณาการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\mu}$ เมื่อกำหนดพารามิเตอร์ $\left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_\mu, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}$จากทฤษฎีของเบส์จะได้ว่า

$$p\left(\underline{\mu}|\left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_\mu, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}\right)p\left(\underline{\mu}|\kappa, \Sigma_\mu\right) \times p\left(\left\{\underline{y}_{ij}\right\}|\underline{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}\right) \quad (4.11)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\underline{\mu}$ มีการแจกแจงแบบปกติกล่าวคือ $\underline{\mu} \sim Normal(\kappa, \Sigma_\mu)$ จะได้ว่า

$$p\left(\underline{\mu}|\kappa, \Sigma_{\mu}\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\mu}-\kappa\right)^{T} \Sigma_{\mu}^{-1}\left(\underline{\mu}-\kappa\right)\right\} \tag{4.12}$$

และฟังก์ชันภาวะความควรจะเป็นคือ

$$p\left(\left\{\underline{y}_{ij}\right\}|\underline{\mu}, \kappa, \Sigma_{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}\right) \propto \prod_{j=1}^{J} \prod_{i=1}^{n_j} p\left(\underline{y}_{ij}|\underline{\mu}, \Lambda_j, \underline{\omega}_{ij}, \Psi_{\epsilon j}\right)$$

$$= \prod_{j=1}^{J} \prod_{i=1}^{n_j} exp\left\{-\frac{1}{2}\left(\underline{y}_{ij}-\underline{\mu}-\Lambda_j\underline{\omega}_{ij}\right)^{T} \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij}-\underline{\mu}-\Lambda_j\underline{\omega}_{ij}\right)\right\}$$

$$= exp\left\{-\frac{1}{2}\left[-2\underline{\mu}^{T} \sum_{j=1}^{J} \Psi_{\epsilon j}^{-1} \sum_{i=1}^{n_j}\left(\underline{y}_{ij}-\Lambda_j\underline{\omega}_{ij}\right)+\underline{\mu}^{T}\left[\sum_{j=1}^{J} n_j \Psi_{\epsilon j}^{-1}\right]\underline{\mu}\right]\right\} \tag{4.13}$$

แทนผลลัพธ์ที่ได้จากสมการที่ (4.12) และสมการที่ (4.13) ลงในสมการที่ (4.11) จะได้การแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\mu}$ เมื่อกำหนด $\left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}$ เป็น

$$p\left(\underline{\mu}|\left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\mu}-\kappa\right)^{T} \Sigma_{\mu}^{-1}\left(\underline{\mu}-\kappa\right)\right\}$$

$$\times exp\left\{-\frac{1}{2}\left[-2\underline{\mu}^{T} \sum_{j=1}^{J} \Psi_{\epsilon j}^{-1} \sum_{i=1}^{n_j}\left(\underline{y}_{ij}-\Lambda_j\underline{\omega}_{ij}\right)+\underline{\mu}^{T}\left[\sum_{j=1}^{J} n_j \Psi_{\epsilon j}^{-1}\right]\underline{\mu}\right]\right\} \tag{4.14}$$

จากสมการที่ (4.14) ทำการกระจายพจน์กำลังสองสมบูรณ์ในส่วนของการแจกแจงความน่าจะเป็นก่อนหน้า และตัดพจน์ที่เป็นค่าคงที่ออกจากสมการจะได้

$$p\left(\underline{\mu}|\left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}\right)$$

$$\propto exp\left\{-\frac{1}{2}\left[\underline{\mu}^{T}\left[\Sigma_{\mu}^{-1}+\sum_{j=1}^{J} n_j \Psi_{\epsilon j}^{-1}\right]\underline{\mu}-2\underline{\mu}^{T}\left[\Sigma_{\mu}^{-1}\kappa+\sum_{j=1}^{J} \Psi_{\epsilon j}^{-1} \sum_{i=1}^{n_j}\left(\underline{y}_{ij}-\Lambda_j\underline{\omega}_{ij}\right)\right]\right]\right\} \tag{4.15}$$

เมื่อเปรียบเทียบรูปแบบของสมการข้างต้นกับการแจกแจงแบบปกติหลายตัวแปรจะได้ว่า สำหรับหน่วยตัวอย่าง i ในหน่วยที่ j

$$\underline{\mu} \quad |\left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\} \sim Normal\left(\hat{\kappa}, \hat{\Sigma}_{\mu}\right) \tag{4.16}$$

โดยที่ $\hat{\Sigma}_{\mu} = \left[\Sigma_{\mu}^{-1}+\sum_{j=1}^{J} n_j \Psi_{\epsilon j}^{-1}\right]^{-1}$

และ $\hat{\kappa} = \hat{\Sigma}_\mu \left[\Sigma_\mu^{-1}\kappa + \sum_{j=1}^{J} \Psi_{\epsilon j}^{-1} \sum_{i=1}^{n_j} \left(\underline{y}_{ij} - \Lambda_j \underline{\omega}_{ij}\right)\right]$

2. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์* $\{\Lambda_j\}$ *เมื่อกำหนด* $\{\underline{y}_{ij}\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}$

เนื่องจากพารามิเตอร์น้ำหนักองค์ประกอบ Λ_j เป็นเมทริกซ์ ซึ่งการพิจารณาการแจกแจงความน่าจะเป็นของเมทริกซ์สุ่ม (random matrix) นั้นมีความซับซ้อนและกระทำได้ยาก เพื่อหลีกเลี่ยงความยุ่งยากดังกล่าวผู้วิจัยจึงกำหนดข้อสมมติให้น้ำหนักองค์ประกอบในแถวที่ k และแถวที่ k’ ของเมทริกซ์น้ำหนักองค์ประกอบ Λ_j เป็นอิสระซึ่งกันและกัน การสุ่มตัวอย่างพารามิเตอร์เมทริกซ์น้ำหนักองค์ประกอบที่ต้องการนั้นจึงสามารถสุ่มตัวอย่างจากเวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์ Λ_j ทีละเวกเตอร์ จากแนวคิดในข้างต้นจึงกำหนดสัญลักษณ์ดังต่อไปนี้

$\underline{\Lambda}_{kj}$ คือ เวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์ Λ_j เฉพาะที่เป็น free parameters

$\underline{\widetilde{\omega}}_{ijk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น fix parameters ในแถวที่ k ของเมทริกซ์ Λ_j

$\underline{\omega}_{-ijk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น free parameters ในแถวที่ k ของเมทริกซ์ Λ_j (กล่าวคือสอดคล้องกับเวกเตอร์ $\underline{\Lambda}_{kj}$)

$\underline{\tilde{y}}_{ijk}$ คือ เวกเตอร์ของคะแนนตัวแปรสังเกตได้ที่มีการปรับค่าเนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบในเวกเตอร์ $\underline{\Lambda}_{kj}$ โดยที่

$$\underline{\tilde{y}}_{ijk} = y_{ijk} - \underline{1}^T \underline{\widetilde{\omega}}_{ijk} - \mu_k$$

พิจารณาการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\Lambda}_{kj}$ เมื่อกำหนด $\{\underline{\tilde{y}}_{ijk}\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}$ จากทฤษฎีของเบส์จะได้ว่า

$$p\left(\underline{\Lambda}_{kj} | \{\underline{\tilde{y}}_{ijk}\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}, \underline{g}_{kj}, H_{kj}\right) \propto p\left(\underline{\Lambda}_{kj} | \underline{g}_{kj}, H_{kj}\right) p\left(\{\underline{\tilde{y}}_{ijk}\} | \underline{\Lambda}_{kj}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}\right) \quad (4.17)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\underline{\Lambda}_{kj}$ มีการแจกแจงแบบปกติ กล่าวคือ $\underline{\Lambda}_{kj} \sim Normal\left(\underline{g}_{kj}, H_{kj}\right)$ ซึ่งเขียนได้ดังนี้

$$p\left(\underline{\Lambda}_{kj}|\underline{g}_{kj}, H_{kj}\right) \propto exp\{-\tfrac{1}{2}\left(\underline{\Lambda}_{kj} - \underline{g}_{kj}\right)^T H_{kj}^{-1}\left(\underline{\Lambda}_{kj} - \underline{g}_{kj}\right)\} \quad (4.18)$$

และฟังก์ชันภาวะความควรจะเป็นคือ

$$p\left(\left\{\underline{\tilde{y}}_{ijk}\right\}|\underline{\Lambda}_{kj}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}\right) \propto \prod_{i=1}^{n_j} p\left(\underline{\tilde{y}}_{ijk}|\underline{\Lambda}_{kj}, \underline{\omega}_{-ijk}, \psi_{\epsilon kj}\right)$$

$$= \prod_{i=1}^{n_j} exp\left\{-\frac{1}{2\psi_{\epsilon kj}}\left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T\underline{\Lambda}_{kj}\right)^T\left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T\underline{\Lambda}_{kj}\right)\right\}$$

$$= exp\left\{-\frac{1}{2\psi_{\epsilon kj}}\sum_{i=1}^{n_j}\left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T\underline{\Lambda}_{kj}\right)^T\left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T\underline{\Lambda}_{kj}\right)\right\} \quad (4.19)$$

แทนสมการที่ (4.18) และ (4.19) ลงในสมการที่ (4.17) จากนั้นทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออก จะได้ $p\left(\underline{\Lambda}_{kj}|\left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}, \underline{g}_{kj}, H_{kj}\right)$

$$\propto exp\left\{-\frac{1}{2}\left[\underline{\Lambda}_{kj}^T\left(H_{kj}^{-1} + \psi_{\epsilon kj}^{-1}\sum_{i=1}^{n_j}\underline{\omega}_{-ijk}\underline{\omega}_{-ijk}^T\right)\underline{\Lambda}_{kj}\right] - 2\underline{\Lambda}_{kj}^T\left(H_{kj}^{-1}\underline{g}_{kj} + \psi_{\epsilon kj}^{-1}\sum_{i=1}^{n_j}\underline{\omega}_{-ijk}\underline{\tilde{y}}_{ijk}\right)\right\} \quad (4.20)$$

เมื่อเปรียบเทียบสมการในข้างต้นกับรูปแบบการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรจะได้ว่า สำหรับตัวแปรสังเกตได้ที่ k ในหน่วยที่ j

$$\underline{\Lambda}_{kj}|\left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj} \sim Normal(\underline{\hat{g}}_{kj}, \hat{H}_{kj}) \quad (4.21)$$

โดยที่ $\hat{H}_{kj} = \left[H_{kj}^{-1} + \psi_{\epsilon kj}^{-1}\sum_{i=1}^{n_j}\underline{\omega}_{-ijk}\underline{\omega}_{-ijk}^T\right]^{-1}$

และ $\underline{\hat{g}}_{kj} = \hat{H}_{kj}\left[H_{kj}^{-1}\underline{g}_{kj} + \psi_{\epsilon kj}^{-1}\sum_{i=1}^{n_j}\underline{\omega}_{-ijk}\underline{\tilde{y}}_{ijk}\right]$

3. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรแฝง $\underline{\omega}_{ij}$ เมื่อกำหนดพารามิเตอร์ $\underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij}$*

จากทฤษฎีของเบส์จะได้การแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของตัวแปรแฝง $\underline{\omega}_{ij}$ เมื่อกำหนดพารามิเตอร์ $\underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij}$ คือ

$$p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij}\right) \propto p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}\right) \times p\left(\underline{y}_{ij} | \underline{\omega}_{ij}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}\right) \qquad (4.22)$$

การแจกแจงความน่าจะเป็นก่อนหน้าของตัวแปรแฝง $\underline{\omega}_{ij}$ มีการแจกแจงความน่าจะเป็นแบบปกติ คือ

$$p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}\right) \propto exp\left\{-\tfrac{1}{2}\left(\underline{\omega}_{ij} - \underline{\theta}_j\right)^T \Sigma_{\omega j}^{-1}\left(\underline{\omega}_{ij} - \underline{\theta}_j\right)\right\} \qquad (4.23)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\underline{y}_{ij}$ เมื่อกำหนด $\underline{\omega}_{ij}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}$ คือ

$$p\left(\underline{y}_{ij} | \underline{\omega}_{ij}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}\right) \propto exp\left\{-\tfrac{1}{2}\left(\underline{y}_{ij} - \underline{\mu} - \Lambda_j \underline{\omega}_{ij}\right)^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu} - \Lambda_j \underline{\omega}_{ij}\right)\right\} \qquad (4.24)$$

แทนค่าสมการที่ (4.23) และ (4.24) ลงในสมการที่ (4.22) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออกจากการพิจารณาจะได้ว่า $p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij}\right)$

$$\propto exp\left\{-\tfrac{1}{2}\left[\underline{\omega}_{ij}^T\left(\Sigma_{\omega j}^{-1} + \Lambda_j^T \Psi_{\epsilon j}^{-1} \Lambda_j\right)\underline{\omega}_{ij} - 2\underline{\omega}_{ij}^T\left(\Sigma_{\omega j}^{-1}\underline{\theta}_j + \Lambda_j^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}\right)\right)\right]\right\} \qquad (4.25)$$

เปรียบเทียบสมการที่ (4.25) กับรูปแบบการแจกแจงความน่าจะเป็นแบบปกติจะได้ว่า สำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j

$$\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}\ , \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij} \sim Normal\left(\underline{\hat{\theta}}_{ij}, \hat{\Sigma}_{\omega j}\right) \qquad (4.26)$$

โดยที่ $\hat{\Sigma}_{\omega j} = \left[\Sigma_{\omega j}^{-1} + \Lambda_j^T \Psi_{\epsilon j}^{-1} \Lambda_j\right]^{-1}$

และ $\underline{\hat{\theta}}_{ij} = \hat{\Sigma}_{\omega j}\left[\Sigma_{\omega j}^{-1}\underline{\theta}_j + \Lambda_j^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}\ \right)\right]$

4. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\nu}_j$ เมื่อกำหนด $\left\{\underline{\xi}_{ij}\right\}, \Phi, \Delta$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ค่าเฉลี่ยตัวแปรแฝง $\underline{\nu}_j$ เมื่อกำหนด $\left\{\underline{\xi}_{ij}\right\}, \Phi, \Delta$ คือ

$$p\left(\underline{\nu}_j | \left\{\underline{\xi}_{ij}\right\}, \Phi, \Delta\right) \propto p\left(\underline{\nu}_j | \Delta\right) \times p\left(\left\{\underline{\xi}_{ij}\right\} | \underline{\nu}_j, \Phi\right) \qquad (4.27)$$

การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\underline{\nu}_j$ มีการแจกแจงแบบปกติ คือ

$$p\left(\underline{\nu}_j | \Delta\right) \propto exp\left\{-\tfrac{1}{2}\underline{\nu}_j^T \Delta^{-1} \underline{\nu}_j\right\} \qquad (4.28)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\left\{\underline{\xi}_{ij}\right\}$ เมื่อกำหนดพารามิเตอร์ $\underline{\nu}_j, \Phi$ คือ

$$p\left(\left\{\underline{\xi}_{ij}\right\}|\underline{\nu}_j,\Phi\right) \propto \prod_{i=1}^{n_j} exp\left\{-\frac{1}{2}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T \Phi^{-1}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\right\} \quad (4.29)$$

แทนสมการที่ (4.28) และ (4.29) ลงในสมการที่ (4.27) ทำการกระจายพจน์กำลังสองสมบูรณ์ในสมการที่ (4.29) และตัดพจน์ที่เป็นค่าคงที่ออกจะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\nu}_j$ เมื่อกำหนด $\left\{\underline{\xi}_{ij}\right\},\Phi,\Delta$ คือ

$$p\left(\underline{\nu}_j|\left\{\underline{\xi}_{ij}\right\},\Phi,\Delta\right) \propto \exp\left\{-\frac{1}{2}\left[\underline{\nu}_j^T\left(\Delta^{-1}+n_j\Phi^{-1}\right)\underline{\nu}_j - 2\underline{\nu}_j^T\Phi^{-1}\sum_{i=1}^{n_j}\underline{\xi}_{ij}\right]\right\} \quad (4.30)$$

เมื่อเปรียบเทียบสมการที่ (4.30) กับรูปแบบของการแจกแจงแบบปกติหลายตัวแปรจะได้ว่าสำหรับหน่วยที่ j

$$\underline{\nu}_j|\left\{\underline{\xi}_{ij}\right\},\Phi,\Delta \sim Normal(\hat{\underline{\nu}}_j,\hat{\Delta}) \quad (4.31)$$

โดยที่ $\hat{\Delta} = \left[\Delta^{-1}+n_j\Phi^{-1}\right]^{-1}$ และ $\hat{\underline{\nu}}_j = \hat{\Delta}\left[\Phi^{-1}\sum_{i=1}^{n_j}\underline{\xi}_{ij}\right]$

5. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์* Φ^{-1} *เมื่อกำหนดพารามิเตอร์* $\underline{\nu}_j,\left\{\underline{\xi}_{ij}\right\},\rho_\phi,R_\phi$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ Φ^{-1} เมื่อกำหนด $\underline{\nu}_j,\left\{\underline{\xi}_{ij}\right\},\rho_\phi,R_\phi$ คือ

$$p\left(\Phi^{-1}|\left\{\underline{\xi}_{ij}\right\},\underline{\nu}_j,\rho_\phi,R_\phi\right) \propto p\left(\Phi^{-1}|\rho_\phi,R_\phi\right)\times p\left(\left\{\underline{\xi}_{ij}\right\}|\underline{\nu}_j,\Phi\right) \quad (4.32)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ Φ^{-1} คือ

$$p\left(\Phi^{-1}|\rho_\phi,R_\phi\right) \propto |\Phi|^{-\frac{\rho_\phi-q_2-1}{2}}\exp\left\{-\frac{1}{2}tr\left(R_\phi^{-1}\Phi^{-1}\right)\right\} \quad (4.33)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\left\{\underline{\xi}_{ij}\right\}$ เมื่อกำหนด $\underline{\nu}_j,\Phi$ คือ

$$p\left(\left\{\underline{\xi}_{ij}\right\}|\underline{\nu}_j,\Phi\right) \propto \prod_{j=1}^{J}\prod_{i=1}^{n_j}|\Phi|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T\Phi^{-1}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\right\}$$

$$= |\Phi|^{-\frac{\sum_{j=1}^{J}n_j}{2}} exp\left\{-\frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T\Phi^{-1}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\right\} \quad (4.34)$$

แทนสมการที่ (4.33) และ (4.34) ลงในสมการที่ (4.32) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ Φ^{-1} เมื่อกำหนด $\{\underline{\xi}_{ij}\}, \underline{\nu}_j, \rho_\phi, R_\phi$ คือ

$$p\left(\Phi^{-1}|\{\underline{\xi}_{ij}\}, \underline{\nu}_j, \rho_\phi, R_\phi\right) \propto |\Phi|^{-\frac{\sum_{j=1}^{J} n_j+\rho_\phi-q_2-1}{2}} exp\left\{-\frac{1}{2}\left[tr(R_\phi^{-1}\Phi^{-1}) + \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T \Phi^{-1}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\right]\right\}$$

$$= |\Phi^{-1}|^{\frac{\sum_{j=1}^{J} n_j+\rho_\phi-q_2-1}{2}} exp\left\{-\frac{1}{2}\left[tr(\Phi^{-1}R_\phi^{-1}) + tr\left(\Phi^{-1}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T\right)\right]\right\}$$

$$= |\Phi^{-1}|^{\frac{\sum_{j=1}^{J} n_j+\rho_\phi-q_2-1}{2}} \exp\left\{-\frac{1}{2}tr\left(\Phi^{-1}R_\phi^{-1} + \Phi^{-1}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T\right)\right\}$$

$$= |\Phi^{-1}|^{\frac{\sum_{j=1}^{J} n_j+\rho_\phi-q_2-1}{2}} exp\left\{-\frac{1}{2}tr\left(\Phi^{-1}\left[R_\phi^{-1} + \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T\right]\right)\right\} \qquad (4.35)$$

เปรียบเทียบสมการที่ (4.35) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบ Wishart จะได้ว่าสำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j

$$\Phi^{-1}|\{\underline{\xi}_{ij}\}, \underline{\nu}_j, \rho_\phi, R_\phi \sim Wishart\left(\hat{\rho}_\phi, \hat{R}_\phi\right) \qquad (4.36)$$

โดยที่ $\hat{\rho}_\phi = n + \rho_\phi$ และ $\hat{R}_\phi = \left[R_\phi^{-1} + \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)\left(\underline{\xi}_{ij}-\underline{\nu}_j\right)^T\right]^{-1}$

6. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\psi_{\epsilon kj}^{-1}$ เมื่อกำหนด $\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \{\underline{y}_{ij}\}, \alpha_{kj}, \beta_{kj}$

เนื่องจากเมทริกซ์ $\Psi_{\epsilon j}$ เป็นเมทริกซ์ความแปรปรวนร่วมของความคลาดเคลื่อนจากการวัด ซึ่งเป็นเมทริกซ์ทแยงมุม (diagonal matrix) โดยที่สมมติให้สมาชิกในแนวทแยงมุม $\psi_{\epsilon kj}^{-1}$ และ $\psi_{\epsilon k'j}^{-1}$ เป็นอิสระซึ่งกันและกัน ดังนั้นการสุ่มตัวอย่างเมทริกซ์ $\Psi_{\epsilon j}$ จึงสามารถสุ่มตัวอย่างพารามิเตอร์ทแยงมุม $\psi_{\epsilon kj}^{-1}$ ทีละตัว จากทฤษฎีของเบส์จะได้ว่า การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\psi_{\epsilon kj}^{-1}$ เมื่อกำหนดพารามิเตอร์ $\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \{\underline{y}_{ij}\}, \alpha_{kj}, \beta_{kj}$ คือ

$$p\left(\psi_{\epsilon kj}^{-1}|\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \{\underline{y}_{ij}\}, \alpha_{kj}, \beta_{kj}\right) \propto p(\psi_{\epsilon kj}^{-1}|\alpha_{kj}, \beta_{kj}) \times p\left(\{\underline{y}_{ij}\}|\mu_k, \underline{\Lambda}_k, \psi_{\epsilon kj}^{-1}, \{\underline{\omega}_{ij}\}\right) \quad (4.37)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\psi_{\epsilon kj}^{-1}$ คือ

$$p(\psi_{\epsilon kj}^{-1}|\alpha_{kj}, \beta_{kj}) \propto (\psi_{\epsilon kj}^{-1})^{\alpha_{kj}-1} exp\{-\beta_{kj}\psi_{\epsilon kj}^{-1}\} \qquad (4.38)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\{\underline{y}_{ij}\}$ เมื่อกำหนด $\mu_k, \underline{\Lambda}_k, \psi_{\epsilon kj}^{-1}, \{\underline{\omega}_{ij}\}$ คือ

$$p\left(\{\underline{y}_{ij}\}|\mu_k, \underline{\Lambda}_k, \psi_{\epsilon kj}^{-1}, \{\underline{\omega}_{ij}\}\right) \propto$$
$$\prod_{i=1}^{n_j}(\psi_{\epsilon kj}^{-1})^{-\frac{1}{2}} exp\left\{-\frac{1}{2\psi_{\epsilon kj}}\left(\underline{y}_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)^T\left(\underline{y}_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)\right\}$$

$$= \left(\psi_{\epsilon kj}^{-1}\right)^{-\frac{n_j}{2}} exp\left\{-\frac{1}{2\psi_{\epsilon kj}}\sum_{i=1}^{n_j}\left(\underline{y}_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)^T\left(\underline{y}_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)\right\} \quad (4.39)$$

แทนสมการที่ (4.38) และ (4.39) ลงในสมการที่ (4.37) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\psi_{\epsilon kj}^{-1}$ เมื่อกำหนด $\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \{\underline{y}_{ij}\}, \alpha_{kj}, \beta_{kj}$ คือ

$$p\left(\psi_{\epsilon kj}^{-1}|\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \{\underline{y}_{ij}\}, \alpha_{kj}, \beta_{kj}\right) \propto \left(\psi_{\epsilon kj}^{-1}\right)^{\frac{n_j}{2}+\alpha_{kj}-1} exp\left\{-\psi_{\epsilon kj}^{-1}\left[\beta_{kj} + \frac{1}{2}\sum_{i=1}^{n_j}\left(\underline{y}_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)^T\left(\underline{y}_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)\right]\right\} \quad (4.40)$$

เมื่อเปรียบเทียบสมการที่ (4.40) กับการแจกแจงความน่าจะเป็นแบบแกมมาจะได้ว่าสำหรับตัวแปรสังเกตได้ที่ k ในหน่วยที่ j

$$\psi_{\epsilon kj}^{-1}|\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \{\underline{y}_{ij}\}, \alpha_{kj}, \beta_{kj} \sim Gamma\left(\hat{\alpha}_{\epsilon kj}, \hat{\beta}_{\epsilon kj}\right) \quad (4.41)$$

โดยที่ $\hat{\alpha}_{\epsilon kj} = \frac{n_j}{2} + \alpha_{kj}$

และ $\hat{\beta}_{\epsilon kj} = \beta_{kj} + \frac{1}{2}\sum_{i=1}^{n_j}\left(y_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)^T\left(y_{ij}-\mu_k-\underline{\Lambda}_k^T\underline{\omega}_{ij}\right)$

7. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\gamma}$ เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\gamma}$ เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u, \underline{\gamma_0}, \Sigma_{\gamma 0}$ คือ

$$p\left(\underline{\gamma}|\{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u\right) \propto p\left(\underline{\gamma}|\underline{\gamma_0}, \Sigma_{\gamma 0}\right) \times p\left(\{\underline{\omega}_{ij}\}|\underline{\gamma}, \underline{u}_j, \Sigma_u, \sigma_\delta^2\right) \quad (4.42)$$

โดยที่การแจกแจงความน่าจะเป็นของพารามิเตอร์จุดตัดแบบสุ่ม β_{0j} คือ

$$p\left(\underline{\gamma}|\underline{\gamma_0}, \Sigma_{\gamma 0}\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\gamma}-\underline{\gamma_0}\right)^T \Sigma_{\gamma 0}^{-1}\left(\underline{\gamma}-\underline{\gamma_0}\right)\right\} \quad (4.43)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\{\underline{\omega}_{ij}\}$ เมื่อกำหนด $\underline{\gamma}, \underline{u}_j, \Sigma_u, \sigma_\delta^2$

$$p\left(\{\underline{\omega}_{ij}\}|\beta_{0j}, \underline{\beta}_j, \sigma_\delta^2\right) \propto \prod_{j=1}^{J}\prod_{i=1}^{n_j} exp\left\{-\frac{1}{2\sigma_\delta^2}\left(\eta_{ij}-\Xi_{ij}^T\underline{\gamma}-\Xi_{ij}^T\underline{u}_j\right)^T\left(\eta_{ij}-\Xi_{ij}^T\underline{\gamma}-\Xi_{ij}^T\underline{u}_j\right)\right\}$$

$$= exp\left\{-\frac{1}{2\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij}-\Xi_{ij}^T\underline{\gamma}-\Xi_{ij}^T\underline{u}_j\right)^2\right\} \quad (4.44)$$

เมื่อ $\Xi = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \xi_{11} & \xi_{21} & \cdots & \xi_{n_1 1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{1J} & \xi_{2J} & \cdots & \xi_{n_J J} \end{bmatrix}^T$

แทนสมการที่ (4.43) และ (4.44) ลงในสมการที่ (4.42) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออก จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ β_{0j} เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{\beta}_j, \gamma_{00}, \sigma_\delta^2, \tau_{00}$ คือ

$$p\left(\underline{\gamma}|\{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\gamma}-\underline{\gamma_0}\right)^T \Sigma_{\gamma 0}^{-1}\left(\underline{\gamma}-\underline{\gamma_0}\right)\right\}$$

$$\times = exp\left\{-\frac{1}{2\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij}-\underline{\Xi}_{ij}^T\underline{\gamma}-\underline{\Xi}_{ij}^T\underline{u}_j\right)^2\right\}$$

$$= exp\left\{-\frac{1}{2}\left[\underline{\gamma}^T\left(\Sigma_{\gamma 0}^{-1}+\Xi\Xi^{\mathrm{T}}\right)\underline{\gamma}-2\underline{\gamma}^T\left[\Sigma_{\gamma 0}^{-1}\underline{\gamma}+\Xi^{\mathrm{T}}(\eta-\Xi u^*)\right]\right]\right\} \quad (4.45)$$

เปรียบเทียบสมการ (4.45) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบปกติจะได้ว่าสำหรับหน่วยที่ j

$$\underline{\gamma}|\{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{\gamma}}\ \ , \hat{\Sigma}_\gamma\right) \quad (4.46)$$

โดยที่ $\hat{\Sigma}_\gamma = \left[\Sigma_u^{-1}+\frac{1}{\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\underline{\Xi}_{ij}\underline{\Xi}_{ij}^T\right]^{-1}$

และ$\underline{\hat{\gamma}}\ \ = \hat{\Sigma}_\gamma\left[\Sigma_u^{-1}\underline{\gamma}+\frac{1}{\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\underline{\Xi}_{ij}(\eta_{ij}-\underline{\Xi}_{ij}^T\underline{u}_j)\right]$ เมื่อ $\underline{\Xi}_{ij} = \begin{pmatrix}1 & \underline{\xi}_{ij}\end{pmatrix}^T$

8. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ u_j เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{u}_j$ เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u$ คือ

$$p\left(\underline{u}_j|\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u\right) \propto p\left(\underline{u}_j|\Sigma_u\right) \times p\left(\{\underline{\omega}_{ij}\}|\underline{\gamma}, \sigma_\delta^2\right) \quad (4.47)$$

โดยที่การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม $\underline{u}_j$ คือ

$$p\left(\underline{u}_j|\Sigma_u\right) \propto exp\left\{-\frac{1}{2}\underline{u}_j^T\Sigma_u^{-1}\underline{u}_j\right\} \quad (4.48)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\{\underline{\omega}_{ij}\}$ เมื่อกำหนด $\underline{\beta}_j, \sigma_\delta^2$ คือ

$$p\left(\{\underline{\omega}_{ij}\}|\underline{\gamma}, \sigma_\delta^2\right) \propto \prod_{i=1}^{n_j} exp\left\{-\frac{1}{2\sigma_\delta^2}\left(\eta_{ij}-\underline{\Xi}_{ij}^T\underline{\gamma}-\underline{\Xi}_{ij}^T\underline{u}_j\right)^T\left(\eta_{ij}-\underline{\Xi}_{ij}^T\underline{\gamma}-\underline{\Xi}_{ij}^T\underline{u}_j\right)\right\}$$

$$= exp\left\{-\frac{1}{2\sigma_\delta^2}\sum_{i=1}^{n_j}\left(\eta_{ij}-\underline{\Xi}_{ij}^T\underline{\gamma}-\underline{\Xi}_{ij}^T\underline{u}_j\right)^T\left(\eta_{ij}-\underline{\Xi}_{ij}^T\underline{\gamma}-\underline{\Xi}_{ij}^T\underline{u}_j\right)\right\} \quad (4.49)$$

แทนสมการที่ (4.48) และ (4.49) ลงในสมการที่ (4.47) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออกจะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของตัวแปรสุ่ม $\underline{u}_j$ เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u$ คือ

$$p\left(\underline{u}_j | \{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u\right)$$

$$\propto exp\left\{-\frac{1}{2}\left[\underline{u}_j^T\left(\Sigma_u^{-1} + \frac{1}{\sigma_\delta^2}\sum_{i=1}^{n_j} \Xi_{ij}\,\underline{\Xi}_{ij}^T\right)\underline{u}_j - 2\underline{u}_j^T\left(\sum_{i=1}^{n_j} \underline{\Xi}_{ij}\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma}\right)\right)\right]\right\} \quad (4.50)$$

เปรียบเทียบสมการ (4.50) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรจะได้ว่าสำหรับหน่วยที่ j

$$\underline{u}_j | \{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{u}}_j, \hat{\Sigma}_{uj}\right) \quad (4.51)$$

โดยที่ $\hat{\Sigma}_{uj} = \left[\frac{1}{\sigma_\delta^2}\sum_{i=1}^{n_j} \underline{\Xi}_{ij}\underline{\Xi}_{ij}^T + \Sigma_u^{-1}\right]^{-1}$

และ $\underline{\hat{u}}_j = \frac{1}{\sigma_\delta^2}\hat{\Sigma}_{uj}\left[\sum_{i=1}^{n_j} \underline{\Xi}_{ij}^T\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma}\right)\right]$

9. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์* Σ_u^{-1} *เมื่อกำหนด* $\{\underline{u}_j\}, \rho_u, R_u$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ Σ_u^{-1} เมื่อกำหนด $\underline{u}_j, \rho_u, R_u$ คือ

$$p\left(\Sigma_u | \{\underline{u}_j\}, \rho_u, R_u\right) \propto p\left(\Sigma_u^{-1} | \rho_u, R_u\right) \times p\left(\{\underline{u}_j\} | \Sigma_u\right) \quad (4.52)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ Σ_u^{-1} มีการแจกแจงความน่าจะเป็นแบบ Wishart ที่มีพารามิเตอร์ ρ_u, R_u กล่าวคือ

$$p\left(\Sigma_u^{-1} | \rho_u, R_u\right) \propto |\Sigma_u^{-1}|^{\frac{\rho_u - q_2 - 1}{2}} exp\left\{-\frac{1}{2}tr\left(R_u^{-1}\Sigma_u^{-1}\right)\right\} \quad (4.53)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\underline{u}_j$ เมื่อกำหนด Σ_u คือ

$$p\left(\{\underline{u}_j\} | \Sigma_u\right) \propto \prod_{j=1}^{J} p\left(\underline{u}_j | \Sigma_u\right) \quad (4.54)$$

$$= \prod_{j=1}^{J} |\Sigma_u|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\underline{u}_j^T \Sigma_u^{-1} \underline{u}_j\right\}$$

$$= |\Sigma_u|^{-\frac{J}{2}} exp\left\{-\frac{1}{2}\sum_{j=1}^{J}\underline{u}_j^T\Sigma_u^{-1}\underline{u}_j\right\} \tag{4.55}$$

แทนสมการที่ (4.55) และ (4.57) ลงในสมการที่ (4.54) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ Σ_u เมื่อกำหนด $\underline{u}_j, \rho_u, R_u$ คือ

$$p\left(\Sigma_u|\{\underline{u}_j\}, \rho_u, R_u\right) \propto |\Sigma_u^{-1}|^{\frac{J+\rho_u-q_2-1}{2}} \exp\left\{-\frac{1}{2}tr\left(\Sigma_u^{-1}\left[R_u^{-1}+\sum_{j=1}^{J}\underline{u}_j^T\underline{u}_j\right]\right)\right\} \tag{4.56}$$

เปรียบเทียบสมการที่ (4.58) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบ Wishart จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ Σ_u^{-1} เมื่อกำหนด $\{\underline{u}_j\}$, $\hat{\rho}_u, \hat{R}_u$ คือ

$$\Sigma_u^{-1}|\{\underline{u}_j\}, \rho_u, R_u \sim Wishart\left(\hat{\rho}_u, \hat{R}_u\right) \tag{4.57}$$

โดยที่ $\hat{\rho}_u = J + \rho_u$ และ $\hat{R}_u = \left[R_u^{-1} + \sum_{j=1}^{J}\underline{u}_j^T\underline{u}_j\right]^{-1}$

10. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ σ_δ^{-2} *เมื่อกำหนด* $\{\underline{\omega}_{ij}\}, \left\{\underline{\beta}_j\right\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ σ_δ^{-2} เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \left\{\underline{\beta}_j\right\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta$ คือ

$$p\left(\sigma_\delta^{-2}|\{\underline{\omega}_{ij}\}, \left\{\underline{\beta}_j\right\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta\right) \propto p(\sigma_\delta^{-2}|\alpha_\delta, \beta_\delta) \times p\left(\{\underline{\omega}_{ij}\}|\sigma_\delta^{-2}, \left\{\underline{\beta}_j\right\}, \{\underline{u}_j\}\right) \tag{4.58}$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ σ_δ^{-2} เมื่อกำหนดพารามิเตอร์ $\alpha_\delta, \beta_\delta$ คือ

$$p(\sigma_\delta^{-2}|\alpha_\delta, \beta_\delta) \propto (\sigma_\delta^{-2})^{\alpha_\delta - 1} exp\{-\beta_\delta\sigma_\delta^{-2}\} \tag{4.59}$$

และฟังก์ชันภาวะความควรจะเป็นของ $\{\underline{\omega}_{ij}\}$ เมื่อกำหนด $\sigma_\delta^{-2}, \left\{\underline{\beta}_j\right\}, \{\underline{u}_j\}$ คือ

$$p\left(\{\underline{\omega}_{ij}\}|\sigma_\delta^{-2}, \left\{\underline{\beta}_j\right\}, \{\underline{u}_j\}\right) \propto \prod_{j=1}^{J}\prod_{i=1}^{n_j} p\left(\underline{\omega}_{ij}|\sigma_\delta^{-2}, \underline{\beta}_j, \underline{u}_j\right)$$

$$= \prod_{j=1}^{J}\prod_{i=1}^{n_j}(\sigma_\delta^{-2})^{-\frac{1}{2}} exp\left\{-\frac{1}{2\sigma_\delta^2}\left(\eta_{ij} - \Xi_{ij}^T\underline{\gamma} - \Xi_{ij}^T\underline{u_j}\right)^T\left(\eta_{ij} - \Xi_{ij}^T\underline{\gamma} - \Xi_{ij}^T\underline{u_j}\right)\right\}$$

$$= (\sigma_\delta^2)^{-\frac{\sum_{j=1}^{J}n_j}{2}} exp\left\{-\frac{1}{2\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \Xi_{ij}^T\underline{\gamma} - \Xi_{ij}^T\underline{u_j}\right)^T\left(\eta_{ij} - \Xi_{ij}^T\underline{\gamma} - \Xi_{ij}^T\underline{u_j}\right)\right\} \tag{4.60}$$

แทนสมการที่ (4.61) และ (4.62) ลงในสมการ (4.60) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ σ_δ^{-2} เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \{\underline{\beta_j}\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta$ คือ

$$p\left(\sigma_\delta^{-2} | \{\underline{\omega}_{ij}\}, \{\underline{\beta_j}\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta\right) \propto \left(\sigma_\delta^{-2}\right)^{\frac{\sum_{j=1}^{J} n_j}{2} + \alpha_\delta - 1} exp\left\{-\sigma_\delta^{-2}\left[\beta_\delta + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma} - \underline{\Xi}_{ij}^T\underline{u_j}\right)^T\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma} - \underline{\Xi}_{ij}^T\underline{u_j}\right)\right]\right\} \quad (4.61)$$

เปรียบเทียบสมการที่ (4.61) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบแกมมา จะได้ว่า

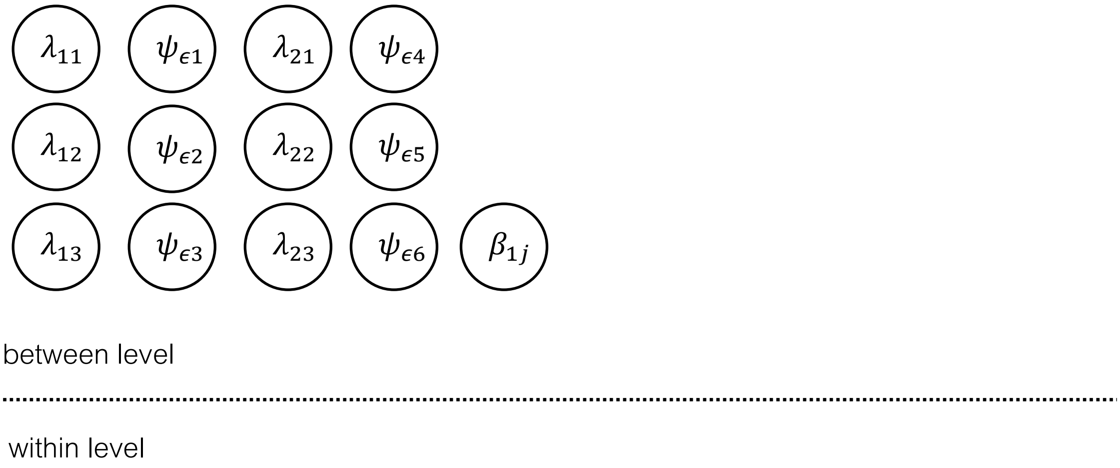
$$\sigma_\delta^{-2} | \left\{\underline{\omega}_{ij}\right\}, \left\{\underline{\beta}_j\right\}, \left\{\underline{u}_j\right\}, \alpha_\delta, \beta_\delta \sim Gamma(\widehat{\alpha}_\delta, \widehat{\beta}_\delta) \quad (4.62)$$

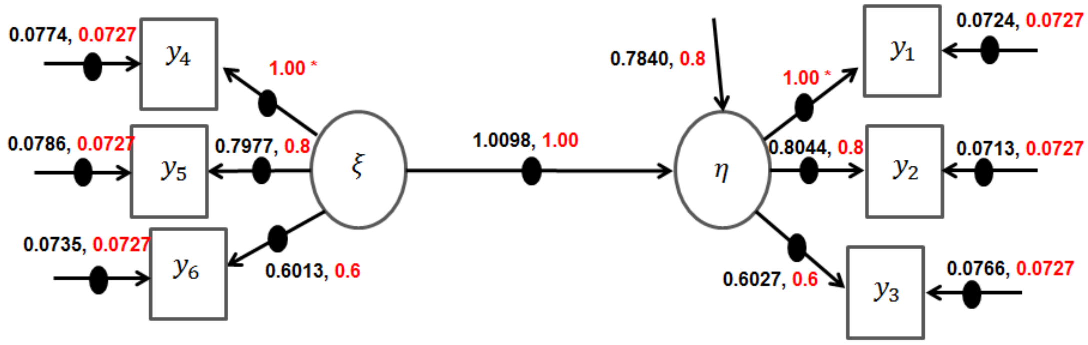
โดยที่ $\hat{\alpha}_\delta = \frac{\sum_{j=1}^{J} n_j}{2} + \alpha_\delta$ และ $\hat{\beta}_\delta = \beta_\delta + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma} - \underline{\Xi}_{ij}^T\underline{u_j}\right)^2$

1.4 การตรวจสอบความถูกต้องของอัลกอริทึมที่ได้พัฒนาขึ้น

ผู้วิจัยได้ทำการตรวจสอบความถูกต้องของค่าประมาณพารามิเตอร์ที่ได้จากอัลกอริทึมที่พัฒนาขึ้นโดยใช้วิธีการเชิงจำลองแบบมอนติคาร์โล ทำการพิจารณาค่าประมาณพารามิเตอร์ที่ได้จาก 2 สถานการณ์ได้แก่ สถานการณ์ที่มีตัวแปรตามแฝงและตัวแปรอิสระจำนวน 1 ตัวแปร และสถานการณ์ที่มีตัวแปรตามแฝงจำนวน 1 ตัวและตัวแปรอิสระแฝงจำนวน 3 ตัวแปร ในแต่ละสถานการณ์กระทำซ้ำจำนวน 100 ครั้ง ผลการประมาณค่าพารามิเตอร์ที่ได้จาก 100 ครั้งจะสรุปเป็นค่าเฉลี่ยและนำมาเปรียบเทียบกับค่าจริงของพารามิเตอร์ที่ใช้ในการสร้างข้อมูลจำลอง

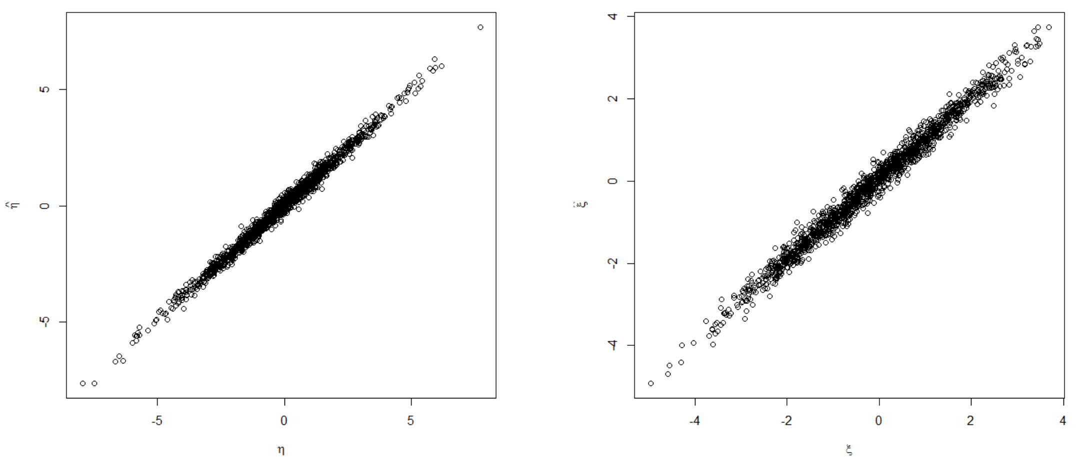
ผลการประมาณค่าพารามิเตอร์จากรูปที่ 4.1 พบว่าในสถานการณ์จำลองที่มีตัวแปรตามแฝงจำนวน 1 ตัวแปรและมีตัวแปรอิสระแฝงจำนวน 1 ตัวแปร วิธีการประมาณค่าพารามิเตอร์ที่พัฒนาขึ้นให้ค่าประมาณพารามิเตอร์ (สีดำ) ใกล้เคียงกับค่าจริงของพารามิเตอร์ (สีแดง) ซึ่งถือเป็นการยืนยันได้ว่าอัลกอริทึมที่พัฒนาขึ้นมีความถูกต้องและน่าเชื่อถือ





รูปที่ 4.1 เปรียบเทียบค่าประมาณพารามิเตอร์กับค่าจริงของพารามิเตอร์ในกรณีที่มีตัวแปรตามแฝงและตัวแปรอิสระแฝงจำนวน 1 ตัวแปร

รูปที่ 4.2 ก และ ข แสดงแผนภาพการกระจายของระหว่างค่าประมาณตัวแปรแฝงและค่าจริงของตัวแปรแฝง โดยรูปที่ 4.2 ก เป็นแผนภาพการกระจายระหว่างค่าประมาณตัวแปรตามแฝงกับค่าจริงของตัวแปรตามแฝง ส่วนรูปที่ 4.2 ข เป็นแผนภาพการกระจายระหว่างค่าประมาณตัวแปรอิสระแฝงกับค่าจริงของตัวแปรอิสระแฝง ซึ่งจากแผนภาพแสดงให้เห็นว่าค่าประมาณตัวแปรตามแฝงและตัวแปรอิสระแฝงมีความใกล้เคียงกับค่าจริงอยู่ในระดับที่สูงมาก

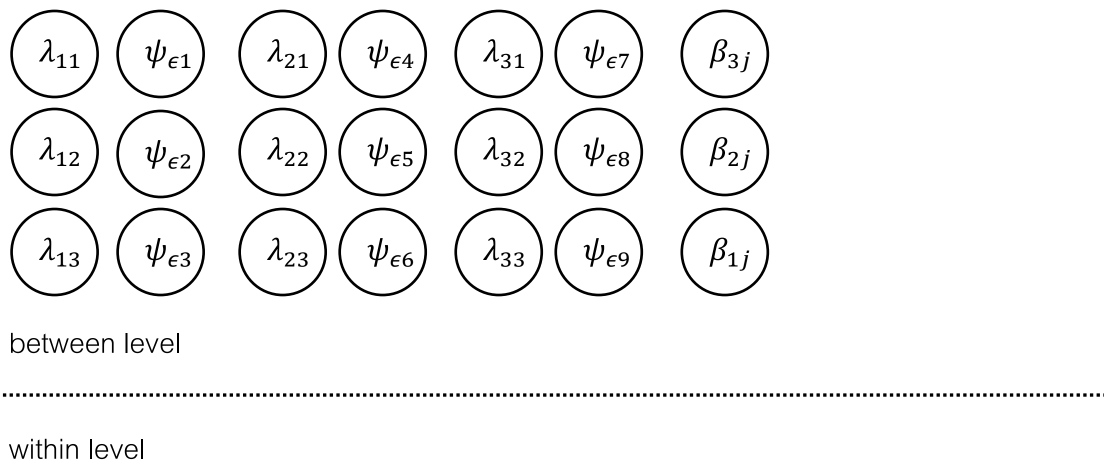


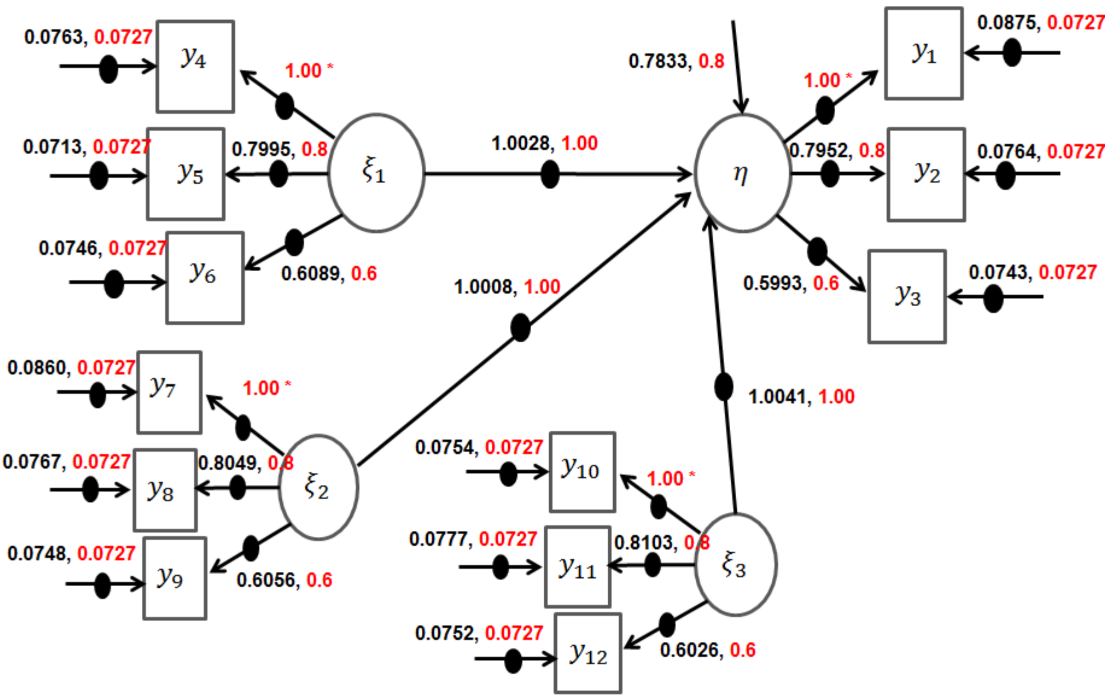
ก. ข.

รูปที่ 4.2 แผนภาพการกระจายของระหว่างค่าประมาณตัวแปรแฝงและค่าจริงของตัวแปรแฝงใน

สถานการณ์จำลองที่มีตัวแปรตามแฝงและตัวแปรอิสระแฝงเท่ากับ 1 ตัวแปร

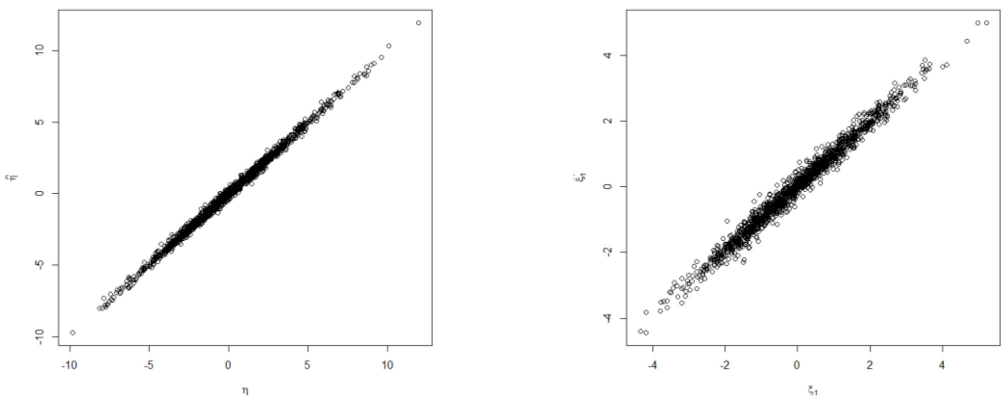
ในกรณีที่มีตัวแปรอิสระแฝงมากกว่า 1 ตัวแปร ผู้วิจัยได้ลองทำการยืนยันความถูกต้องในกรณีที่มีตัวแปรอิสระแฝงเท่ากับ 3 ตัวแปร ผลการประมาณค่าพารามิเตอร์จากรูปที่ 4.3 พบว่าในสถานการณ์จำลองที่มีตัวแปรตามแฝงจำนวน 1 ตัวแปรและมีตัวแปรอิสระแฝงจำนวน 3 ตัวแปร วิธีการประมาณค่าพารามิเตอร์ที่พัฒนาขึ้นให้ค่าประมาณพารามิเตอร์ (สีดำ) ใกล้เคียงกับค่าจริงของพารามิเตอร์ (สีแดง) ซึ่งถือเป็นการยืนยันได้ว่าอัลกอริทึมที่พัฒนาขึ้นสามารถประมาณค่าพารามิเตอร์ในกรณีที่มีตัวแปรอิสระแฝงมากกว่า 1 ตัวแปรได้อย่างมีความถูกต้องและน่าเชื่อถือ



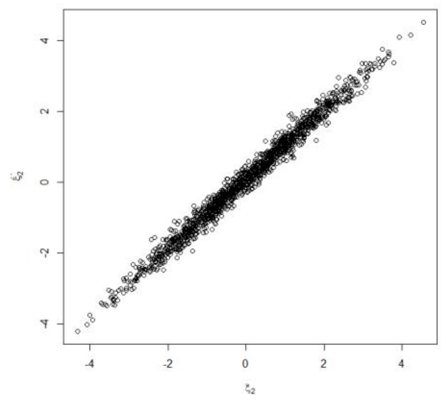


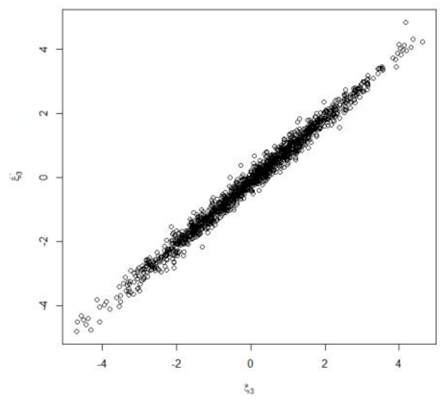
รูปที่ 4.3 เปรียบเทียบค่าประมาณพารามิเตอร์กับค่าจริงของพารามิเตอร์ในกรณีที่มีตัวแปรตามแฝงจำนวน 1 ตัวแปรและตัวแปรอิสระแฝงจำนวน 3 ตัวแปร

รูปที่ 4.4 ก ถึง ง แสดงแผนภาพการกระจายของระหว่างค่าประมาณตัวแปรแฝงและค่าจริงของตัวแปรแฝง โดยรูปที่ 4.2 ก เป็นแผนภาพการกระจายระหว่างค่าประมาณตัวแปรตามแฝงกับค่าจริงของตัวแปรตามแฝง ส่วนรูปที่ 4.2 ข ถึง ง เป็นแผนภาพการกระจายระหว่างค่าประมาณตัวแปรอิสระแฝงกับค่าจริงของตัวแปรอิสระแฝง ซึ่งจากแผนภาพแสดงให้เห็นว่าค่าประมาณตัวแปรตามแฝงและตัวแปรอิสระแฝงมีความใกล้เคียงกับค่าจริงอยู่ในระดับที่สูงมาก



ก. ข.





ค. ง.

รูปที่ 4.4 แผนภาพการกระจายของระหว่างค่าประมาณตัวแปรแฝงและค่าจริงของตัวแปรแฝงในสถานการณ์จำลองที่มีตัวแปรตามแฝงเท่ากับ 1 ตัวแปรและตัวแปรอิสระแฝงเท่ากับ 3 ตัวแปร

ตอนที่ 2: การวิเคราะห์ข้อมูลกรณีการจำลองข้อมูลด้วยวิธีมอนติคาร์โล

การศึกษานี้มีวัตถุประสงค์เพื่อตรวจสอบความสามารถในการประมาณค่าพารามิเตอร์โดยวิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้น และเพื่อเปรียบเทียบความสามารถระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์กับวิธีประมาณค่าแบบภาวะความควรจะเป็นสูงสุด โดยวิธีการประมาณค่าแบบเบส์จะประมาณค่าภายใต้โมเดลการวัดที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่ม ส่วนวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดจะประมาณค่าภายใต้โมเดลการวัดที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์คงที่

ในการศึกษาผู้วิจัยใช้เกณฑ์การพิจารณาค่าความคลาดเคลื่อนกำลังสอง (MSE) ของค่าประมาณพารามิเตอร์ในโมเดล และค่าอัตราส่วนของผลต่างค่าคลาดเคลื่อนกำลังสองระหว่างวิธีการประมาณค่าแบบเบส์กับวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted เทียบกับค่าคลาดเคลื่อนกำลังสองของวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted ($RDMSE$) ซึ่งได้แบ่งออกเป็น 6 เซตได้แก่ เซตของพารามิเตอร์น้ำหนัก

องค์ประกอบ เซตของพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด เซตของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมของตัวแปรแฝง เซตของพารามิเตอร์อิทธิพลคงที่ เซตของพารามิเตอร์ความแปรปรวนในระดับที่หนึ่ง และเซตของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมในระดับที่สอง

สำหรับค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่มที่ได้จากวิธีการประมาณแบบเบส์ ผู้วิจัยจะทำการเฉลี่ยค่าประมาณของทุกกลุ่มเข้าด้วยกันแล้วจึงคำนวณเป็นค่า MSE เพื่อให้สามารถเปรียบเทียบกับค่าประมาณที่ได้จากวิธีภาวะความควรจะเป็นสูงสุดได้ โดยจะนำเสนอในรูปแบบตารางและกราฟ สัญลักษณ์ดังต่อไปนี้ใช้แทนความหมายต่างๆในผลการวิจัย

J แทน จำนวนตัวอย่างในระดับที่ 2

Λ_j แทน พารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่ม

$\Psi_{\varepsilon j}$ แทน พารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่ม

Φ แทน พารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมของตัวแปรอิสระแฝง

γ แทน พารามิเตอร์อิทธิพลคงที่

σ_δ^2 แทน พารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1

Σ_u แทน พารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมของโมเดลระดับที่ 2

η แทน ตัวแปรตามแฝง

ξ แทน ตัวแปรอิสระแฝง

$\overline{\rho^2}$ แทน ค่าเฉลี่ยของสัมประสิทธิ์ความเที่ยงรวม (average reliability)

MSE แทน ค่าความคลาดเคลื่อนกำลังสอง (mean square error)

$\overline{MSE}$ แทน ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (mean square error)

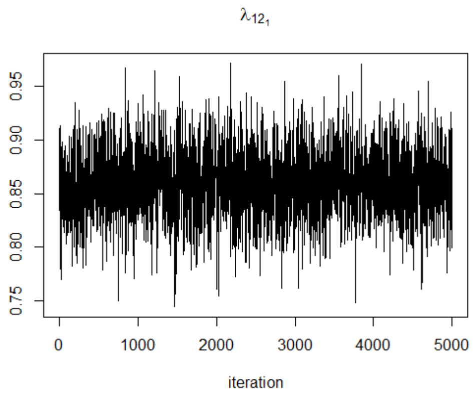
$RAMSE$ แทน ค่าอัตราส่วนค่าคลาดเคลื่อนกำลังสองของวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted เทียบกับวิธีการประมาณค่าแบบเบส์

Bayes แทน วิธีการประมาณค่าพารามิเตอร์แบบเบส์ (ประมาณค่าภายใต้โมเดลการวัดที่มีน้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่ม)

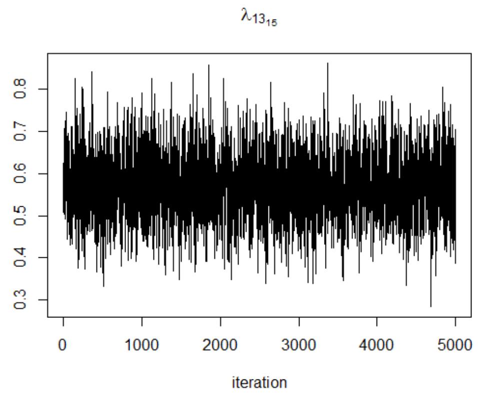
Mplus แทน วิธีการประมาณค่าพารามิเตอร์ภาวะความควรจะเป็นสูงสุด (ประมาณค่าภายใต้โมเดลการวัดที่มีน้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดคงที่)

ค่าประมาณของพารามิเตอร์ในโมเดลคำนวณจากตัวอย่างพารามิเตอร์ที่สุ่มจากการแจกแจงความน่าจะเป็นภายหลังร่วมจำนวน 5000 ตัวอย่างด้วยอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ที่พัฒนาขึ้น ก่อนใช้ตัวอย่างดังกล่าวเพื่อประมาณค่าพารามิเตอร์ในโมเดล ผู้วิจัยทำการตัดตัวอย่างพารามิเตอร์ในส่วนแรก (burn-in) จำนวน 1000 ตัวอย่าง และทำการตรวจสอบความเหมาะสมของตามคุณสมบัติของลูกโซ่มาร์คอฟโดยการใช้ trace plot, marginal posterior density plot, auto-correlation plot, partial-auto correlation plotm, การทดสอบของ Geweke และการทดสอบของ Heidelberger-Welch

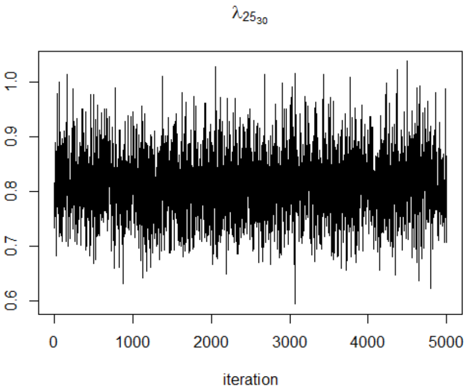
รูปที่ 4.5 แสดงตัวอย่าง trace plot ของตัวอย่างพารามิเตอร์ในโมเดล ซึ่งจะเห็นว่าตัวอย่างของพารามิเตอร์ที่สร้างขึ้นโดยอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์มีลักษณะที่ลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังตามที่ต้องการ



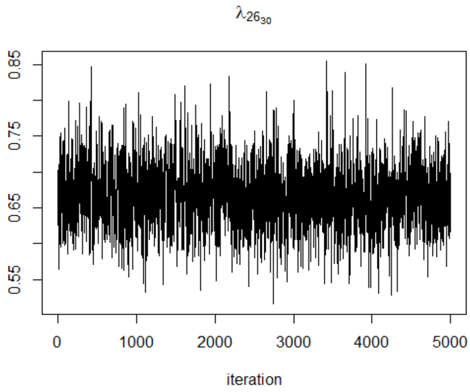
ก. trace plot ของพารามิเตอร์ $\lambda_{12(1)}$



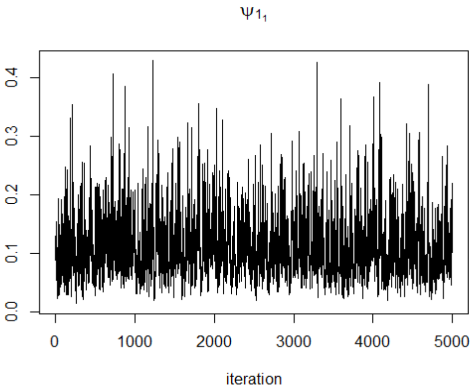
ข. trace plot ของพารามิเตอร์ $\lambda_{13(15)}$



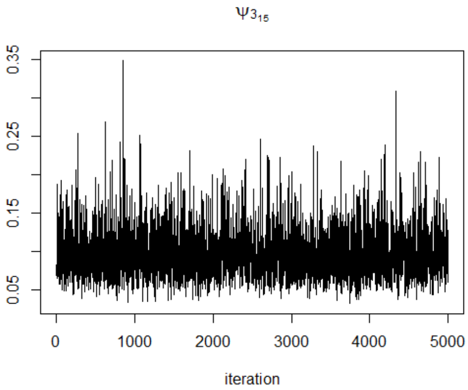
ค. trace plot ของพารามิเตอร์ $\lambda_{25(30)}$



ง. trace plot ของพารามิเตอร์ $\lambda_{26(50)}$

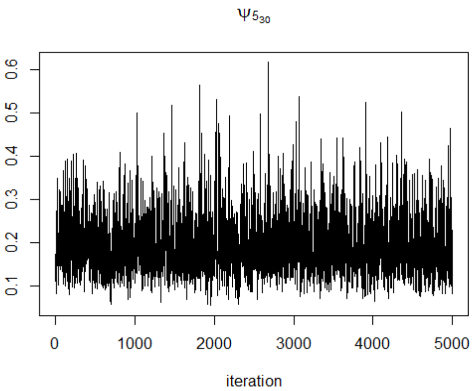


จ. trace plot ของพารามิเตอร์ $\psi_{1(1)}$

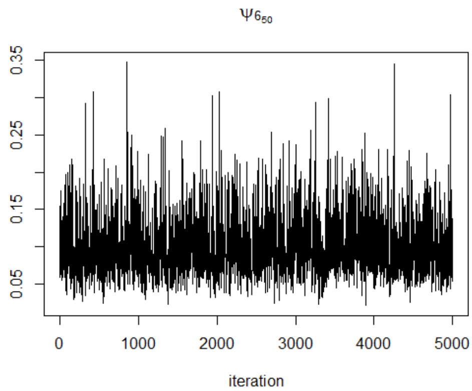


ฉ. trace plot ของพารามิเตอร์ $\psi_{3(15)}$

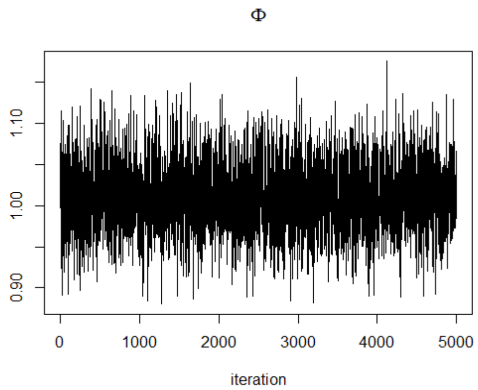
<u>รูปที่ 4.5</u> ตัวอย่าง trace plot ของตัวอย่างพารามิเตอร์ในโมเดล



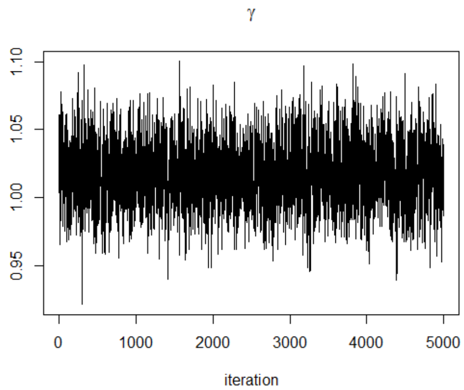
จ. trace plot ของพารามิเตอร์ $\psi_{5(30)}$



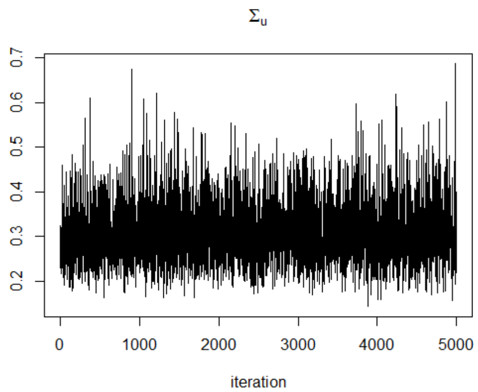
ฉ. trace plot ของพารามิเตอร์ $\psi_{6(50)}$



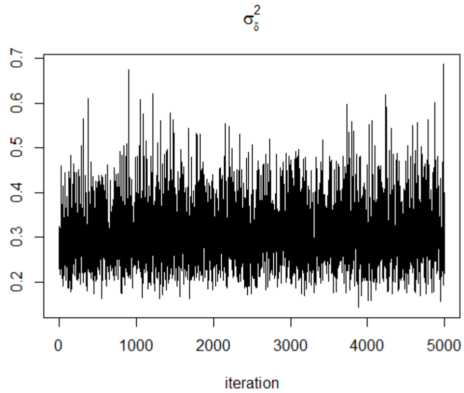
ช. trace plot ของพารามิเตอร์ Φ



ซ. trace plot ของพารามิเตอร์ γ



ฌ. trace plot ของพารามิเตอร์ Σ_u



ญ. trace plot ของพารามิเตอร์ σ_δ^2

รูปที่ 4.5 ตัวอย่าง trace plot ของตัวอย่างพารามิเตอร์ในโมเดล (ต่อ)

ผลการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์น้ำหนักองค์ประกอบแบบเบส์และวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted

จากตารางที่ 4.1 เมื่อพิจารณาค่า MSE และค่า RAMSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบให้ผลการวิจัยดังต่อไปนี้

เมื่อพิจารณาวิธี Bayes ในกรณีที่จำนวนหน่วยมีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบ λ_{mk} จะมีค่าอยู่ในช่วง [0.1118,0.1599], [0.0214,0.0340], [0.0080,0.0133] และ [0.0022,0.0038] ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนัก

องค์ประกอบ λ_{mk} จะมีค่าอยู่ในช่วง [0.0638,0.0925], [0.0220,0.0332], [0.0081,0.0140] และ [0.0023,0.0034] ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบ องค์ประกอบ λ_{mk} จะมีค่าอยู่ในช่วง [0.0559,0.0736], [0.0203,0.0296], [0.0102,0.0112] และ [0.0016,0.0027] ตามลำดับ

เมื่อพิจารณาวิธี Mplus ในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบ λ_{mk} จะมีค่าอยู่ในช่วง [0.1346,0.1600], [0.07560.1239], [0.0702,0.1103] และ [0.06810.1001] ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบ λ_{mk} จะมีค่าอยู่ในช่วง [0.0708,0.1088], [0.0580,0.0845], [0.0535,0.0751] และ [0.0505,0.0695] ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบ องค์ประกอบ λ_{mk} จะมีค่าอยู่ในช่วง [0.0707,0.0970], [0.0482,0.0578], [0.0421,0.0503] และ [0.0393,0.0447] ตามลำดับ

เมื่อพิจารณาเปรียบเทียบค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบระหว่างวิธี Bayes กับ Mplus พบว่า วิธี Bayes จะให้ค่า MSE ที่ต่ำกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา โดยเมื่อพิจารณาจากค่า RAMSE พบว่าเมื่อเปรียบเทียบกับวิธี Bayes วิธี Mplus มีค่า MSE แตกต่างจากวิธี Bayes น้อยที่สุดเท่ากับ 1.10 เท่า และมากที่สุดเท่ากับ 28.28 เท่า ผลการวิจัยที่ได้ทำให้กล่าวได้ว่าวิธี Bayes จะให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่มีความแม่นยำและมีประสิทธิภาพมากกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยค่าเฉลี่ยความเที่ยงรวม พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อระดับค่าเฉลี่ยความเที่ยงรวมเพิ่มขึ้น โดยจะมีค่า MSE ใกล้เคียงกันที่สุดเมื่อระดับค่าเฉลี่ยความเที่ยงมีค่าเท่ากับ 0.3 และสังเกตได้ว่าเมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวมมีค่าเพิ่มขึ้นวิธีการทั้งสองจะมีค่า MSE ที่แตกต่างกันมากขึ้น จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยค่าเฉลี่ยความเที่ยงรวมเป็นปัจจัยที่ส่งผลให้วิธีการประมาณค่าพารามิเตอร์ทั้งสองมีความแม่นยำและประสิทธิภาพมากยิ่งขึ้น

และวิธี Bayes จะมีอัตราการลดลงของค่า MSE เมื่อระดับค่าเฉลี่ยความเที่ยงรวมที่รวดเร็วมากกว่าวิธี Mplus

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีค่าเพิ่มขึ้น โดยจะมีค่า MSE ที่ใกล้เคียงกันมากที่สุด เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วยตัวอย่าง และสังเกตได้ว่าเมื่อปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเพิ่มขึ้น วิธี Mplus จะมีค่า MSE ที่ลดลงเข้าใกล้วิธี Bayes จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 เป็นปัจจัยที่ส่งผลให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบของทั้งสองมีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น และยังเป็นปัจจัยที่ทำให้วิธี Mplus มีค่าประมาณน้ำหนักองค์ประกอบที่ใกล้เคียงวิธี Bayes มากยิ่งขึ้น เมื่อจำนวนหน่วยตัวอย่างมีขนาดใหญ่

<u>ผลการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบเบส์และวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted</u>

จากตารางที่ 4.1 เมื่อพิจารณาค่า MSE และค่า RAMSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดให้ผลการวิจัยดังต่อไปนี้

เมื่อพิจารณาวิธี Bayes จะพบว่าในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด $\psi_{\epsilon k}$ จะมีค่าอยู่ในช่วง [0.0180,0.0268], [0.0046,0.0086], [0.0013,0.0044] และ [0.0003,0.0010] ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด $\psi_{\epsilon k}$ จะมีค่าอยู่ในช่วง [0.0106,0.0162], [0.0041,0.0064], [0.0009,0.0018] และ [0.0002,0.0009] ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด $\psi_{\epsilon k}$ จะมีค่าอยู่ในช่วง [0.0064,0.0097], [0.0021,0.0039], [0.0008,0.0023] และ [0.0002,0.0007] ตามลำดับ

เมื่อพิจารณาวิธี Mplus พบว่าเมื่อจำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด $\psi_{\epsilon k}$ จะมีค่าอยู่ในช่วง [0.1635,0.1836], [0.0750,0.0838], [0.0498,0.0660] และ [0.0263,0.0409] ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด $\psi_{\epsilon k}$ จะมีค่าอยู่ในช่วง [0.1015,0.1165], [0.0430,0.0608], [0.0314,0.0416] และ [0.0296,0.0399] ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด $\psi_{\epsilon k}$ จะมีค่าอยู่ในช่วง [0.0692,0.0805], [0.0432,0.0510], [0.0212,0.0340] และ [0.0115,0.0191] ตามลำดับ

เมื่อพิจารณาเปรียบเทียบค่า MSE ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบระหว่างวิธี Bayes กับ Mplus พบว่า วิธี Bayes จะให้ค่า MSE ที่ต่ำกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา โดยเมื่อพิจารณาจากค่า RAMSE พบว่าเมื่อเปรียบเทียบกับวิธี Bayes วิธี Mplus มีค่า MSE แตกต่างจากวิธี Bayes น้อยที่สุดเท่ากับ 7.92 เท่า และมากที่สุดเท่ากับ 68.00 เท่า ผลการวิจัยที่ได้ทำให้กล่าวได้ว่าวิธี Bayes จะให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่มีความแม่นยำและมีประสิทธิภาพมากกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยค่าเฉลี่ยความเที่ยงรวม พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อระดับค่าเฉลี่ยความเที่ยงรวมเพิ่มขึ้น โดยจะมีค่า MSE ใกล้เคียงกันมากขึ้นเมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวมมีค่าเพิ่มขึ้น จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยค่าเฉลี่ยความเที่ยงรวมเป็นปัจจัยที่ส่งผลให้วิธีการประมาณค่าพารามิเตอร์ทั้งสองมีความแม่นยำและประสิทธิภาพมากยิ่งขึ้น และทำให้วิธี Mplus มีค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ใกล้เคียงกับวิธี Bayes มากยิ่งขึ้น

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีค่าเพิ่มขึ้น โดยจะมีค่า MSE ที่ใกล้เคียงกันมากที่สุด เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วยตัวอย่าง และสังเกตได้ว่าเมื่อปัจจัย

จำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเพิ่มขึ้น วิธี Mplus จะมีค่า MSE ที่ลดลงเข้าใกล้วิธี Bayes จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 เป็นปัจจัยที่ส่งผลให้ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดของทั้งสองมีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น และยังเป็นปัจจัยที่ทำให้วิธี Mplus มีค่าประมาณความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ใกล้เคียงวิธี Bayes มากยิ่งขึ้น เมื่อจำนวนหน่วยตัวอย่างมีขนาดใหญ่

<u>ผลการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงจากการวัดแบบเบส์และวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted</u>

จากตารางที่ 4.1 เมื่อพิจารณาค่า MSE และค่า RAMSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงให้ผลการวิจัยดังต่อไปนี้

เมื่อพิจารณาวิธี Bayes ในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงจะมีค่าเท่ากับ 0.0731, 0.0150, 0.0080 และ 0.0055 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงจะมีค่าเท่ากับ 0.0267, 0.0075, 0.0042 และ 0.0029 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงจะมีค่าอยู่ในช่วง 0.0235, 0.0051, 0.0037 และ 0.0023 ตามลำดับ

เมื่อพิจารณาวิธี Mplus ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงจะมีค่าเท่ากับ 0.2026, 0.1270, 0.0985 และ 0.0810 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงจะมีค่าเท่ากับ 0.1383, 0.0880, 0.0673 และ 0.0542 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงจะมีค่าอยู่ในช่วง 0.0989, 0.0712, 0.0634 และ 0.0451 ตามลำดับ

เมื่อพิจารณาเปรียบเทียบค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงระหว่างวิธี Bayes กับ Mplus พบว่า วิธี Bayes จะให้ค่า MSE ที่ต่ำกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา โดยเมื่อพิจารณาจากค่า RAMSE พบว่าเมื่อเปรียบเทียบกับวิธี Bayes วิธี Mplus มีค่า MSE แตกต่างจากวิธี Bayes น้อยที่สุดเท่ากับ 2.77 เท่า และมากที่สุดเท่ากับ 19.61 เท่า ผลการวิจัยที่ได้ทำให้กล่าวได้ว่าวิธี Bayes จะให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่มีความแม่นยำและมีประสิทธิภาพมากกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยค่าเฉลี่ยความเที่ยงรวม พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อระดับค่าเฉลี่ยความเที่ยงรวมเพิ่มขึ้น โดยจะมีค่า MSE ใกล้เคียงกันมากขึ้นเมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวมมีค่าเพิ่มขึ้น จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยค่าเฉลี่ยความเที่ยงรวมเป็นปัจจัยที่ส่งผลให้วิธีการประมาณค่าพารามิเตอร์ทั้งสองมีความแม่นยำและประสิทธิภาพมากยิ่งขึ้น และทำให้วิธี Mplus มีค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงที่ใกล้เคียงกับวิธี Bayes มากยิ่งขึ้น

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีค่าเพิ่มขึ้น โดยจะมีค่า MSE ที่ใกล้เคียงกันมากที่สุด เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วยตัวอย่าง และสังเกตได้ว่าเมื่อปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเพิ่มขึ้น วิธี Mplus จะมีค่า MSE ที่ลดลงเข้าใกล้วิธี Bayes จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 เป็นปัจจัยที่ส่งผลให้ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดของทั้งสองมีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น และยังเป็นปัจจัยที่ทำให้วิธี Mplus มีค่าประมาณความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ใกล้เคียงวิธี Bayes มากยิ่งขึ้น เมื่อจำนวนหน่วยตัวอย่างมีขนาดใหญ่

<u>ผลการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์อิทธิพลคงที่จากการวัดแบบเบส์และวิธีภาวะความควรจะเป็นสูงสุดแบบ restricted</u>

จากตารางที่ 4.1 เมื่อพิจารณาค่า MSE และค่า RAMSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่ให้ผลการวิจัยดังต่อไปนี้

เมื่อพิจารณาวิธี Bayes ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่จะมีค่าเท่ากับ 0.0217, 0.0155, 0.0141 และ 0.0140 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่จะมีค่าเท่ากับ 0.0053, 0.0048, 0.0039 และ 0.0024 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่จะมีค่าอยู่ในช่วง 0.0039, 0.0032, 0.0026 และ 0.0019 ตามลำดับ

เมื่อพิจารณาวิธี Mplus ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่จะมีค่าเท่ากับ 0.1921, 0.1469, 0.1331 และ 0.1266 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่จะมีค่าเท่ากับ 0.1333, 0.1013, 0.0928 และ 0.0907 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่จะมีค่าอยู่ในช่วง 0.1009, 0.0734, 0.0679 และ 0.0655 ตามลำดับ

เมื่อพิจารณาเปรียบเทียบค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงระหว่างวิธี Bayes กับ Mplus พบว่า วิธี Bayes จะให้ค่า MSE ที่ต่ำกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา โดยเมื่อพิจารณาจากค่า RAMSE พบว่าเมื่อเปรียบเทียบกับวิธี Bayes วิธี Mplus มีค่า MSE แตกต่างจากวิธี Bayes น้อยที่สุดเท่ากับ 8.85 เท่า และมากที่สุดเท่ากับ 37.79 เท่า ผลการวิจัยที่ได้ทำให้กล่าวได้ว่าวิธี Bayes จะให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่มีความแม่นยำและมีประสิทธิภาพมากกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยค่าเฉลี่ยความเที่ยงรวม พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อระดับค่าเฉลี่ยความเที่ยงรวมเพิ่มขึ้น โดยจะมีค่า MSE ใกล้เคียงกันมากขึ้นเมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวมมีค่าเพิ่มขึ้น จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยค่าเฉลี่ยความเที่ยงรวมเป็นปัจจัยที่ส่งผล

ให้วิธีการประมาณค่าพารามิเตอร์ทั้งสองมีความแม่นยำและประสิทธิภาพมากยิ่งขึ้น และทำให้วิธี Mplus มีค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงที่ใกล้เคียงกับวิธี Bayes มากยิ่งขึ้น

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีค่าเพิ่มขึ้น โดยจะมีค่า MSE ที่ใกล้เคียงกันมากที่สุด เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วยตัวอย่าง และสังเกตได้ว่าเมื่อปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเพิ่มขึ้น วิธี Mplus จะมีค่า MSE ที่ลดลงเข้าใกล้วิธี Bayes จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 เป็นปัจจัยที่ส่งผลให้ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดของทั้งสองมีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น และยังเป็นปัจจัยที่ทำให้วิธี Mplus มีค่าประมาณความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ใกล้เคียงวิธี Bayes มากยิ่งขึ้น เมื่อจำนวนหน่วยตัวอย่างมีขนาดใหญ่

<u>ผลการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1 จากการวัดแบบเบส์และวิธีภาวะความควรจะเป็นสูงสุด</u>

จากตารางที่ 4.1 เมื่อพิจารณาค่า MSE และค่า RAMSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1 ให้ผลการวิจัยดังต่อไปนี้

เมื่อพิจารณาวิธี Bayes ในกรณีที่จำนวนตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณความแปรปรวนของโมเดลในระดับที่ 1จะมีค่าเท่ากับ 0.0595, 0.0153, 0.0066 และ 0.0036 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1จะมีค่าเท่ากับ 0.0212, 0.0083, 0.0032 และ 0.0023 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1จะมีค่าอยู่ในช่วง 0.0165, 0.0038, 0.0029 และ 0.0012 ตามลำดับ

เมื่อพิจารณาวิธี Mplus ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1จะมีค่าเท่ากับ 0.2107, 0.1359, 0.1113 และ 0.0943 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1จะมีค่าเท่ากับ 0.1554, 0.1031, 0.0820 และ 0.0662 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1จะมีค่าอยู่ในช่วง 0.1174, 0.0822, 0.0652 และ 0.0540 ตามลำดับ

เมื่อพิจารณาเปรียบเทียบค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลระดับที่ 1 ระหว่างวิธี Bayes กับ Mplus พบว่า วิธี Bayes จะให้ค่า MSE ที่ต่ำกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา โดยเมื่อพิจารณาจากค่า RAMSE พบว่าเมื่อเปรียบเทียบกับวิธี Bayes วิธี Mplus มีค่า MSE แตกต่างจากวิธี Bayes น้อยที่สุดเท่ากับ 3.54 เท่า และมากที่สุดเท่ากับ 45.00 เท่า ผลการวิจัยที่ได้ทำให้กล่าวได้ว่าวิธี Bayes จะให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่มีความแม่นยำและมีประสิทธิภาพมากกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยค่าเฉลี่ยความเที่ยงรวม พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อระดับค่าเฉลี่ยความเที่ยงรวมเพิ่มขึ้น โดยจะมีค่า MSE ใกล้เคียงกันมากขึ้นเมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวมมีค่าเพิ่มขึ้น จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยค่าเฉลี่ยความเที่ยงรวมเป็นปัจจัยที่ส่งผลให้วิธีการประมาณค่าพารามิเตอร์ทั้งสองมีความแม่นยำและประสิทธิภาพมากยิ่งขึ้น และทำให้วิธี Mplus มีค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1 ที่ใกล้เคียงกับวิธี Bayes มากยิ่งขึ้น

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีค่าเพิ่มขึ้น โดยจะมีค่า MSE ที่ใกล้เคียงกันมากที่สุด เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วยตัวอย่าง และสังเกตได้ว่าเมื่อปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเพิ่มขึ้น วิธี Mplus จะมีค่า MSE ที่ลดลงเข้าใกล้วิธี Bayes จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 เป็นปัจจัยที่

ส่งผลให้ค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1 ของทั้งสองมีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น และยังเป็นปัจจัยที่ทำให้วิธี Mplus มีค่าประมาณความแปรปรวนของโมเดลในระดับที่ 1 ที่ใกล้เคียงวิธี Bayes มากยิ่งขึ้น เมื่อจำนวนหน่วยตัวอย่างมีขนาดใหญ่

<u>ผลการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 จากการวัดแบบเบส์และวิธีภาวะความควรจะเป็นสูงสุด</u>

จากตารางที่ 4.1 เมื่อพิจารณาค่า MSE และค่า RAMSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 ให้ผลการวิจัยดังต่อไปนี้

เมื่อพิจารณาวิธี Bayes ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณความแปรปรวนของโมเดลในระดับที่ 2 จะมีค่าเท่ากับ 0.0200, 0.0089, 0.0072 และ 0.0053 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 จะมีค่าเท่ากับ 0.0053, 0.0045, 0.0037 และ 0.0032 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1จะมีค่าอยู่ในช่วง 0.0058, 0.0034, 0.0021 และ 0.0018 ตามลำดับ

เมื่อพิจารณาวิธี Mplus ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 15 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 จะมีค่าเท่ากับ 0.1404, 0.1209, 0.1158 และ 0.1059 ตามลำดับ ในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 30 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 จะมีค่าเท่ากับ 0.0959, 0.0845, 0.0811 และ 0.0729 ตามลำดับ และในกรณีที่จำนวนตัวอย่างในระดับที่ 2มีจำนวนเท่ากับ 50 หน่วย ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 จะมีค่าอยู่ในช่วง 0.0707, 0.0609, 0.0583 และ 0.0534 ตามลำดับ

เมื่อพิจารณาเปรียบเทียบค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลระดับที่ 2 ระหว่างวิธี Bayes กับ Mplus พบว่า วิธี Bayes จะให้ค่า MSE ที่ต่ำกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา โดยเมื่อพิจารณาจากค่า RAMSE พบว่าเมื่อเปรียบเทียบกับวิธี Bayes

วิธี Mplus มีค่า MSE แตกต่างจากวิธี Bayes น้อยที่สุดเท่ากับ 7.02 เท่า และมากที่สุดเท่ากับ 29.67 เท่า ผลการวิจัยที่ได้ทำให้กล่าวได้ว่าวิธี Bayes จะให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่มีความแม่นยำและมีประสิทธิภาพมากกว่าวิธี Mplus ในทุกกรณีที่ทำการศึกษา

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยค่าเฉลี่ยความเที่ยงรวม พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อระดับค่าเฉลี่ยความเที่ยงรวมเพิ่มขึ้น โดยจะมีค่า MSE ใกล้เคียงกันมากขึ้นเมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวมมีค่าเพิ่มขึ้น จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยค่าเฉลี่ยความเที่ยงรวมเป็นปัจจัยที่ส่งผลให้วิธีการประมาณค่าพารามิเตอร์ทั้งสองมีความแม่นยำและประสิทธิภาพมากยิ่งขึ้น และทำให้วิธี Mplus มีค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 ที่ใกล้เคียงกับวิธี Bayes มากยิ่งขึ้น

พิจารณาเปรียบเทียบแนวโน้มของค่า MSE ระหว่างวิธี Bayes กับ Mplus โดยพิจารณาจากมิติของปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 พบว่า วิธีการทั้งสองจะมีค่า MSE ลดลง เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีค่าเพิ่มขึ้น โดยจะมีค่า MSE ที่ใกล้เคียงกันมากที่สุด เมื่อจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเท่ากับ 50 หน่วยตัวอย่าง และสังเกตได้ว่าเมื่อปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 มีจำนวนเพิ่มขึ้น วิธี Mplus จะมีค่า MSE ที่ลดลงเข้าใกล้วิธี Bayes จากผลการวิจัยนี้ทำให้กล่าวได้ว่า ปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 เป็นปัจจัยที่ส่งผลให้ค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 ของทั้งสองมีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น และยังเป็นปัจจัยที่ทำให้วิธี Mplus มีค่าประมาณความแปรปรวนของโมเดลในระดับที่ 2 ที่ใกล้เคียงวิธี Bayes มากยิ่งขึ้น เมื่อจำนวนหน่วยตัวอย่างมีขนาดใหญ่

<u>สรุปผลการศึกษาด้วยข้อมูลจำลอง</u>

จากผลการจำลองในข้างต้นสามารถสรุปได้ว่าวิธี Bayes เป็นวิธีประมาณพารามิเตอร์ที่ให้ค่าประมาณในส่วนต่างๆของโมเดลอย่างแม่นยำและมีประสิทธิภาพเหนือวิธี Mplus ในทุกกรณีที่ทำการศึกษา หรือกล่าวได้ว่าวิธี Bayes เป็นวิธีที่มีความสามารถเหนือกว่าวิธี Mplus ในทุกสถานการณ์จำลอง สาเหตุเนื่องมาจาก การประมาณค่าพารามิเตอร์ด้วยวิธี Bayes ได้มีการ

คำนึงถึงความผันแปรระหว่างกลุ่มของพารามิเตอร์ในโมเดลการวัด ในขณะที่การประมาณค่าพารามิเตอร์ด้วยวิธี Mplus กำหนดให้เป็นค่าคงที่ วิธี Bayes จึงเป็นการประมาณค่าพารามิเตอร์ภายใต้โมเดลที่มีความสอดคล้องกับธรรมชาติของข้อมูลมากกว่าวิธี Mplus นอกจากนี้ผลการศึกษายังบ่งชี้ว่าจะเห็นว่าปัจจัยค่าเฉลี่ยความเที่ยงรวม และจำนวนตัวอย่างในระดับที่ 2เป็นปัจจัยที่มีผลกระทบต่อความแม่นยำและประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ทั้งสอง โดยปัจจัยทั้งสองจะทำให้วิธีการทั้งสองมีค่า MSE ที่ลดต่ำลง และวิธี Mplus จะมีค่า MSE ที่ลู่เข้าใกล้วิธี Bayes เมื่อปัจจัยทั้งสองมีค่าเพิ่มขึ้น หรือกล่าวได้ว่า วิธี Mplus จะมีความถูกต้องมากยิ่งขึ้น เปรียบเทียบกับวิธี Bayes เมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวม หรือปัจจัยจำนวนหน่วยตัวอย่างในระดับที่ 2 มีค่าเพิ่มขึ้น สาเหตุเนื่องจากการเพิ่มระดับของค่าเฉลี่ยความเที่ยงจะทำให้มีความคงเส้นคงวาในการวัดตัวแปรแฝงจากตัวแปรสังเกตได้มากขึ้น ทำให้ตัวแปรแฝงที่ประมาณค่าได้จากวิธีการต่างๆมีความใกล้เคียงตัวแปรแฝงที่แท้จริงมากยิ่งขึ้น ส่งผลให้ค่าประมาณพารามิเตอร์ที่ได้มีความถูกต้องมากยิ่งขึ้น ในส่วนของจำนวนตัวอย่างในระดับที่ 2 การที่หน่วยตัวอย่างมีจำนวนเพิ่มมากขึ้น ย่อมทำให้มีสารสนเทศที่ให้ในการประมาณค่าพารามิเตอร์ในโมเดลมากยิ่งขึ้นตามไปด้วย จึงส่งผลโดยตรงต่อคุณภาพของค่าประมาณพารามิเตอร์ อย่างไรก็ตามเป็นที่สังเกตได้ว่าปัจจัยขนาดหน่วยตัวอย่างในระดับที่ 2 มีผลต่อค่า MSE ของวิธี Bayes น้อยกว่าวิธี Mplus ทั้งนี้เพราะวิธี Mplus เป็นวิธีที่ขึ้นกับทฤษฎีการลู่เข้าเชิงกำกับ (asymptotic theory) กล่าวคือคุณสมบัติที่ดีของวิธี Mplus จะเกิดขึ้นเมื่อจำนวนหน่วยตัวอย่างมีจำนวนที่ใหญ่เพียงพอ ในขณะที่วิธี Bayes ไม่ได้ขึ้นกับทฤษฎีดังกล่าวจึงมีความแกร่งต่อกรณีที่จำนวนหน่วยตัวอย่างมีขนาดตัวอย่างเล็กมากกว่า ทำให้ปัจจัยจำนวนหน่วยตัวอย่างจึงทำให้ปัจจัยที่มีความสำคัญต่อวิธี Mplus มากกว่าวิธี Bayes

<u>ตารางที่ 4.1</u> ค่าความคลาดเคลื่อนกำลังสอง (MSE) ของพารามิเตอร์เปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุด แบบ restricted จากโปรแกรม Mplus

J	พารามิเตอร์	Bayes				Mplus (ML)			
		$\overline{\rho^2}$ =0.3	$\overline{\rho^2}$=0.5	$\overline{\rho^2}$=0.7	$\overline{\rho^2}$=0.9	$\overline{\rho^2}$=0.3	$\overline{\rho^2}$=0.5	$\overline{\rho^2}$=0.7	$\overline{\rho^2}$=0.9

15	λ_{12}	0.1365	0.0282	0.0101	0.0025	0.16	0.1239	0.1103	0.1001
	λ_{13}	0.1118	0.0214	0.008	0.0022	0.1462	0.0756	0.0702	0.0681
	λ_{22}	0.1599	0.034	0.0133	0.0038	0.1545	0.1129	0.0962	0.0858
	λ_{23}	0.1343	0.0286	0.0111	0.0031	0.1346	0.0944	0.0814	0.0738
	เฉลี่ย	0.1356	0.0281	0.0106	0.0029	0.1488	0.1017	0.0895	0.082
	RDMSE	1.10	3.62	8.44	28.28				
30	λ_{12}	0.0903	0.0276	0.0106	0.0025	0.1088	0.0845	0.0751	0.0695
	λ_{13}	0.0638	0.022	0.0081	0.0023	0.0708	0.058	0.0535	0.0505
	λ_{22}	0.0925	0.0332	0.014	0.0034	0.0955	0.0698	0.0613	0.0572
	λ_{23}	0.0815	0.025	0.0122	0.0028	0.097	0.0688	0.0585	0.0529
	เฉลี่ย	0.082	0.0269	0.0112	0.0028	0.093	0.0703	0.0621	0.0575
	RDMSE	1.13	2.61	5.54	20.54				
50	λ_{12}	0.0733	0.0245	0.0102	0.002	0.0921	0.0525	0.047	0.0444
	λ_{13}	0.0559	0.0203	0.011	0.0016	0.0707	0.0482	0.0445	0.0422
	λ_{22}	0.0736	0.0296	0.0109	0.0027	0.097	0.0578	0.0503	0.0447
	λ_{23}	0.0638	0.0253	0.0112	0.0022	0.0852	0.049	0.0421	0.0393
	เฉลี่ย	0.0667	0.0249	0.0108	0.0021	0.0863	0.0519	0.046	0.0427
	RDMSE	1.29	2.08	4.26	20.33				

ตารางที่ 4.1 (ต่อ) ค่าความคลาดเคลื่อนกำลังสอง (MSE) ของพารามิเตอร์เปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus

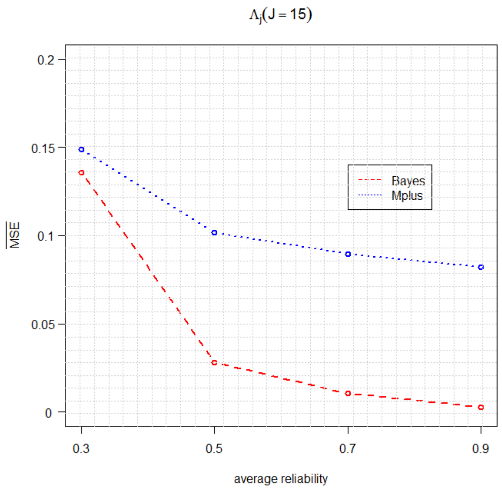
J	พารามิเตอร์	Bayes				Mplus (ML)			
		$\overline{\rho^2}$=0.3	$\overline{\rho^2}$=0.5	$\overline{\rho^2}$=0.7	$\overline{\rho^2}$=0.9	$\overline{\rho^2}$=0.3	$\overline{\rho^2}$=0.5	$\overline{\rho^2}$=0.7	$\overline{\rho^2}$=0.9
15	$\psi_{\epsilon 1}$	0.0268	0.0086	0.0044	0.0009	0.1836	0.0838	0.0563	0.0323
	$\psi_{\epsilon 2}$	0.026	0.0055	0.0015	0.0004	0.1799	0.0776	0.0564	0.0409
	$\psi_{\epsilon 3}$	0.0227	0.0051	0.0015	0.0003	0.1767	0.0811	0.0606	0.0309
	$\psi_{\epsilon 4}$	0.0208	0.0073	0.0033	0.001	0.1785	0.0812	0.066	0.0276

	$\psi_{\epsilon 5}$	0.018	0.0051	0.0015	0.0004	0.1635	0.0759	0.0505	0.0312
	$\psi_{\epsilon 6}$	0.0194	0.0046	0.0013	0.0003	0.1782	0.075	0.0498	0.0263
	เฉลี่ย	0.0223	0.006	0.0023	0.0006	0.1767	0.0791	0.0566	0.0315
	RDMSE	7.92	13.18	24.61	52.50				
30	$\psi_{\epsilon 1}$	0.0134	0.0045	0.0014	0.0009	0.1165	0.052	0.034	0.0334
	$\psi_{\epsilon 2}$	0.0129	0.0064	0.0012	0.0003	0.1015	0.0608	0.0387	0.0399
	$\psi_{\epsilon 3}$	0.0106	0.0041	0.0009	0.0002	0.1147	0.0599	0.0416	0.0348
	$\psi_{\epsilon 4}$	0.0162	0.0058	0.0018	0.0008	0.1131	0.0543	0.0314	0.0296
	$\psi_{\epsilon 5}$	0.014	0.0063	0.0013	0.0003	0.1085	0.0501	0.0332	0.0362
	$\psi_{\epsilon 6}$	0.0118	0.0042	0.0011	0.0002	0.1061	0.043	0.0382	0.0301
	เฉลี่ย	0.0132	0.0052	0.0013	0.0005	0.1101	0.0534	0.0362	0.034
	RDMSE	8.34	10.27	27.85	68.00				
50	$\psi_{\epsilon 1}$	0.0097	0.0037	0.0014	0.0007	0.0754	0.0449	0.0248	0.0173
	$\psi_{\epsilon 2}$	0.008	0.0026	0.0017	0.0003	0.0708	0.051	0.0266	0.0122
	$\psi_{\epsilon 3}$	0.0077	0.0033	0.001	0.0002	0.0783	0.0443	0.034	0.0115
	$\psi_{\epsilon 4}$	0.0085	0.0039	0.0023	0.0006	0.0805	0.0438	0.0253	0.0152
	$\psi_{\epsilon 5}$	0.0064	0.0021	0.0011	0.0003	0.0786	0.0486	0.0255	0.0191
	$\psi_{\epsilon 6}$	0.0068	0.0026	0.0008	0.0002	0.0692	0.0432	0.0212	0.0149
	เฉลี่ย	0.0078	0.003	0.0014	0.0004	0.0755	0.046	0.0262	0.015
	RDMSE	9.68	15.33	18.71	37.50				

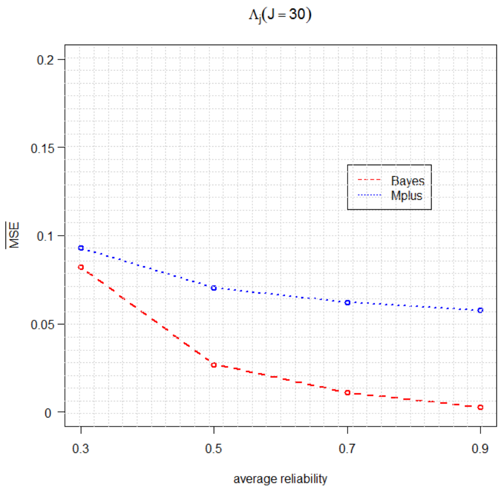
ตารางที่ 4.1 (ต่อ)ค่าความคลาดเคลื่อนกำลังสอง (MSE) ของพารามิเตอร์เปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุด แบบ restricted จากโปรแกรม Mplus

J	พารามิเตอร์	Bayes				Mplus (ML)			
		$\overline{\rho^2}$=0.3	$\overline{\rho^2}$=0.5	$\overline{\rho^2}$=0.7	$\overline{\rho^2}$=0.9	$\overline{\rho^2}$=0.3	$\overline{\rho^2}$=0.5	$\overline{\rho^2}$=0.7	$\overline{\rho^2}$=0.9
15	Φ	0.0731	0.015	0.008	0.0055	0.2026	0.127	0.0985	0.081
	RDMSE	2.77	8.47	12.31	14.73				
30	Φ	0.0267	0.0075	0.0042	0.0029	0.1383	0.088	0.0673	0.0542
	RDMSE	5.18	11.73	16.02	18.69				
50	Φ	0.0235	0.0051	0.0037	0.0023	0.0989	0.0712	0.0634	0.0451
	RDMSE	4.21	13.96	17.14	19.61				

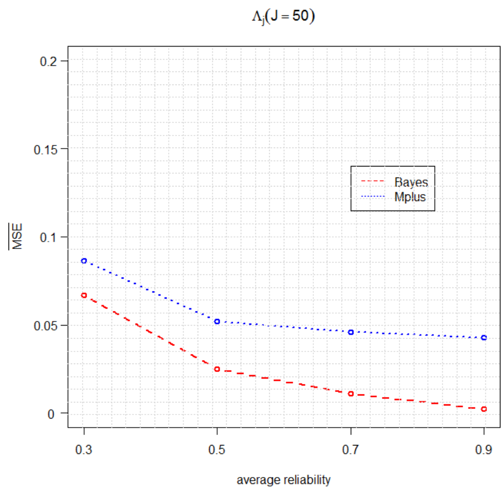
15	γ	0.0217	0.0155	0.0141	0.014	0.1921	0.1469	0.1331	0.1266
	RDMSE	8.85	9.48	9.44	9.04				
30	γ	0.0053	0.0048	0.0039	0.0024	0.1333	0.1013	0.0928	0.0907
	RDMSE	25.15	21.10	23.79	37.79				
50	γ	0.0039	0.0032	0.0026	0.0019	0.1009	0.0734	0.0679	0.0655
	RDMSE	25.87	22.94	26.12	34.47				
15	σ_δ^2	0.0595	0.0153	0.0066	0.0036	0.2107	0.1359	0.1113	0.0943
	RDMSE	3.54	8.88	16.86	26.19				
30	σ_δ^2	0.0212	0.0083	0.0032	0.0023	0.1554	0.1031	0.082	0.0662
	RDMSE	7.33	12.42	25.63	28.78				
50	σ_δ^2	0.0165	0.0038	0.0029	0.0012	0.1174	0.0822	0.0652	0.054
	RDMSE	7.12	21.63	22.48	45.00				
15	Σ_u	0.02	0.0089	0.0072	0.0053	0.1404	0.1209	0.1158	0.1059
	RDMSE	7.02	13.58	16.08	19.98				
30	Σ_u	0.0053	0.0045	0.0037	0.0032	0.0959	0.0845	0.0811	0.0729
	RDMSE	18.09	18.78	21.92	22.78				
50	Σ_u	0.0058	0.0034	0.0021	0.0018	0.0707	0.0609	0.0583	0.0534
	RDMSE	12.19	17.91	27.76	29.67				



ก. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 15 กลุ่ม

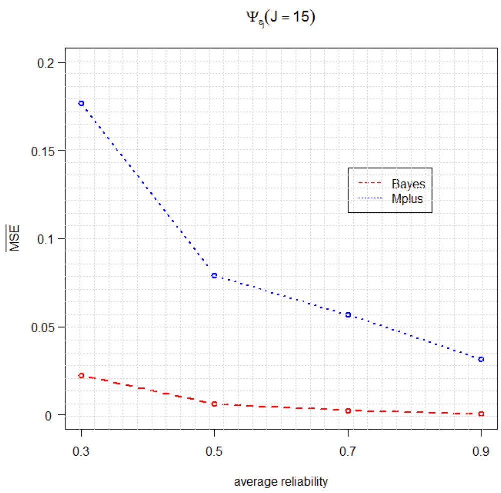


ข. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 30 กลุ่ม

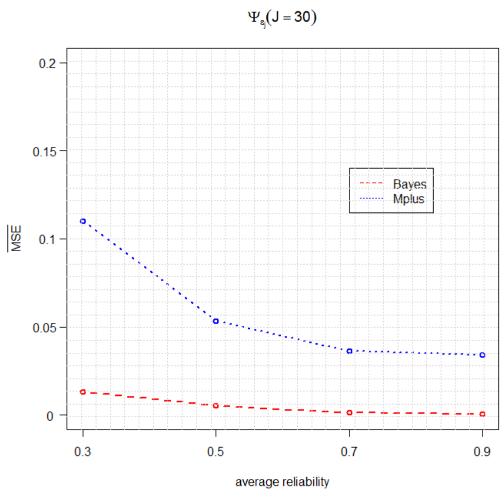


ค. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 50 กลุ่ม

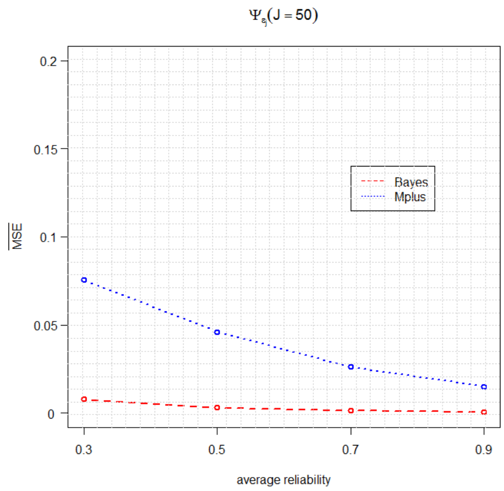
รูปที่ 4.6 ค่า $\overline{MSE}$ ของค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบเปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม เมื่อกำหนดให้ $\overline{\rho^2}$ มีค่าเท่ากับ 0.3, 0.5, 0.7 และ 0.9 และ J=15, 30 และ 50



ก. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 15 กลุ่ม

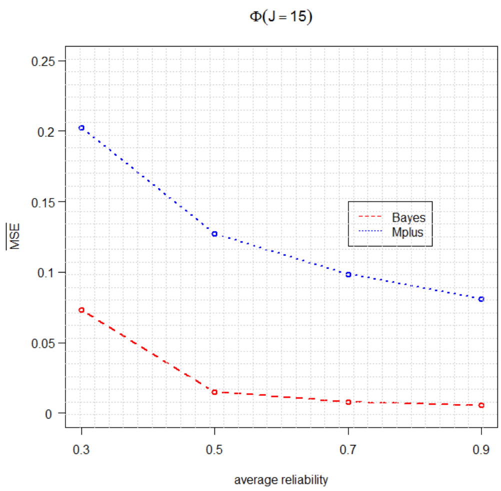


ข. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 30 กลุ่ม

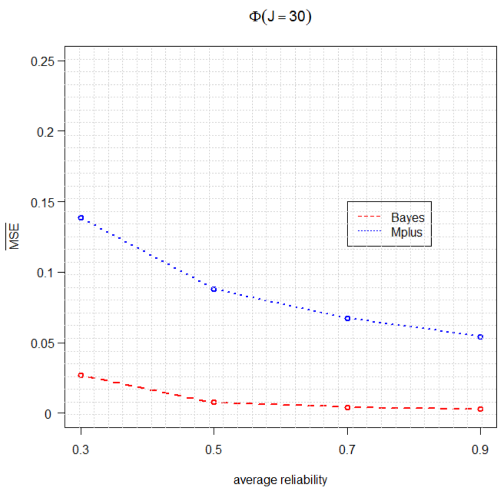


ค. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 50 กลุ่ม

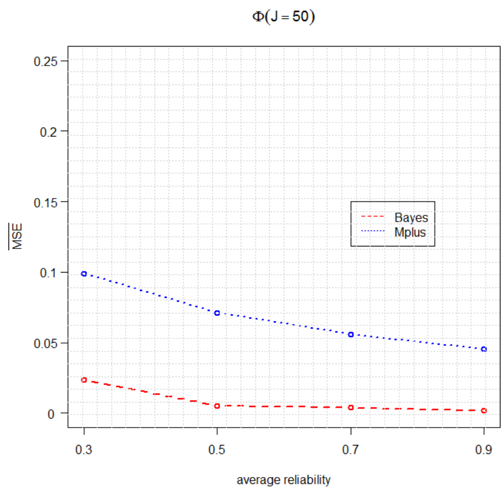
รูปที่ 4.7 ค่า $\overline{MSE}$ ของค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด เปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus เมื่อกำหนดให้ $\overline{\rho^2}$ มีค่าเท่ากับ 0.3, 0.5, 0.7 และ 0.9 และ J=15, 30 และ 50



ก. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 15 กลุ่ม

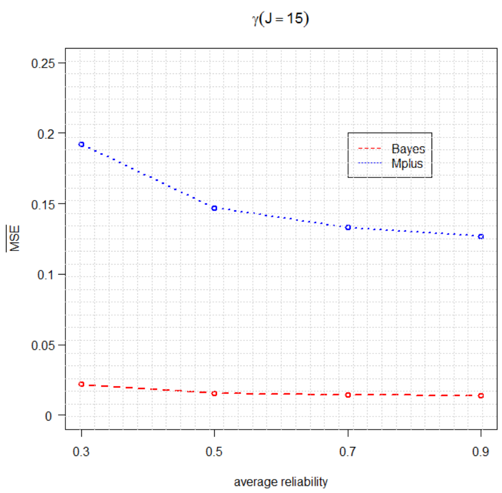


ข. ค่า $\overline{MSE}$ เมื่อหน่วยเท่ากับ 30 กลุ่ม

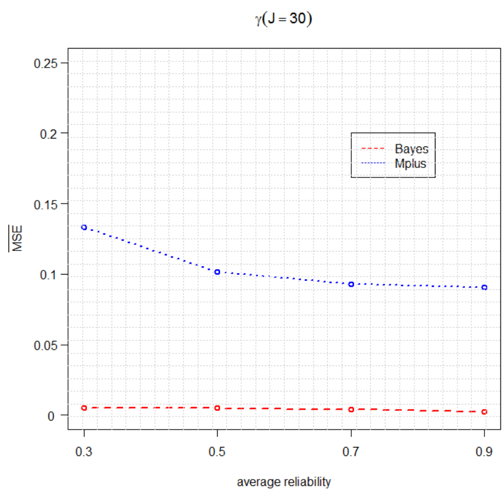


ค. ค่า MSE เมื่อหน่วยเท่ากับ 50 กลุ่ม

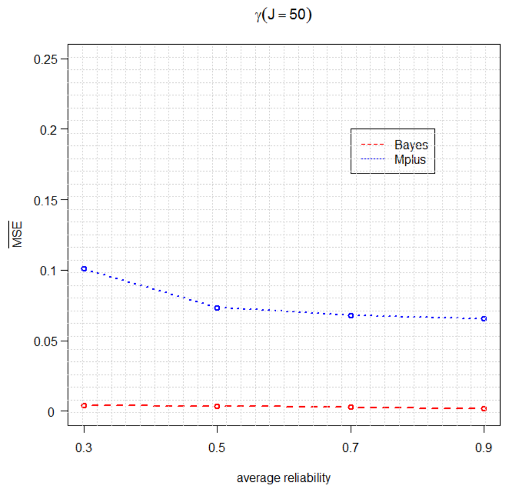
รูปที่ 4.8 ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝงเปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus เมื่อกำหนดให้ $\overline{\rho^2}$ มีค่าเท่ากับ 0.3, 0.5, 0.7 และ 0.9 และ J=15, 30 และ 50



ก. ค่า MSE เมื่อหน่วยเท่ากับ 15 กลุ่ม

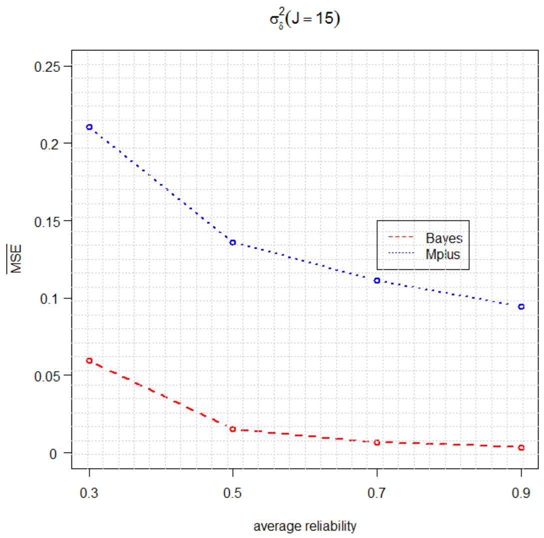


ข. ค่า MSE เมื่อหน่วยเท่ากับ 30 กลุ่ม

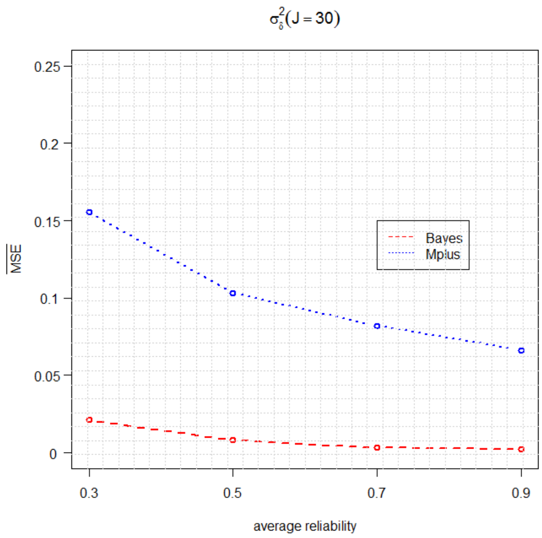


ค. ค่า MSE เมื่อหน่วยเท่ากับ 50 กลุ่ม

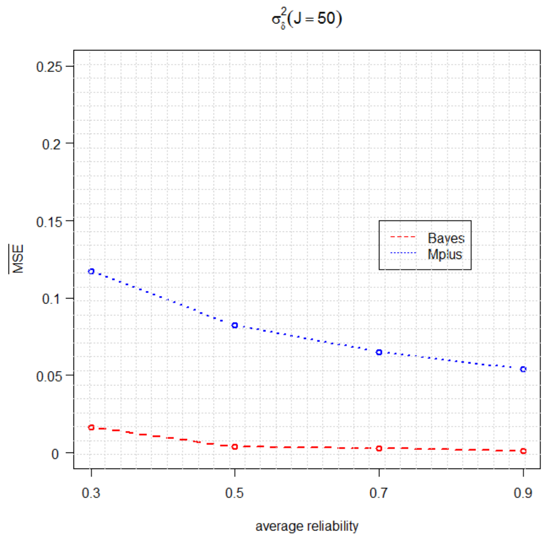
<u>รูปที่ 4.9</u> ค่า MSE ของค่าประมาณพารามิเตอร์อิทธิพลคงที่เปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus เมื่อกำหนดให้ $\overline{\rho^2}$ มีค่าเท่ากับ 0.3, 0.5, 0.7 และ 0.9 และ J=15, 30 และ 50



ก. ค่า MSE เมื่อหน่วยเท่ากับ 15 กลุ่ม

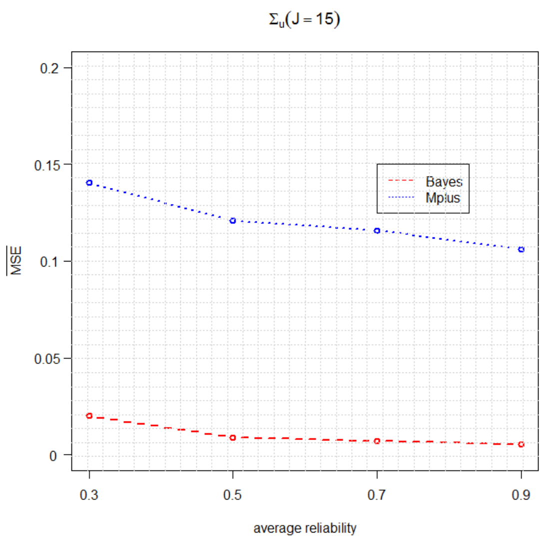


ข. ค่า MSE เมื่อหน่วยเท่ากับ 30 กลุ่ม

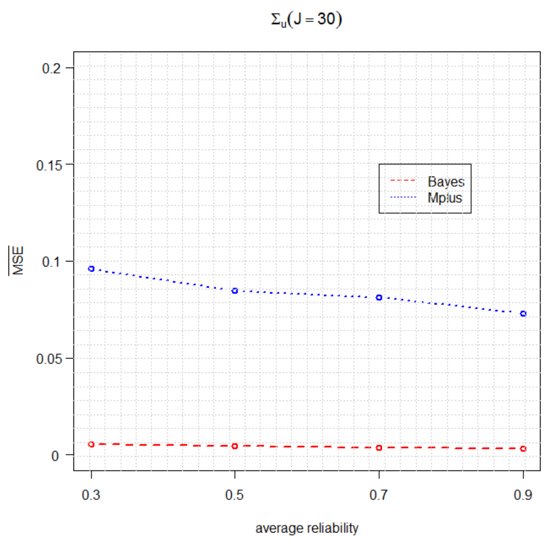


ค. ค่า MSE เมื่อหน่วยเท่ากับ 50 กลุ่ม

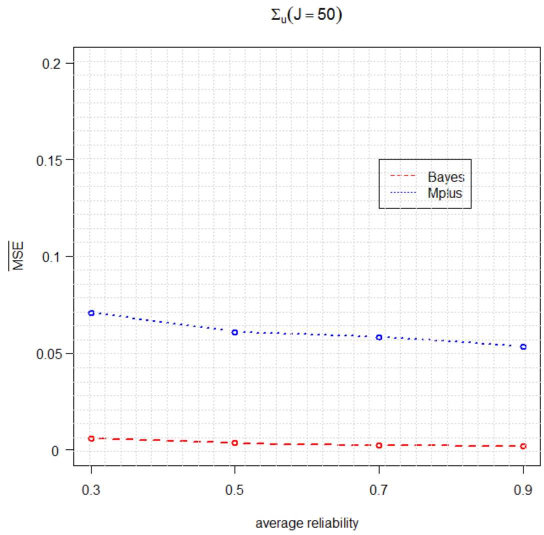
รูปที่ 4.10 ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1 เปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus เมื่อกำหนดให้ $\overline{\rho^2}$ มีค่าเท่ากับ 0.3, 0.5, 0.7 และ 0.9 และ J=15, 30 และ 50



ก. ค่า MSE เมื่อหน่วยเท่ากับ 15 กลุ่ม



ข. ค่า MSE เมื่อหน่วยเท่ากับ 30 กลุ่ม



ค. ค่า MSE เมื่อหน่วยเท่ากับ 50 กลุ่ม

<u>รูปที่ 4.11</u> ค่า MSE ของค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2 เปรียบเทียบระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ กับวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดแบบ restricted จากโปรแกรม Mplus เมื่อกำหนดให้ $\overline{\rho^2}$ มีค่าเท่ากับ 0.3, 0.5, 0.7 และ 0.9 และ J=15, 30 และ 50

ตอนที่ 3: การวิเคราะห์โดยใช้ข้อมูลทุติยภูมิ

ในกรณีนี้ผู้วิจัยจะเปรียบเทียบวิธีการประมาณค่าแบบเบส์และวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted โดยใช้ข้อมูลทุติยภูมิจากวิทยานิพนธ์ระดับดุษฎีบัณฑิต สาขาวิชาวิธีวิทยาวิจัยทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เรื่อง “อิทธิพลของพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลและสุขภาวะครูที่มีต่อสุขภาวะของนักเรียน : โมเดลการปรับและการส่งผ่านพหุระดับ” โดย นางสาวถมรัตน์ ศิริภาพ (ถมรัตน์ ศิริภาพ, 2554) ข้อมูลตัวอย่างที่ใช้เป็นข้อมูลของนักเรียนรวมทั้งสิ้น 2706

คน ติดอยู่ในหน่วยครูจำนวน 71 คนจาก 71 ห้องเรียนใน 36 โรงเรียนทั่วทุกภาคของประเทศ โดยผู้วิจัยเลือกใช้ตัวแปรจากงานวิจัยดังกล่าวจำนวน 3 ตัวแปรดังนี้

1) ตัวแปรสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) วัดได้จากตัวแปรสังเกตได้จำนวน 6 ตัวแปรได้แก่ เจตคติและอารมณ์เชิงบวก (AFF) ความเพลิดเพลิน (ENJ) อัตมโนทัศน์เชิงวิชาการ (SEL) ปัญหาทางสังคม (SOC) ความวิตกกังวล (WOR) และปัญหาสุขภาพกาย (PHY)
2) ตัวแปรพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) และตัวแปรผลสัมฤทธิ์ทางการเรียนในระดับนักเรียน (ACH_w) วัดได้จากตัวแปรสังเกตได้จำนวน 4 ตัวแปรได้แก่ การใช้อำนาจ-การร่วมมือ (DC) การร่วมมือ-การคล้อยตาม (CS) การคล้อยตาม-การต่อต้าน (SO) และการต่อต้าน-การใช้อำนาจ (OD)
3) ตัวแปรผลสัมฤทธิ์ทางการเรียนพิจารณาจากเกรดเฉลี่ยสะสมของนักเรียน

โดยวิธีการประมาณค่าพารามิเตอร์แบบเบส์ผู้วิจัยจะวิเคราะห์โดยใช้โปรแกรม R ส่วนวิธีการประมาณค่าพารามิเตอร์แบบภาวะความควรจะเป็นสูงสุดแบบ restricted ผู้วิจัยจะวิเคราะห์โดยใช้โปรแกรม Mplus ในการวิเคราะห์ข้อมูลผู้วิจัยกำหนดโมเดลการวิเคราะห์ดังสมการต่อไปนี้

<u>โมเดลการวัด</u>

$$\underline{y}_{ij} = \underline{\mu} + \Lambda_j \underline{\omega}_{ij} + \underline{\varepsilon}_{ij} \qquad (4.67)$$

เมื่อ $\underline{y}_{ij} = (AFF, ENJ, SEL, SOC, WOR, PHY, DC, CS, SO, OD)^T$,

$\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_{10})^T$,

$\Lambda_j = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & \lambda_{14} & 1^* & \lambda_{16} & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 0^* & 0^* & \lambda_{21} & \lambda_{22} & \lambda_{23} & \lambda_{24} \end{bmatrix}_j^T$,

$\underline{\omega}_{ij} = (SWB_w, SIB_w)_{ij}^T$,

$\underline{\varepsilon}_{ij} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{10})_{ij}^T$ โดยที่ $\varepsilon_{kij} \sim N(\underline{0}, \psi_{\varepsilon k(j)})$

และ ε_{kij} กับ $\varepsilon_{ki'j'}$ เป็นอิสระซึ่งกันและกัน

<u>โมเดลระดับที่ 1 (level-1 model)</u>

$$SWB_{wij} = \Xi_{ij} \underline{\beta_j} + \delta_{ij} \qquad (4.68\ ก)$$

เมื่อ $\Xi_{ij} = (ACH_w \quad SIB_w)_{ij}$, $\underline{\beta_j} = (\beta_1, \beta_2)_j^T$ และ $\delta_{ij} \sim N(0, \sigma_\delta^2)$

<u>โมเดลระดับที่ 2 (level-2 model)</u>

$$\beta_j = \underline{\gamma} + \underline{u}_j \quad (4.68\ ข)$$

เมื่อ $\underline{\gamma} = (\gamma_{01}, \gamma_{02})^T$, $\underline{u}_j = (u_{1j}, 0)^T \sim MVN(\underline{0}, \Sigma_u)$ โดยที่ $\Sigma_u = \begin{bmatrix} \tau_{11} & 0 \\ 0 & 0 \end{bmatrix}$

และ $\underline{u}_j$ กับ $\underline{u}_{j'}$ เป็นอิสระซึ่งกันและกัน

ในกรณีที่ประมาณค่าด้วยวิธีการประมาณค่าแบบเบส์ โมเดลที่ใช้ในการประมาณคือโมเดลการวิเคราะห์ที่พารามิเตอร์ในโมเดลการวัดมีความผันแปรระหว่างหน่วยดังในสมการที่ 4.67 ถึง 4.68 แต่ในกรณีที่ประมาณค่าด้วยโปรแกรม Mplus โมเดลที่ใช้ในการประมาณคือโมเดลที่ลดรูปลงกล่าวคือเป็นโมเดลที่พารามิเตอร์ในโมเดลการวัดไม่มีความผันแปรระหว่างหน่วย เหตุผลที่ผู้วิจัยมีความจำเป็นต้องวิเคราะห์ด้วยโมเดลที่แตกต่างกันเนื่องจากโปรแกรม Mplus มีข้อจำกัดทางเทคนิคที่ไม่สามารถประมาณค่าพารามิเตอร์ในโมเดลที่พารามิเตอร์ในโมเดลการวัดมีความผันแปรระหว่างกลุ่ม กล่าวคือโมเดลการวัดในกรณีที่ประมาณค่าด้วยโปรแกรม Mplus จะลดรูปเป็นดังสมการที่ (4.69) ดังนี้

$$\underline{y}_{ij} = \underline{\mu} + \Lambda \quad \underline{\omega}_{ij} + \underline{\varepsilon}_{ij} \quad (4.69)$$

เมื่อ $\underline{y}_{ij} = (AFF, ENJ, SEL, SOC, WOR, PHY, DC, CS, SO, OD)^T$,

$\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_{10})^T$,

$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & \lambda_{14} & 1^* & \lambda_{16} & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 0^* & 0^* & \lambda_{21} & \lambda_{22} & \lambda_{23} & \lambda_{24} \end{bmatrix}^T$,

$\underline{\omega}_{ij} = (SWB_w, SIB_w)_{ij}^T$,

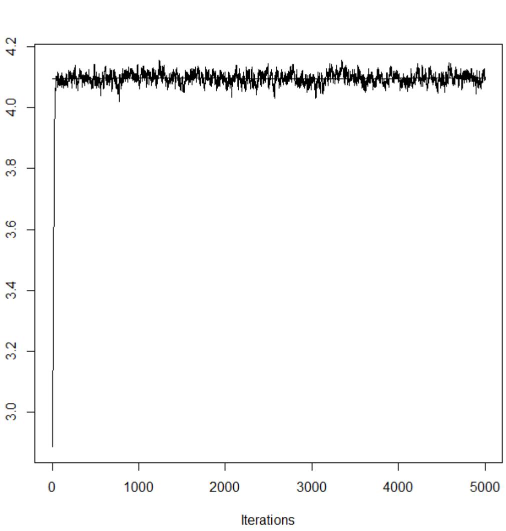
$\underline{\varepsilon}_{ij} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{10})_{ij}^T$ โดยที่ $\varepsilon_{kij} \sim N(\underline{0}, \psi_{\varepsilon k})$

และ ε_{kij} กับ $\varepsilon_{ki'j'}$ เป็นอิสระซึ่งกันและกัน

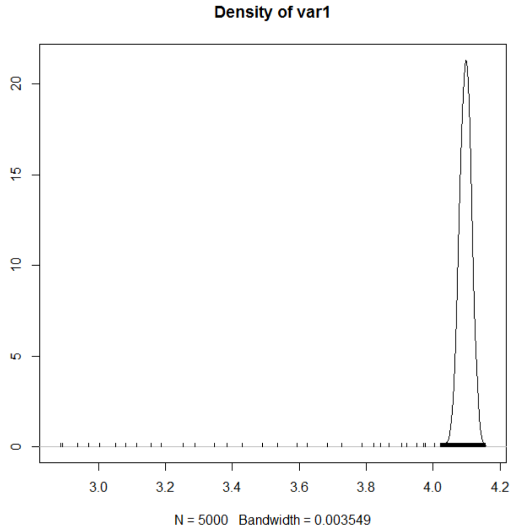
ในการประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณค่าแบบเบส์ ผู้วิจัยเลือกกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าโดยใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบวงศ์คู่สังยุค (noninformative conjugate prior distribution) ที่ไม่ให้สารสนเทศ ดังนี้ การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์จุดตัดแกนในโมเดลการวัดกำหนดให้ $\underline{\mu} \sim N(\underline{0}, 100I)$ การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่มกำหนดให้ $\psi_{\varepsilon kj}^{-1} \sim Gamma(0.001, 0.001)$ การแจกแจงความน่าจะเป็นก่อนหน้าของ

พารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝง $\Phi \sim Wishart(n * J, I)$ การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่ม $\Lambda_{kj} \sim N(0.5,100)$ การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปวนของโมเดลระดับที่หนึ่ง $\sigma_{\delta}^{-2} \sim Gamma(0.001,0.001)$, $\underline{\gamma} \sim N(\underline{0}, 100)$ และการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนในระดับที่สอง $\Sigma_u^{-1} \sim Wishart(J, I)$ โดยที่ I คือเมทริกซ์เอกลักษณ์

ผู้วิจัยใช้วิธีการลูกโซ่มาร์คอฟมอนติคาร์โล (MCMC) ในการประมาณค่าพารามิเตอร์ อัลกอริทึมที่ใช้ในการประมาณคือ อัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ (Gibbs-sampling algorithm) วิธีการดังกล่าวเป็นวิธีการแบบทวนซ้ำซึ่งจะทำการสุ่มตัวอย่างพารามิเตอร์จากการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ในโมเดล ค่าประมาณพารามิเตอร์ที่ได้คำนวณจากตัวอย่างของพารามิเตอร์ที่สุ่มจากวิธีการข้างต้นจำนวน 5000 ตัวอย่าง ผู้วิจัยได้ทำการตรวจสอบคุณสมบัติลู่เข้าสู่การแจกแจงสถานะคงที่ (stationary distribution) ซึ่งในที่นี้คือการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ในโมเดลที่ 4.69 ถึง 4.70 โดยใช้เครื่องมือตรวจสอบจาก package CODA ในโปรแกรม R (Plummer, Best, Cowles, Vines, Sarkar and Almond, 2012) ประกอบไปด้วยการพิจารณา trace plot กับ Cumulative quantile plot และตัวสถิติ Geweke กับ Heidelberger-Welch เพื่อกำหนดจำนวน burn-in และตรวจสอบคุณสมบัติการลู่เข้าสู่การแจงแจกความน่าจะเป็นภายหลังของพารามิเตอร์ ผลการตรวจสอบพบว่าหลังจากกระทำการตัดตัวอย่างของพารามิเตอร์ในส่วนแรกจำนวน 1000 ตัวอย่าง ไม่มีหลักฐานยืนยันว่าลูกโซ่มาร์คอฟที่สร้างขึ้นไม่มีคุณสมบัติลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลัง ผู้วิจัยจึงกำหนดจำนวน burn-in จำนวน 1000 ตัวอย่าง และใช้ตัวอย่างของพารามิเตอร์ที่เหลือจำนวน 4000 ตัวอย่างในการวิเคราะห์ รูปที่ 4.12 ถึง 4.18 แสดงตัวอย่างของ trace plot และ marginal posterior distribution plot ของพารามิเตอร์ในโมเดลบางส่วนซึ่งจะเห็นว่าตัวอย่างของพารามิเตอร์ที่สร้างขึ้นจากวิธีการลูกโซ่มาร์คอฟมอรติคาร์โลลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังตามที่ต้องการ

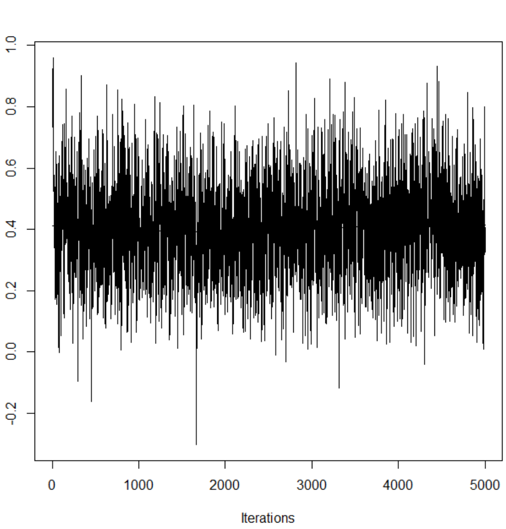


ก. trace plot

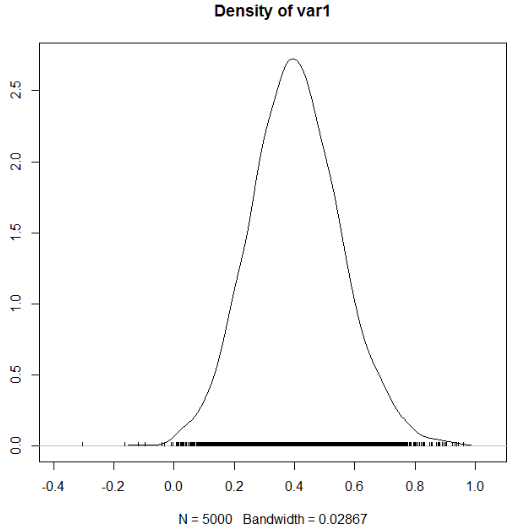


ข. marginal posterior density plot

รูปที่ 4.12 trace plot และ marginal posterior density plot ของพารามิเตอร์ μ_1

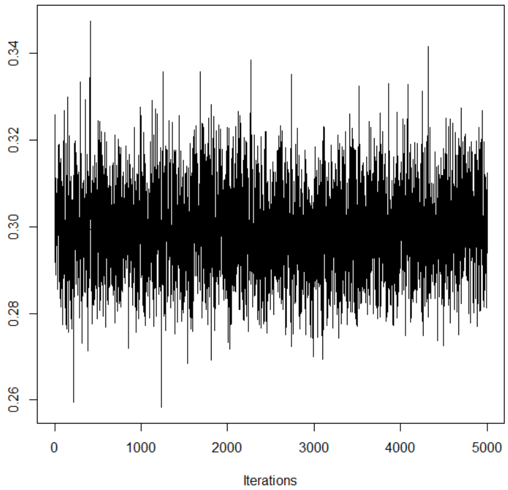


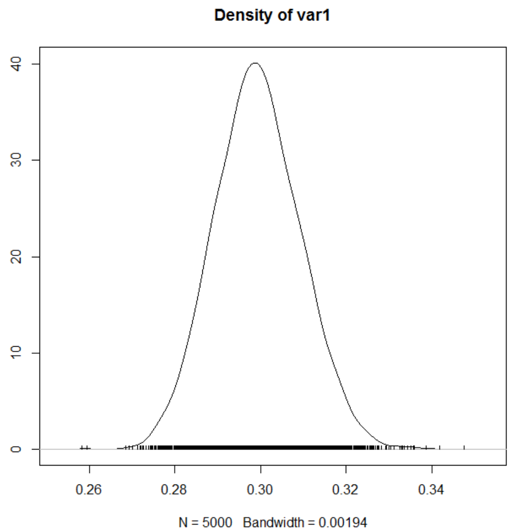
ก. trace plot



ข. marginal posterior density plot

รูปที่ 4.13 trace plot และ marginal posterior density plot ของพารามิเตอร์ $\lambda_{11(j=1)}$

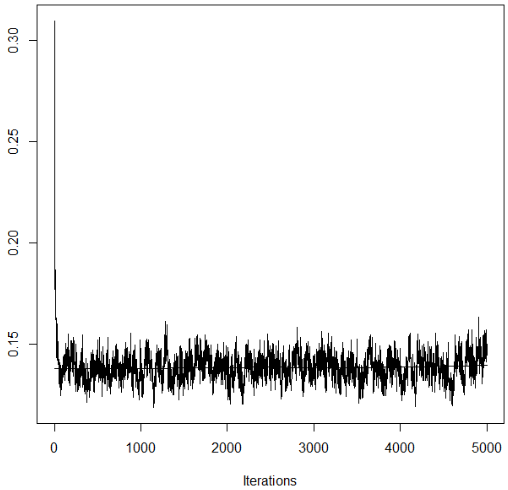


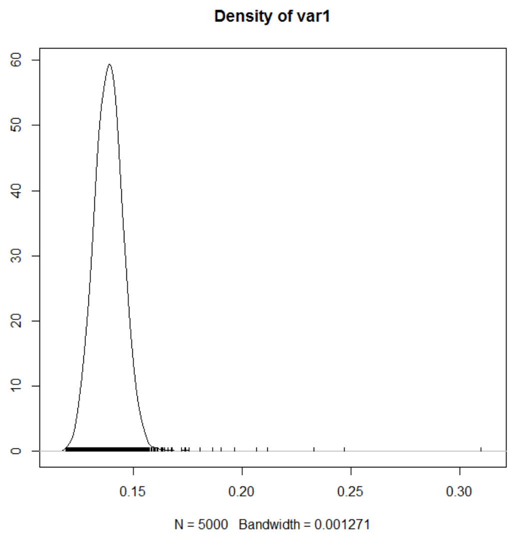


ก. trace plot

ข. marginal posterior density plot

รูปที่ 4.14 trace plot และ marginal posterior density plot ของพารามิเตอร์ $\psi_{\varepsilon 1(j=1)}$

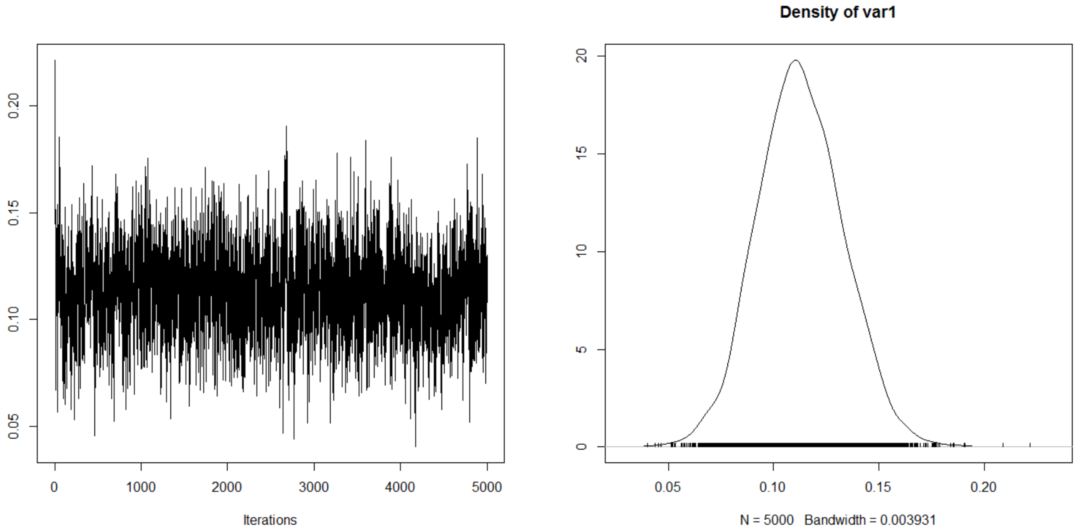




ก. trace plot

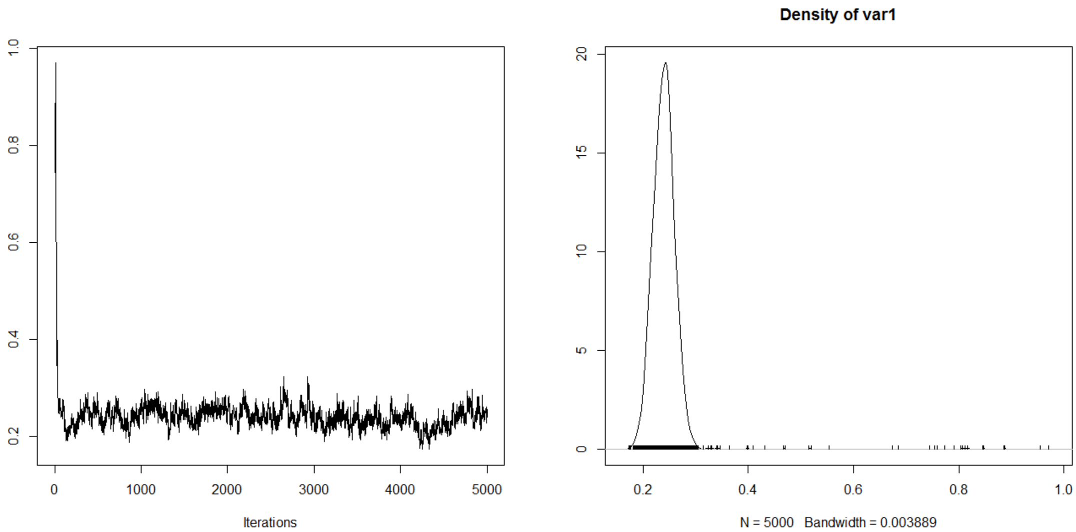
ข. marginal posterior density plot

รูปที่ 4.15 trace plot และ marginal posterior density plot ของพารามิเตอร์ Φ



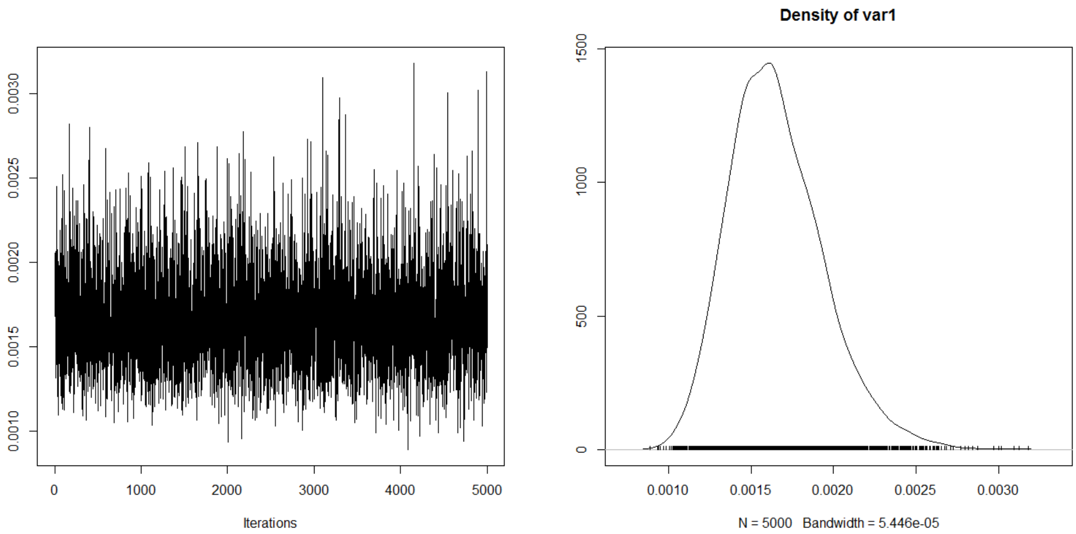
ก. trace plot ข. marginal posterior density plot

รูปที่ 4.16 trace plot และ marginal posterior density plot ของพารามิเตอร์ γ_{01}



ก. trace plot ข. marginal posterior density plot

รูปที่ 4.17 trace plot และ marginal posterior density plot ของพารามิเตอร์ σ_{δ}^{2}



ก. trace plot ข. marginal posterior density plot

รูปที่ 4.18 trace plot และ marginal posterior density plot ของพารามิเตอร์ τ_{11}

ตารางที่ 4.2 แสดงค่าประมาณพารามิเตอร์ ส่วนเบี่ยงเบนมาตรฐานภายหลัง (มีเฉพาะวิธีเบส์) และค่าคลาดเคลื่อนมาตรฐาน ที่ได้จากวิธีเบส์และโปรแกรม Mplus ค่าประมาณพารามิเตอร์ที่ได้จากวิธีเบส์จะหาจากค่าเฉลี่ยภายหลัง (posterior mean) ของการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ และเนื่องจากพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ประมาณด้วยวิธีเบส์มีจำนวนมาก ผู้วิจัยจึงนำเสนอเป็นค่าเฉลี่ยของทุกกลุ่มเพื่อให้ง่ายต่อการพิจารณา รายละเอียดของผลการวิเคราะห์มีดังนี้

พารามิเตอร์จุดตัดแกน (measurement intercepts)

เมื่อพิจารณาค่าประมาณพารามิเตอร์จุดตัดแกนของแต่ละตัวแปรสังเกตได้ทั้ง 6 ตัวของตัวแปรแฝงสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) พบว่าวิธีการประมาณค่าแบบเบส์และโปรแกรม Mplus ให้ค่าประมาณที่มีค่าอยู่ในช่วง [3.168,4.046] และ [3.088,4.045] ตามลำดับ และค่าประมาณพารามิเตอร์จุดตัดแกนของแต่ละตัวแปรสังเกตได้ทั้ง 4 ตัวของตัวแปรแฝงพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) พบว่าวิธีการประมาณค่าแบบเบส์และโปรแกรม Mplus ให้ค่าประมาณที่มีค่าอยู่ในช่วง [3.217,3.971] และ

[3.226,3.959] ตามลำดับ จากผลการประมาณค่าดังกล่าวจะเห็นว่าวิธีการประมาณค่าพารามิเตอร์ทั้งสองให้ค่าประมาณพารามิเตอร์จุดตัดแกนที่ใกล้เคียงกัน และจากการทดสอบสมมติฐานพบว่าทั้งสองวิธียังให้ผลการทดสอบที่สอดคล้องกันคือกล่าวคือมีนัยสำคัญเชิงสถิติที่ระดับ 0.01 ทุกตัว

พารามิเตอร์น้ำหนักองค์ประกอบ

จะพิจารณาค่าประมาณของน้ำหนักองค์ประกอบที่ได้จากวิธีการประมาณค่าแบบเบส์และโปรแกรม Mplus แยกตามตัวแปรแฝงดังต่อไปนี้

<u>*ตัวแปรสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w)*</u>

เมื่อพิจารณาค่าเฉลี่ยของค่าประมาณของพารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่มของแต่ละตัวแปรสังเกตได้ทั้ง 6 ตัวของตัวแปรแฝงสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) ที่ได้จากวิธีการประมาณค่าแบบเบส์พบว่ามีค่าอยู่ในช่วง [0.375,1.000] และจากการใช้ช่วงความน่าเชื่อถือ 99% (99% posterior credible interval) เป็นเครื่องมือในการทดสอบสมมติฐานซึ่งพบว่า มีนัยสำคัญเชิงสถิติที่ระดับ 0.01 ทุกตัว ตัวแปรที่มีค่าน้ำหนักองค์ประกอบมากที่สุดคือความวิตกกังวล ($\bar{\lambda}_{15} = 1.000$) รองลงมาคือ ปัญหาทางสังคม ($\bar{\lambda}_{14} = 0.816$) เจตคติและอารมณ์เชิงบวก ($\bar{\lambda}_{11} = 0.600$) ความเพลิดเพลิน ($\bar{\lambda}_{12} = 0.472$) อัตมโนทัศน์เชิงวิชาการ ($\bar{\lambda}_{13} = 0.394$) และปัญหาสุขภาพกาย ($\bar{\lambda}_{16} = 0.375$) ตามลำดับ

เมื่อพิจารณาค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบของแต่ละตัวแปรสังเกตได้ทั้ง 6 ตัวของตัวแปรแฝงตัวแปรแฝงสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) ที่ได้จากวิธีการประมาณค่าด้วยโปรแกรม Mplus พบว่ามีนัยสำคัญเชิงสถิติที่ระดับ 0.01 เช่นเดียวกับวิธีการประมาณค่าแบบเบส์ แต่ค่าประมาณที่ได้และลำดับของความสำคัญของน้ำหนักองค์ประกอบมีความแตกต่างกัน โดยมีค่าอยู่ในช่วง [0.234,1.000] ตัวแปรที่มีน้ำหนักองค์ประกอบมากที่สุดคือความวิตกกังวล ($\lambda_{15} = 1.000$) รองลงมาคือ ปัญหาทางสังคม ($\lambda_{14} = 0.812$) เจตคติและอารมณ์เชิงบวก ($\lambda_{11} = 0.402$) ปัญหาสุขภาพกาย ($\lambda_{16} = 0.384$) ความเพลิดเพลิน ($\lambda_{12} = 0.291$) และอัตมโนทัศน์เชิงวิชาการ ($\lambda_{13} = 0.234$) ตามลำดับ

จากผลการวิเคราะห์ในข้างต้นจะเห็นว่าค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบในจากวิธีการทั้งสองได้แก่ ค่าประมาณน้ำหนักองค์ประกอบ λ_{11}, λ_{12} และ λ_{13} มีค่าที่ค่อนข้างแตกต่างกัน และทำให้ลำดับของความสำคัญของน้ำหนักองค์ประกอบมีความแตกต่างกันในบางส่วน

ตัวแปรพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w)

เมื่อพิจารณาค่าเฉลี่ยของค่าประมาณของพารามิเตอร์น้ำหนักองค์ประกอบของแต่ละตัวแปรสังเกตได้ทั้ง 4 ตัวของตัวแปรแฝงพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) พบว่ามีค่าเฉลี่ยค่าประมาณน้ำหนักองค์ประกอบมีค่าอยู่ในช่วง [0.754,1.078] และจากการใช้ช่วงความน่าเชื่อถือ 99% (99% posterior credible interval) เป็นเครื่องมือในการทดสอบสมมติฐานของค่าเฉลี่ยค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบ พบว่า มีนัยสำคัญทางสถิติที่ระดับ 0.01 ทุกตัว ตัวแปรที่มีค่าน้ำหนักองค์ประกอบมากที่สุดคือการใช้อำนาจ-การร่วมมือ ($\bar{\lambda}_{21} = 1.078$) รองลงมาได้แก่ การร่วมมือ-การคล้อยตาม ($\bar{\lambda}_{22} = 1.000$) การคล้อยตาม-การต่อต้าน ($\bar{\lambda}_{23} = 0.778$) และการต่อต้าน-การใช้อำนาจ ($\bar{\lambda}_{24} = 0.754$) ตามลำดับ

เมื่อพิจารณาค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบของแต่ละตัวแปรสังเกตได้ทั้ง 4 ตัวของตัวแปรแฝงพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) ที่ได้จากวิธีการประมาณค่าด้วยโปรแกรม Mplus พบว่ามีนัยสำคัญเชิงสถิติที่ระดับ 0.01 เช่นเดียวกับวิธีการประมาณค่าแบบเบส์ แต่ค่าประมาณที่ได้มีความแตกต่างกับ โดยมีค่าอยู่ในช่วง [0.805,1.000] ตัวแปรที่มีน้ำหนักองค์ประกอบมากที่สุดคือ การร่วมมือ-การคล้อยตาม ($\lambda_{22} = 1.000$) รองลงมาคือ การใช้อำนาจ-การร่วมมือ ($\lambda_{21} = 0.954$) การคล้อยตาม-การต่อต้าน ($\lambda_{23} = 0.851$) และการต่อต้าน-การใช้อำนาจ ($\lambda_{24} = 0.805$) ตามลำดับ

จากผลการวิเคราะห์ในข้างต้นจะเห็นว่าค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบในจากวิธีการทั้งสองมีขนาดความสำคัญที่ไม่แตกต่างกันมากนักแต่ลำดับความสำคัญของน้ำหนักองค์ประกอบมีความแตกต่างกันกล่าวคือค่าประมาณพารามิเตอร์ λ_{21} ที่ได้จากวิธีการประมาณค่าแบบเบส์จะมีขนาดอยู่ในลำดับที่หนึ่ง ในขณะที่วิธีการประมาณค่าด้วยโปรแกรม Mplus จะมีขนาดอยู่ในลำดับที่สอง

พารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด

เมื่อพิจารณาค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดแต่ละตัวแปรสังเกตได้ทั้ง 6 ตัวของตัวแปรแฝงสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) พบว่าส่วนใหญ่วิธีการประมาณค่าแบบเบส์และโปรแกรม Mplus ให้ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ใกล้เคียงกัน โดยมีค่าอยู่ในช่วง [0.131,0.350] และ [0.127,0.280] ตามลำดับ ยกเว้น $\psi_{\varepsilon 4}$ (Bayes=0.206, Mplus=0.127) และ $\psi_{\varepsilon 5}$ (Bayes=0.350, Mplus=0.198)

เมื่อพิจารณาค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดของแต่ละตัวแปรสังเกตได้ทั้ง 4 ตัวของตัวแปรแฝงพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) พบว่าวิธีการประมาณค่าแบบเบส์และโปรแกรม Mplus ให้ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ใกล้เคียงกัน โดยมีค่าอยู่ในช่วง [0.029,0.144] และ [0.009,0.176] ตามลำดับ ยกเว้น $\psi_{\varepsilon 4}$ (Bayes=0.144, Mplus=0.176), $\psi_{\varepsilon 9}$ (Bayes=0.029, Mplus=0.009) และ $\psi_{\varepsilon 10}$ (Bayes=0.031, Mplus=0.015)

พารามิเตอร์ความแปรปรวนของตัวแปรอิสระแฝง

เมื่อพิจารณาค่าประมาณพารามิเตอร์ความแปรปรวนของตัวแปรพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน (SIB_w) พบว่าวิธีการประมาณค่าแบบเบส์มีค่าประมาณความแปรปรวนของพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียนที่ใกล้เคียงกันโดยมีค่าเท่ากับ 0.138 และ 0.137 ตามลำดับ

พารามิเตอร์อิทธิพลคงที่

เมื่อพิจารณาค่าประมาณของอิทธิพลคงที่ของผลสัมฤทธิ์ทางการเรียนที่ส่งผลต่อสุขภาวะของนักเรียน (γ_{01}) พบว่าวิธีการทั้งสองมีค่าประมาณที่ใกล้เคียงกัน (Bayes = 0.113, Mplus = 0.138) และมีนัยสำคัญเชิงสถิติที่ระดับ 0.01 ทำให้สามารถได้ข้อสรุปที่สอดคล้องกันว่า นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนที่ดี จะมีแนวโน้มที่มีสุขภาวะที่ดี ซึ่งมีความสมเหตุสมผล

เมื่อพิจารณาค่าประมาณพารามิเตอร์อิทธิพลคงที่ของพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน ที่ส่งผลต่อสุขภาวะของนักเรียน *(γ_{02})* พบว่าวิธีการทั้งสองมีค่าประมาณที่แตกต่างกันอย่างเห็นได้ชัด (Bayes = 0.089, Mplus = -0.025) นอกจากนี้ข้อสรุปของผลการทดสอบสมมติฐานของวิธีการประมาณค่าแบบเบส์พบว่ามีนัยสำคัญเชิงสถิติที่ระดับ 0.01 ในขณะที่ผลการทดสอบสมมติฐานที่ได้จากโปรแกรม Mplus พบว่าไม่มีนัยสำคัญเชิงสถิติ

พารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 2

เมื่อพิจารณาค่าประมาณความแปรปรวนของสัมประสิทธิ์ความถดถอยแบบสุ่มของผลสัมฤทธิ์ทางการเรียน (ACH_W) ที่ส่งผลต่อสุขภาวะของนักเรียน (SWB_w) พบว่าวิธีการทั้งสองให้ค่าประมาณที่ใกล้เคียงกัน (Bayes = 0.002, Mplus = 0.005) และมีนัยสำคัญเชิงสถิติที่ระดับ 0.01 จากผลการวิเคราะห์ในข้างต้นทำให้สามารถสรุปได้ว่าวิธีการทั้งสองได้ผลสรุปที่สอดคล้องกัน กล่าวคือ มีความผันแปรระหว่างกลุ่มอย่างมีนัยสำคัญเชิงสถิติของพารามิเตอร์สัมประสิทธิ์ความถดถอยของผลสัมฤทธิ์ทางการเรียนที่ส่งผลต่อสุขภาวะของนักเรียน

พารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1

เมื่อพิจารณาค่าประมาณพารามิเตอร์ความแปรปรวนของโมเดลในระดับที่ 1 พบว่าวิธีการทั้งสองให้ค่าประมาณที่แตกต่างกันเล็กน้อย (Bayes = 0.242, Mplus = 0.295) และมีนัยสำคัญทางสถิติที่ระดับ 0.01

เปรียบเทียบการประมาณค่าแบบช่วง

เมื่อพิจารณาผลการประมาณค่าแบบช่วงด้วยวิธีเบส์ (99% posterior interval) และวิธีดั้งเดิม (99% confidence interval) พบว่ามีความแตกต่างของผลการวิเคราะห์ที่ได้ซึ่งสามารถแบ่งออกได้เป็น 2 ประเด็นดังนี้ ประเด็นแรก คือประเด็นเกี่ยวกับประสิทธิภาพของค่าประมาณแบบช่วงที่ได้ ซึ่งจะพบว่าเมื่อพิจารณาจากค่าอัตราส่วนของความยาวช่วงของวิธีดั้งเดิมต่อวิธีเบส์ (width ratio) จะพบว่าการประมาณค่าส่วนใหญ่ วิธีการประมาณค่าแบบเบส์จะให้ช่วงของการประมาณที่แคบกว่าวิธีดั้งเดิมค่อนข้างมาก สาเหตุเนื่องจากการหาค่าประมาณแบบช่วงของวิธีเบส์จะหาจากการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ซึ่งประมาณจากตัวอย่างของ

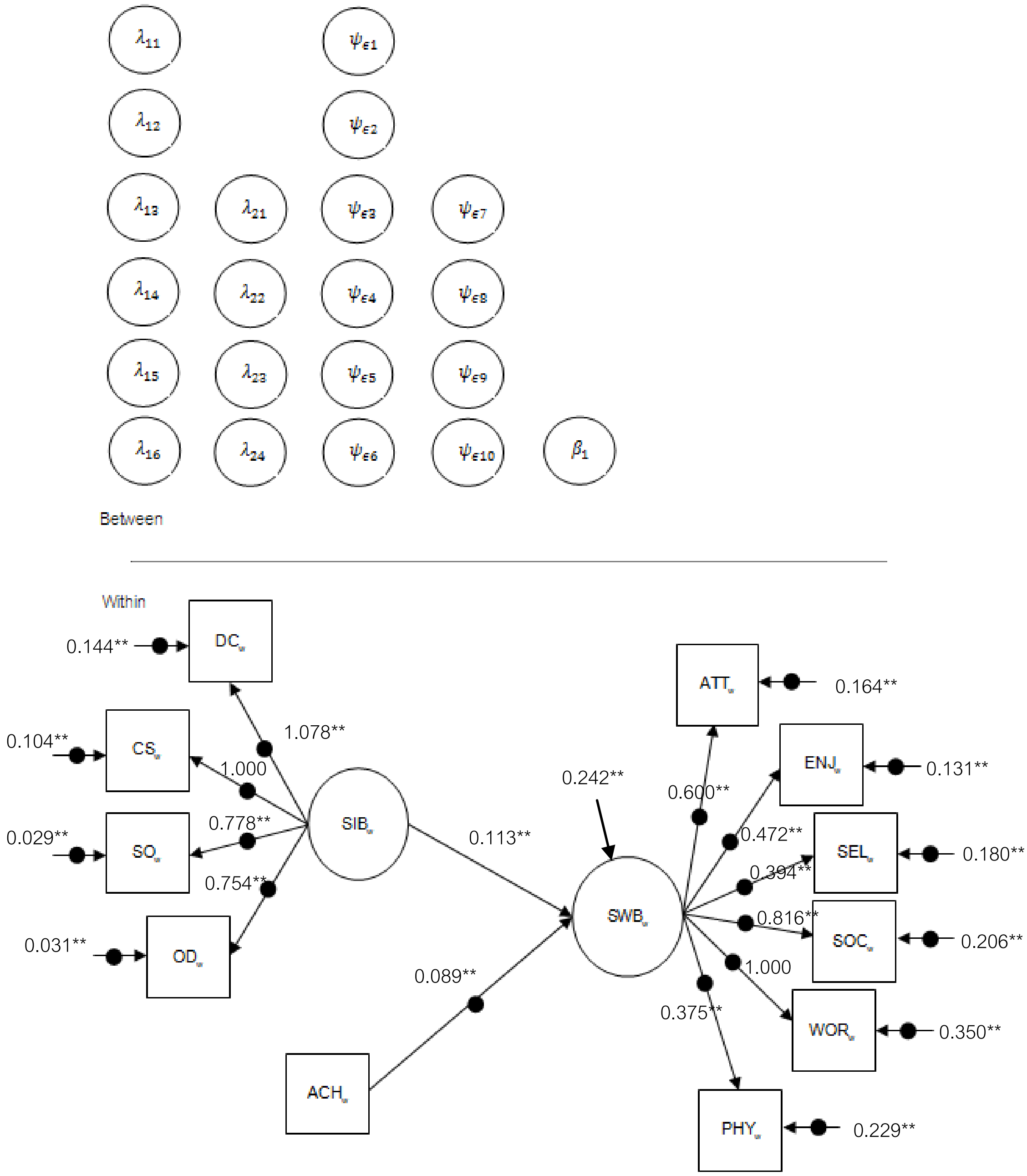
พารามิเตอร์ที่สุ่มจากอัลกอริทึมการประมาณค่าโดยตรง ส่วนวิธีการประมาณค่าแบบช่วงแบบดั้งเดิมจะอาศัยทฤษฎีการลู่เข้าเชิงกำกับของความน่าจะเป็น ซึ่งตัวสถิติที่นำมาสร้างช่วงความเชื่อมั่นจะลู่เข้าสู่การแจกแจงตามทฤษฎีก็ต่อเมื่อมีขนาดตัวอย่างใหญ่เพียงพอ วิธีการประมาณค่าแบบช่วงแบบเบส์จึงมีความสอดคล้องกับข้อมูลมากกว่าทำให้ได้ช่วงการประมาณที่แคบกว่า การนำไปใช้งานในทางปฏิบัติจึงมีประสิทธิภาพมากกว่า ประเด็นที่สอง คือประเด็นเกี่ยวกับการแปลความหมายของค่าประมาณแบบช่วงซึ่งมีความแตกต่างกันอย่างมาก ในกรณีนี้การแปลผลจาก 99% ของช่วงการประมาณแบบเบส์นั้น นักวิจัยหรือนักสถิติสามารถอ้างหรือยืนยันได้ว่าความน่าจะเป็นที่ค่าพารามิเตอร์ที่สนใจจะอยู่ในช่วงความน่าเชื่อถือที่ได้สร้างขึ้นมานั้นมีค่าเท่ากับ 99% ซึ่งเป็นระดับของความน่าเชื่อถือที่ได้กำหนดไว้ จะเห็นว่าการแปลความหมายของช่วงความน่าเชื่อถือดังกล่าวเป็นการแปลความหมายไปที่ตัวพารามิเตอร์ที่สนใจโดยตรง ที่ทำเช่นนี้ได้เพราะสถิติแบบเบส์นั้นสร้างโมเดลความน่าจะเป็นเพื่ออธิบายพารามิเตอร์ที่สนใจโดยตรง ในขณะที่เมื่อเปรียบเทียบกับการแปลความหมายของ 99% ช่วงการประมาณแบบดั้งเดิมหรือความเชื่อมั่นนั้นเป็นการยืนยันในเชิงความถี่สัมพัทธ์ของการสุ่มตัวอย่างว่าช่วงความเชื่อมั่นที่สร้างขึ้นมานั้นจะครอบคลุมค่าพารามิเตอร์ที่สนใจด้วยความน่าจะเป็นเท่ากับ 99% จะเห็นว่าการแปลความหมายของช่วงการประมาณแบบเบส์มีความง่ายและเป็นธรรมชาติมากกว่าช่วงความเชื่อมั่นแบบดั้งเดิม

ตารางที่ 4.2 ค่าประมาณพารามิเตอร์ ค่าส่วนเบี่ยงเบนมาตรฐานภายหลังและค่าคลาดเคลื่อนมาตรฐานที่ได้จากวิธีประมาณค่าแบบเบส์และโปรแกรม Mplus

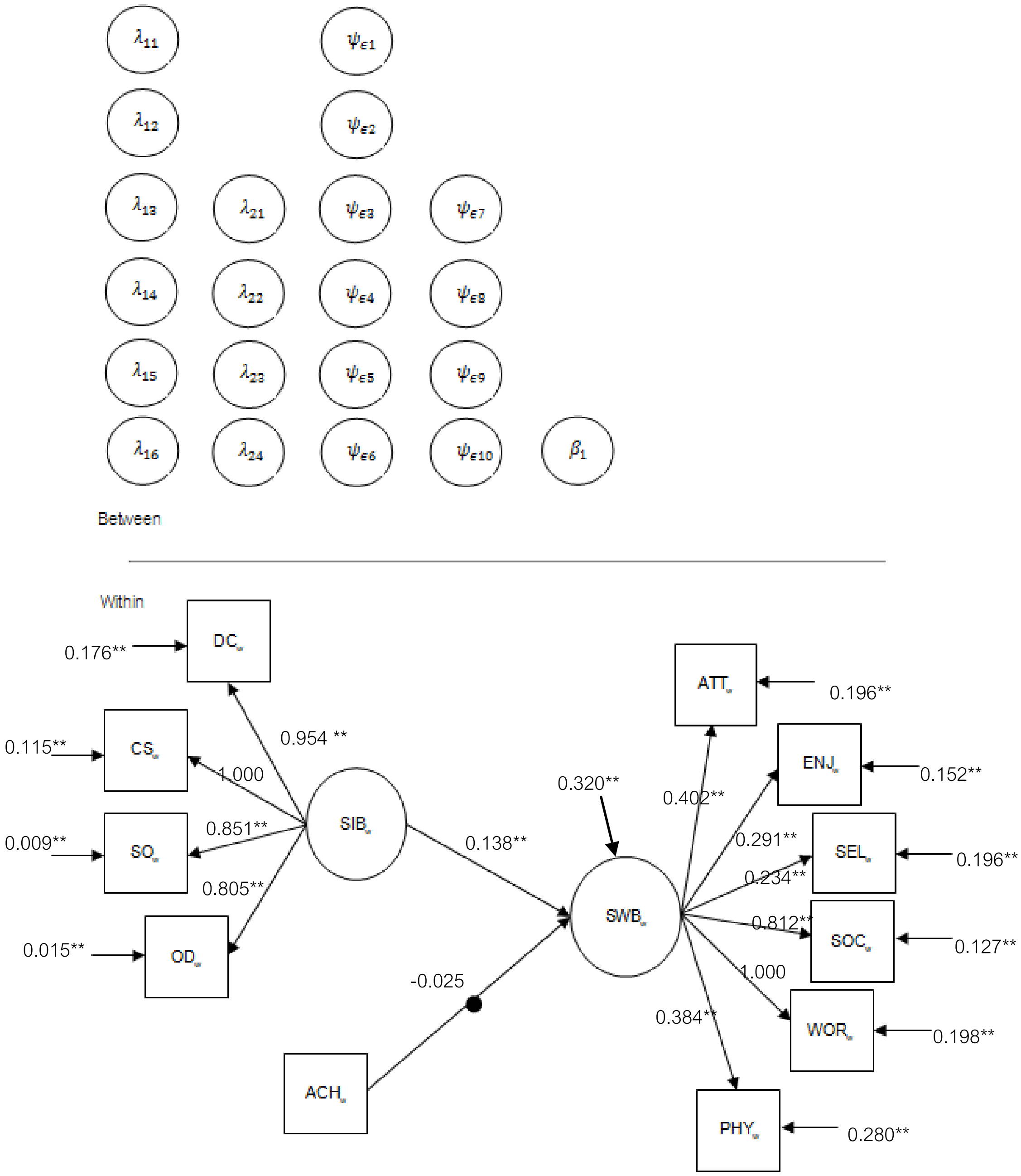
พารามิเตอร์	Bayes					Mplus				Width ratio
	ค่าประมาณ	Posterior SD	Time-series SE	99% posterior interval		ค่าประมาณ	SE	99% confidence interval		
				Lower	Upper			Lower	Upper	
μ_1	4.046**	0.018	0.0036	3.996	4.089	4.045**	0.046	3.938	4.152	2.301
μ_2	3.856**	0.015	0.0028	3.814	3.892	3.877**	0.037	3.791	3.963	2.205
μ_3	3.461**	0.015	0.0024	3.419	3.493	3.492**	0.031	3.420	3.564	1.946
μ_4	3.168**	0.024	0.0028	3.097	3.224	3.088**	0.067	2.932	3.244	2.457
μ_5	3.440**	0.029	0.0045	3.340	3.506	3.316**	0.079	3.132	3.500	2.217
μ_6	3.643**	0.018	0.0017	3.595	3.684	3.593**	0.045	3.488	3.698	2.360
μ_7	3.971**	0.012	0.0007	3.943	3.997	3.959**	0.027	3.896	4.022	2.333
μ_8	3.542**	0.009	0.0006	3.521	3.564	3.593**	0.023	3.480	3.588	2.512
μ_9	3.217**	0.007	0.0004	3.200	3.233	3.226**	0.013	3.196	3.256	1.818
μ_{10}	3.294**	0.007	0.0004	3.279	3.309	3.303**	0.012	3.275	3.331	1.867
$\bar{\lambda}_{11}$	0.600**	0.168	0.0035	0.520	0.690	0.402**	0.041	0.307	0.497	1.118
$\bar{\lambda}_{12}$	0.472**	0.153	0.0034	0.399	0.551	0.291**	0.036	0.207	0.375	1.105
$\bar{\lambda}_{13}$	0.394**	0.164	0.0032	0.318	0.473	0.234**	0.036	0.150	0.318	1.084
$\bar{\lambda}_{14}$	0.816**	0.185	0.0030	0.736	0.901	0.812**	0.025	0.754	0.870	0.703
$\bar{\lambda}_{16}$	0.375**	0.192	0.0019	0.304	0.449	0.384**	0.027	0.321	0.447	0.869
$\bar{\lambda}_{21}$	1.078**	0.210	0.0018	1.015	1.154	0.954**	0.028	0.889	1.019	0.935
$\bar{\lambda}_{23}$	0.778**	0.120	0.0015	0.734	0.827	0.851**	0.032	0.777	0.925	1.591
$\bar{\lambda}_{24}$	0.754**	0.120	0.0015	0.710	0.803	0.805**	0.041	0.710	0.900	2.043
$\bar{\psi}_{\varepsilon k1}$	0.164**	0.053	0.0008	0.142	0.190	0.196**	0.01	0.173	0.219	0.958
$\bar{\psi}_{\varepsilon k2}$	0.131**	0.042	0.0006	0.122	0.155	0.152**	0.008	0.133	0.171	1.152
$\bar{\psi}_{\varepsilon k3}$	0.180**	0.050	0.0005	0.168	0.201	0.196**	0.008	0.177	0.215	1.152
$\bar{\psi}_{\varepsilon k4}$	0.206**	0.069	0.0014	0.170	0.242	0.127**	0.008	0.108	0.146	0.528
$\bar{\psi}_{\varepsilon k5}$	0.350**	0.110	0.0022	0.272	0.377	0.198**	0.014	0.165	0.231	0.629
$\bar{\psi}_{\varepsilon k6}$	0.299**	0.079	0.0004	0.274	0.319	0.280**	0.011	0.254	0.306	1.156
$\bar{\psi}_{\varepsilon k7}$	0.144**	0.045	0.0004	0.128	0.158	0.176**	0.007	0.160	0.192	1.067
$\bar{\psi}_{\varepsilon k8}$	0.104**	0.033	0.0006	0.085	0.120	0.115**	0.008	0.096	0.134	1.086
$\bar{\psi}_{\varepsilon k9}$	0.029**	0.011	0.0003	0.026	0.035	0.009**	0.001	0.006	0.010	0.444
$\bar{\psi}_{\varepsilon k10}$	0.031**	0.113	0.0003	0.028	0.036	0.015**	0.024	0.013	0.017	0.500
Φ	0.138**	0.007	0.0006	0.124	0.155	0.137**	0.013	0.107	0.167	1.935
γ_{01}	0.113**	0.020	0.0007	0.067	0.162	0.138**	0.026	0.078	0.198	1.263
γ_{02}	0.089**	0.033	0.0013	0.014	0.169	-0.025	0.032	-0.099	0.049	0.955
τ_{11}	0.002**	2.92E-4	3.82E-6	0.001	0.002	0.005**	0.001	0.003	0.007	4.000
σ_δ^2	0.242**	0.041	0.0034	0.193	0.295	0.320**	0.024	0.264	0.376	1.098

ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่มและความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่มของที่ได้จากวิธีการประมาณค่าแบบเบส์ ผู้วิจัยได้นำเสนอเป็นค่าเฉลี่ยของทุกกลุ่ม

*** width ratio คืออัตราส่วนของความยาวช่วงของวิธีประมาณค่าแบบเบส์ต่อวิธีดั้งเดิม*



รูปที่ 4.19 ผลการวิเคราะห์ข้อมูลทุติยภูมิด้วยวิธีการประมาณค่าแบบเบส์



รูปที่ 4.20 ผลการวิเคราะห์ข้อมูลทุติยภูมิด้วยวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted

สรุปผลการศึกษาด้วยข้อมูลทุติยภูมิ

จากการเปรียบเทียบผลการวิเคราะห์ในข้างต้นทำให้สามารถสรุปได้ว่า วิธีการประมาณค่าแบบเบส์ซึ่งประมาณค่าภายใต้ข้อสมมติที่พารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวน

ของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่ม และวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted ซึ่งภายใต้ข้อสมมติที่พารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบคงที่ จะให้ค่าประมาณพารามิเตอร์และข้อสรุปที่แตกต่างกันดังต่อไปนี้ 1) ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบและค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดที่มีขนาดแตกต่างกันในบางส่วน โดยส่วนใหญ่วิธีการประมาณค่าแบบเบส์จะให้ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่สูงกว่าและให้ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดที่ต่ำกว่า ซึ่งหมายความว่าวิธีการประมาณค่าแบบเบส์จะให้ค่าประมาณพารามิเตอร์ในส่วนนี้ที่ดีกว่าวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted เล็กน้อย 2) ค่าประมาณพารามิเตอร์อิทธิพลคงที่ของพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน ที่ส่งผลต่อสุขภาวะของนักเรียน (γ_{02}) ซึ่งเป็นความแตกต่างที่เห็นได้ชัดที่สุด กล่าวคือ โดยวิธีการประมาณค่าพารามิเตอร์แบบเบส์จะให้ค่าประมาณที่มีค่าเป็นบวกและมีนัยสำคัญเชิงสถิติซึ่งมีความสมเหตุสมผลและสอดคล้องกับกรอบแนวคิดทฤษฎีเบื้องหลัง ในขณะที่วิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted จะให้ค่าประมาณที่มีค่าติดลบและไม่มีนัยสำคัญเชิงสถิติ 3) ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนในโมเดลระดับที่ 1 ของวิธีการประมาณค่าแบบเบส์มีขนาดเล็กกว่าวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted ซึ่งหมายความว่าวิธีการประมาณค่าพารามิเตอร์แบบเบส์จะให้ความคลาดเคลื่อนที่น้อยกว่า และ 4) เมื่อเปรียบเทียบผลการประมาณค่าแบบช่วงจากวิธีเบส์ (99% posterior interval) และวิธีดั้งเดิม (99% confidence interval) พบว่ามีความแตกต่างของผลการวิเคราะห์ที่ได้ กล่าวคือ วิธีประมาณค่าแบบเบส์จะให้ช่วงการประมาณที่แคบกว่าวิธีดั้งเดิมในส่วนใหญ่ นอกจากนี้จะมีการแปลความหมายที่เป็นธรรมชาติและง่ายกว่าวิธีดั้งเดิม

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่อง “วิธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด: การศึกษาสถานการณ์จำลองแบบมอนติคาร์โลจากข้อมูลจริง” มีวัตถุประสงค์ 3 ประการได้แก่ 1) เพื่อพัฒนาวีธีการประมาณค่าแบบเบส์สำหรับการวิเคราะห์ข้อมูลพหุระดับที่ตัวแปรมีความคลาดเคลื่อนจากการวัด 2) เพื่อศึกษาและตรวจสอบความสามารถของวิธีการประมาณค่าแบบเบส์ที่พัฒนาขึ้นและเปรียบเทียบความสามารถของการประมาณค่าพารามิเตอร์กับวิธีการประมาณที่ใช้ในโปรแกรม Mplus และ 3) เพื่อทดลองใช้วิธีการประมาณค่าแบบเบส์ที่พัฒนาขึ้นเพื่อประมาณค่าพารามิเตอร์โดยใช้ข้อมูลจริง รวมทั้งเปรียบเทียบกับผลการประมาณค่าพารามิเตอร์ที่ได้จากโปรแกรม Mplus

สรุปผลการวิจัย

<u>สรุปผลการวิจัยตามวัตถุประสงค์ข้อที่ 1</u>

ในส่วนของการพัฒนาวิธีการประมาณค่าพารามิเตอร์แบบเบส์เพื่อตอบสนองวัตถุประสงค์การวิจัยในข้อแรกนั้น ผู้วิจัยได้พิจารณาจากธรรมชาติของข้อมูลซึ่งโดยตรรกกะแล้วหากข้อมูลมีความเป็นพหุระดับ พารามิเตอร์ทั้งในโมเดลการวัดและโมเดลโครงสร้างนั้นไม่จำเป็นที่จะต้องมีค่าคงที่ระหว่างกลุ่ม อย่างไรก็ตามการกำหนดให้พารามิเตอร์ทุกตัวในโมเดลมีความผันแปรระหว่างกลุ่มจะทำให้โมเดลดังกล่าวไม่สามารถระบุได้ (unidentified) จากแนวคิดในข้างต้นผู้วิจัยจึงได้กำหนดโมเดลเป้าหมายที่ใช้ในการศึกษาดังนี้

<u>โมเดลการวัด</u> (measurement model)

$$\underline{y}_{ij} = \underline{\mu} + \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij}$$

<u>โมเดลสมการโครงสร้าง</u> (structural model)

$$\eta_{ij} = \beta_j^T \underline{\Xi}_{ij} + \delta_{ij} = \underline{\gamma}^T \underline{\Xi}_{ij} + \underline{u}_j^T \underline{\Xi}_{ij} + \delta_{ij}$$

เมื่อ $\underline{\Xi}_{ij} = \begin{pmatrix} 1 & \underline{\xi}_{ij} \end{pmatrix}^T$ และ $\beta_j = \underline{\gamma} + \underline{u}_j$,

$\underline{\epsilon}_{ij} \sim N(\underline{0}, \Psi_{\epsilon j})$ โดยที่ $\Psi_{\epsilon j} = diag(\psi_{\epsilon kj})$,

$\underline{\omega}_{ij} = \left(\eta_{ij}, \underline{\xi}_{ij}\right)^T \sim N\left(\underline{\theta}_{\omega j}, \Sigma_{\omega j}\right)$

โดยที่ $\underline{\theta}_{\omega j} = \begin{pmatrix} \underline{\beta}_j^T \underline{\nu}_j \\ \underline{\nu}_j \end{pmatrix}$, $\Sigma_{\omega j} = \begin{bmatrix} \underline{\beta}_j \Phi \underline{\beta}_j^T + \sigma_\delta^2 & \underline{\beta}_j^T \Phi \\ \underline{\beta}_j \Phi & \Phi \end{bmatrix}$

และ $\delta_{ij} \sim N\left(\underline{0}, \sigma_\delta^2\right)$

วิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้นผู้วิจัยเลือกการประมาณค่าด้วยวิธีลูกโซ่มาร์คอฟมอนติคาร์โล (Markov Chain Monte Carlo:MCMC) โดยอัลกอริทึมที่ใช้คืออัลกอริทึมการสุ่มตัวอย่างแบบกิบบส์ (Gibbs-sampling algorithm)

วิธีการประมาณค่าแบบเบส์ที่ผู้วิจัยพัฒนาขึ้น ผู้วิจัยได้เลือกใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบวงศ์คู่สังยุค (conjugacy prior distribution) ของแต่ละพารามิเตอร์ในโมเดลดังนี้ 1) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์จุดตัดแกนคือ $\underline{\mu} \sim Normal(\kappa, \Sigma_\mu)$ 2) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่ม คือ $\underline{\Lambda}_{kj} \sim Normal\left(\underline{g}_{kj}, H_{kj}\right)$ โดยที่ $\underline{\Lambda}_{kj}$ คือเวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์น้ำหนักองค์ประกอบ Λ_j ที่เป็นพารามิเตอร์อิสระ (fixed parameter) 3) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความคลาดเคลื่อนจากการวัดแบบสุ่ม คือ $\psi_{\epsilon kj}^{-1} \sim Gamma(\alpha_{kj}, \beta_{kj})$ โดยที่ $\psi_{\epsilon kj} \in \Psi_{\epsilon j}$ 4) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนของโมเดลระดับที่ 1 คือ $\sigma_\delta^{-2} \sim Gamma(\alpha_\delta, \beta_\delta)$ 5) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ค่าเฉลี่ยตัวแปรอิสระแฝงแบบสุ่มคือ $\underline{\nu}_j \sim Normal\left(\underline{0}, \Delta\right)$ โดยที่ Δ เมทริกซ์ความแปรปรวนร่วมก่อนหน้า 6) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมของตัวแปรอิสระแฝงคือ $\Phi^{-1} \sim Wishart\left(\rho_\phi, R_\phi\right)$ 7) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมในของโมเดลระดับที่ 2 คือ $\Sigma_u^{-1} \sim Wishart(\rho_u, R_u)$ 8) การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์อิทธิพลคงที่ $\underline{\gamma} \sim Normal\left(\underline{\gamma}_0, \Sigma_{\gamma 0}\right)$

อัลกอริทึมที่พัฒนาขึ้นเป็นกระบวนการประมาณค่าแบบทวนซ้ำประกอบไปด้วย 10 ขั้นตอนมีรายละเอียดโดยย่อดังนี้

1. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\mu}$ จาก

$$\underline{\mu} \mid \left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_{\mu}, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\} \sim Normal(\hat{\kappa}, \hat{\Sigma}_{\mu})$$

โดยที่ $\hat{\Sigma}_{\mu} = \left[\Sigma_{\mu}^{-1} + \sum_{j=1}^{J} n_j \Psi_{\epsilon j}^{-1}\right]^{-1}$

และ $\hat{\kappa} = \hat{\Sigma}_{\mu}\left[\Sigma_{\mu}^{-1}\kappa + \sum_{j=1}^{J} \Psi_{\epsilon j}^{-1} \sum_{i=1}^{n_j}\left(\underline{y}_{ij} - \Lambda_j \underline{\omega}_{ij}\right)\right]$

2. สำหรับหน่วยที่ j และตัวแปรสังเกตได้ที่ k สุ่มตัวอย่างพารามิเตอร์ $\underline{\Lambda}_{kj}$ จาก

$$\underline{\Lambda}_{kj} \mid \left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj} \sim Normal(\underline{\hat{g}}_{kj}, \hat{H}_{kj})$$

เมื่อ $\underline{\omega}_{-ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ_j และเป็น free parameters, $\underline{\tilde{\omega}}_{ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ_j และเป็น fix parameters, $\underline{\tilde{y}}_{ijk}$ เป็นค่าตัวแปรสังเกตได้ที่ปรับค่าที่เนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบในแถวที่ k พารามิเตอร์เมทริกซ์ Λ_j ซึ่ง $\underline{\tilde{y}}_{ijk} = y_{ijk} - \underline{1}^T \underline{\tilde{\omega}}_{ijk} - \mu_k$

โดยที่ $\hat{H}_{kj} = \left[H_{kj}^{-1} + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\omega}_{-ijk}^T\right]^{-1}$

และ $\underline{\hat{g}}_{kj} = \hat{H}_{kj}\left[H_{kj}^{-1}\underline{g}_{kj} + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\tilde{y}}_{ijk}\right]$

3. สำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j สุ่มตัวอย่างตัวแปรแฝง $\underline{\omega}_{ij}$ จาก

$$\underline{\omega}_{ij} \mid \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}, \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij} \sim Normal(\underline{\hat{\theta}}_{ij}, \hat{\Sigma}_{\omega j})$$

โดยที่ $\hat{\Sigma}_{\omega j} = \left[\Sigma_{\omega j}^{-1} + \Lambda_j^T \Psi_{\epsilon j}^{-1} \Lambda_j\right]^{-1}$

และ $\underline{\hat{\theta}}_{ij} = \hat{\Sigma}_{\omega j}\left[\Sigma_{\omega j}^{-1}\underline{\theta}_j + \Lambda_j^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}\right)\right]$

4. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ค่าเฉลี่ยของตัวแปรแฝง $\underline{\nu}_j$ จาก

$$\underline{\nu}_j \mid \left\{\underline{\xi}_{ij}\right\}, \Phi, \Delta \sim Normal(\underline{\hat{\nu}}_j, \hat{\Delta})$$

โดยที่ $\hat{\Delta} = \left[\Delta^{-1} + n_j \Phi^{-1}\right]^{-1}$ และ $\underline{\hat{\nu}}_j = \hat{\Delta}\left[\Phi^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ij}\right]$

5. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Φ^{-1} จาก

$$\Phi^{-1} \mid \left\{\underline{\xi}_{ij}\right\}, \underline{\nu}_j, \rho_{\phi}, R_{\phi} \sim Wishart(\hat{\rho}_{\phi}, \hat{R}_{\phi})$$

โดยที่ $\hat{\rho}_{\phi} = n + \rho_{\phi}$

และ $\hat{R}_\phi = \left[R_\phi^{-1} + \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij} - \underline{\nu}_j\right)\left(\underline{\xi}_{ij} - \underline{\nu}_j\right)^T\right]^{-1}$

6. สำหรับตัวแปรสังเกตได้ที่ k และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\psi_{\epsilon kj}^{-1}$ จาก

$$\psi_{\epsilon kj}^{-1}|\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \left\{\underline{y}_{ij}\right\}, \alpha_{kj}, \beta_{kj} \sim Gamma\left(\hat{\alpha}_{\epsilon kj}, \hat{\beta}_{\epsilon kj}\right)$$

โดยที่ $\hat{\alpha}_{\epsilon kj} = \frac{n_j}{2} + \alpha_{kj}$

และ $\hat{\beta}_{\epsilon kj} = \beta_{kj} + \frac{1}{2}\sum_{i=1}^{n_j}\left(y_{ij} - \mu_k - \underline{\Lambda}_k^T\underline{\omega}_{ij}\right)^T\left(y_{ij} - \mu_k - \underline{\Lambda}_k^T\underline{\omega}_{ij}\right)$

7. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\gamma}$ จาก

$$\underline{\gamma}|\{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{\gamma}}\ \ , \hat{\Sigma}_\gamma\right)$$

โดยที่ $\hat{\Sigma}_\gamma = \left[\Sigma_{\gamma 0}^{-1} + \frac{1}{\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\underline{\Xi}_{ij}\underline{\Xi}_{ij}^T\right]^{-1}$

และ$\underline{\hat{\gamma}} \quad = \hat{\Sigma}_\gamma\left[\Sigma_{\gamma 0}^{-1}\underline{\gamma} + \frac{1}{\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\underline{\Xi}_{ij}(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{u}_j)\right]$ เมื่อ $\underline{\Xi}_{ij} = \begin{pmatrix}1 & \underline{\xi}_{ij}\end{pmatrix}^T$

8. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{u}_j$ จาก

$$\underline{u}_j|\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{u}}_j, \hat{\Sigma}_{u\mathrm{j}}\right)$$

โดยที่ $\hat{\Sigma}_{u\mathrm{j}} = \left[\frac{1}{\sigma_\delta^2}\sum_{i=1}^{n_j}\underline{\Xi}_{ij}\underline{\Xi}_{ij}^T + \Sigma_u^{-1}\right]^{-1}$

และ $\underline{\hat{u}}_j = \frac{1}{\sigma_\delta^2}\hat{\Sigma}_{u\mathrm{j}}\left[\sum_{i=1}^{n_j}\underline{\Xi}_{ij}^T\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma}\right)\right]$

9. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Σ_u^{-1} จาก

$$\Sigma_u^{-1}|\underline{u}_j, \rho_u, R_u \sim Wishart\left(\hat{\rho}_u, \hat{R}_u\right)$$

โดยที่ $\hat{\rho}_u = J + \rho_u$ และ $\hat{R}_u = \left[R_u^{-1} + \sum_{j=1}^{J}\underline{u}_j^T\underline{u}_j\right]^{-1}$

10. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $(\sigma_\delta^2)^{-1}$ จาก

$$(\sigma_\delta^2)^{-1}|\{\underline{\omega}_{ij}\}, \left\{\underline{\beta}_j\right\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta \sim Gamma\left(\hat{\alpha}_\delta, \hat{\beta}_\delta\right)$$

โดยที่ $\hat{\alpha}_\delta = \frac{\sum_{j=1}^{J} n_j}{2} + \alpha_\delta$ และ $\hat{\beta}_\delta = \beta_\delta + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\Xi}_{ij}^T\underline{\gamma} - \underline{\Xi}_{ij}^T\underline{u}_j\right)^2$

กระบวนการประมาณค่าที่สร้างขึ้นนี้มีความจำเป็นที่ต้องการจำนวนรอบของการทวนซ้ำที่มากพอเพื่อให้ตัวอย่างของพารามิเตอร์ในโมเดลที่สร้างขึ้นลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ที่ต้องการ โดยผู้วิจัยสามารถใช้ตัวอย่างของพารามิเตอร์ที่ลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังดังกล่าวเพื่อสรุปเป็นค่าประมาณและเพื่ออนุมานเกี่ยวกับพารามิเตอร์ในโมเดล

สรุปผลการวิจัยตามวัตถุประสงค์ข้อที่ 2

การศึกษานี้มีวัตถุประสงค์เพื่อตรวจสอบความสามารถในการประมาณค่าพารามิเตอร์โดยวิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่พัฒนาขึ้น และเพื่อเปรียบเทียบความสามารถระหว่างวิธีการประมาณค่าพารามิเตอร์แบบเบส์ที่ประมาณค่าภายใต้โมเดลการวัดที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่มกับวิธีการประมาณค่าจากโปรแกรม Mplus ที่ประมาณค่าภายใต้โมเดลการวัดที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์คงที่

โมเดลประชากรที่ใช้ในการสร้างข้อมูลเพื่อทำการศึกษาประกอบไปด้วยโมเดลการวัดและโมเดลสมการโครงสร้างซึ่งมีรายละเอียดดังนี้

<u>โมเดลการวัด (measurement model)</u>

$$\underline{y}_{ij} = \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij}$$

<u>โมเดลสมการโครงสร้าง (structural model)</u>

$$\eta_{ij} = \underline{\beta}_j^T \underline{\xi}_{ij} + \delta_{ij} = \underline{\gamma}^T \underline{\xi}_{ij} + \underline{u}_j \underline{\xi}_{ij} + \delta_{ij}$$

ปัจจัยที่ใช้ในการกำหนดสถานการณ์จำลองให้มีความแตกต่างกันประกอบไปด้วย 2 ปัจจัยได้แก่ 1) ปัจจัยระดับค่าเฉลี่ยความเที่ยงในการวัดตัวแปรแฝง ($\overline{\rho}^2$) กำหนดระดับที่ใช้ในการศึกษาจำนวน 4 ระดับจากต่ำไปสูงได้แก่ 0.3, 0.5, 0.7 และ 0.9 ตามลำดับ 2) ปัจจัยขนาดหน่วย กำหนดระดับที่ใช้ในการศึกษาไว้ 3 ระดับได้แก่ 15, 30 และ 50 หน่วย (J) จากการกำหนดในข้างต้นคิดเป็นสถานการณ์จำลองที่ใช้ในการศึกษาทั้งหมด 4x3=12 สถานการณ์ การประเมินประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ในโมเดลจะพิจารณาแยกตามกลุ่มพารามิเตอร์ซึ่งแบ่งออกเป็น 6 เซตของพารามิเตอร์ได้แก่ เซตของพารามิเตอร์น้ำหนักองค์ประกอบ เซตของพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัด เซตของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมของตัวแปรแฝง เซตของพารามิเตอร์อิทธิพลคงที่ เซตของพารามิเตอร์ความแปรปรวนในระดับที่หนึ่ง และเซตของพารามิเตอร์ความแปรปรวนและความแปรปรวนร่วมในระดับที่สอง เกณฑ์ที่ใช้ในการพิจารณาคือค่าคลาดเคลื่อนกำลังสอง (mean square error: MSE) ซึ่งเป็นค่าที่ใช้วัดความแม่นยำและประสิทธิภาพในภาพรวม

ผลการวิเคราะห์วิธีการประมาณค่าแบบเบส์และวิธีการประมาณค่าภาวะความควรจะเป็นสูงสุดจากโปรแกรม Mplus จะให้ค่า MSE ที่แตกต่างกัน โดยวิธีการประมาณค่าแบบเบส์จะให้ค่า MSE ของค่าประมาณพารามิเตอร์ในโมเดลต่ำกว่าอย่างเห็นได้ชัดในทุกสถานการณ์จำลอง ผลการศึกษาดังกล่าวทำให้สรุปได้ว่า วิธีการประมาณค่าพารามิเตอร์แบบเบส์สามารถแก้ไขปัญหาการประมาณค่าพารามิเตอร์ในโมเดลสมการโครงสร้างพหุระดับที่โมเดลการวัดเป็นโมเดลที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่มได้ซึ่งในสถานการณ์ที่ศึกษาจะให้ค่าประมาณพารามิเตอร์ในโมเดลที่มีความถูกต้องมากกว่าวิธีการประมาณค่าพารามิเตอร์แบบดั้งเดิมที่ประมาณค่าภายใต้โมเดลการวัดที่มีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดคงที่ นอกจากนี้ยังได้ข้อค้นพบว่าเมื่อปัจจัยค่าเฉลี่ยความเที่ยงรวมและจำนวนหน่วยตัวอย่างในระดับที่ 2 มีขนาดเพิ่มขึ้น วิธีประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted จะมีค่าประมาณที่มีความถูกต้องมากยิ่งขึ้นเมื่อเปรียบเทียบกับค่าประมาณที่ได้จากวิธีการประมาณค่าแบบเบส์

<u>สรุปผลการวิจัยตามวัตถุประสงค์ข้อที่ 3</u>

การศึกษาในส่วนที่ 3 เป็นการทดลองใช้วิธีการประมาณค่าแบบเบส์เพื่อประมาณค่าพารามิเตอร์จากข้อมูลจริง ผู้วิจัยใช้ข้อมูลทุติยภูมิจากวิทยานิพนธ์ระดับดุษฎีบัณฑิต สาขาวิชาวิธีวิทยาวิจัยทางการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เรื่อง “อิทธิพลของพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลและสุขภาวะครูที่มีต่อสุขภาวะของนักเรียน : โมเดลการปรับและการส่งผ่านพหุระดับ” (ถมรัตน์ ศิริภาพ, 2554) ข้อมูลตัวอย่างที่ใช้เป็นข้อมูลของนักเรียนรวมทั้งสิ้น 2706 คน ติดอยู่ในหน่วยครูจำนวน 71 คนจาก 71 ห้องเรียนใน 36 โรงเรียนทั่วทุกภาคของประเทศ โดยผู้วิจัยเลือกใช้ตัวแปรจากงานวิจัยดังกล่าวจำนวน 3 ตัวแปร ได้แก่ ตัวแปรสุขภาวะของนักเรียนในระดับนักเรียน (SWB_w) วัดได้จากตัวแปรสังเกตได้จำนวน 6 ตัวแปรได้แก่ เจตคติและอารมณ์เชิงบวก (AFF) ความเพลิดเพลิน (ENJ) อัตมโนทัศน์เชิงวิชาการ (SEL) ปัญหาทางสังคม (SOC) ความวิตกกังวล (WOR) และปัญหาสุขภาพกาย (PHY) ตัวแปรพฤติกรรมครูด้านสัมพันธภาพระหว่างบุคคลตามการรับรู้ของนักเรียนในระดับ

นักเรียน (SIB_w) และตัวแปรผลสัมฤทธิ์ทางการเรียนในระดับนักเรียน (ACH_w) วัดได้จากตัวแปรสังเกตได้จำนวน 4 ตัวแปรได้แก่ การใช้อำนาจ-การร่วมมือ (DC) การร่วมมือ-การคล้อยตาม (CS) การคล้อยตาม-การต่อต้าน (SO) และการต่อต้าน-การใช้อำนาจ (OD) และตัวแปรผลสัมฤทธิ์ทางการเรียนพิจารณาจากเกรดเฉลี่ยสะสมของนักเรียน

จากการเปรียบเทียบผลการวิเคราะห์ในข้างต้นทำให้สามารถสรุปได้ว่า วิธีการประมาณค่าแบบเบส์ (ภายใต้ข้อสมมติที่พารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่ม) และวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุด (ภายใต้ข้อสมมติที่พารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบคงที่) จะให้ค่าประมาณพารามิเตอร์และข้อสรุปที่แตกต่างกันดังต่อไปนี้ 1) ค่าประมาณพารามิเตอร์น้ำหนักองค์ประกอบที่มีขนาดและลำดับความสำคัญของน้ำหนักองค์ประกอบที่แตกต่างกันในบางส่วน 2) ค่าประมาณพารามิเตอร์ความแปรปรวนของความคลาดเคลื่อนจากการวัดที่มีขนาดของความแปรปรวนแตกต่างกันในบางส่วน และ 3) ค่าประมาณพารามิเตอร์อิทธิพลคงที่ของพฤติกรรมครูด้านสัมพันธภาพตามการรับรู้ของนักเรียนในระดับนักเรียน ที่ส่งผลต่อสุขภาวะของนักเรียน (γ_{02}) ซึ่งเป็นความแตกต่างที่เห็นได้ชัดที่สุดกล่าวคือ โดยวิธีการประมาณค่าพารามิเตอร์แบบเบส์จะให้ค่าประมาณที่มีค่าเป็นบวกและมีนัยสำคัญเชิงสถิติซึ่งมีความสมเหตุสมผลและสอดคล้องกับกรอบแนวคิดทฤษฎีเบื้องหลัง ในขณะที่วิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดจะให้ค่าประมาณที่มีค่าติดลบและไม่มีนัยสำคัญเชิงสถิติ

อภิปรายผลการวิจัย

ในส่วนของการอภิปรายผลผู้วิจัยแบ่งออกเป็นส่วนของการอภิปรายในส่วนของระเบียบวิธีวิจัยและส่วนของการอภิปรายในส่วนองค์ความรู้ที่ได้

<u>การอภิปรายในส่วนของระเบียบวิธีวิจัย</u>

1) วิธีการประมาณค่าแบบเบส์ที่พัฒนาขึ้นมีข้อจำกัดกล่าวคือสามารถใช้ประมาณภายใต้โมเดลที่มีตัวแปรตามแฝง 1 ตัว และมีตัวแปรอิสระแฝงได้หลายตัวเท่านั้น จึงเป็นข้อจำกัดของงานวิจัยในกรณีที่ผู้วิจัยต้องการวิเคราะห์ตัวแปรตามหลายตัว หรือโมเดลในลักษณะอื่นๆ

อย่างไรก็ตามวิธีการดังกล่าวยังสามารถขยายอัลกอริทึมให้รองรับกับสถานการณ์ที่มีความทั่วไปมากขึ้นได้ได้แก่ กรณีตัวแปรตามหลายตัว กรณีที่มีตัวแปรสังเกตได้ในโมเดลสมการโครงสร้าง กรณีที่พารามิเตอร์จุดตัดแกนของโมเดลการวัดเป็นพารามิเตอร์แบบสุ่ม เป็นต้น ซึ่งรายละเอียดจะกล่าวในข้อเสนอแนะสำหรับการวิจัยในครั้งต่อไป

2) การพิสูจน์เพื่อหารูปแบบของการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขในแต่ละขั้นตอนของอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ ผู้วิจัยมีข้อสมมติให้พารามิเตอร์ในโมเดลการวัดและโมเดลสมการโครงสร้างเป็นอิสระแบบมีเงื่อนไขซึ่งกันและกัน(สมการที่ 4.10) กล่าวคือสมมติให้ $p\left(\Theta_y, \Theta_\omega | \Omega, \left\{\underline{y}_{ij}\right\}\right) \propto \left[p(\Theta_y) p\left(\left\{\underline{y}_{ij}\right\} | \Omega, \Theta_y\right)\right] \times [p(\Theta_\omega) p(\Omega|\Theta_\omega)]$ แนวคิดนี้ผู้วิจัยอ้างอิงจากงานวิจัยที่มีผู้พัฒนาวิธีการประมาณค่าพารามิเตอร์แบบเบส์ในโมเดลสมการโครงสร้างก่อนหน้านี้ (Dunson, Palomo, and Bollen, 2005; Lee and Song, 2004) และนอกจากนี้กำหนดให้พารามิเตอร์ภายในโมเดลการวัดและโมเดลสมการโครงสร้างเป็นอิสระแบบมีเงื่อนไขซึ่งกันและกันอีกด้วย โดยอ้างอิงแนวคิดนี้จากงานวิจัยของ Goldstein และคณะ (2008) และ Browne (1995) อย่างไรก็ตามในธรรมชาติความเป็นจริงพารามิเตอร์ในส่วนต่างๆข้างต้นอาจไม่จำเป็นต้องเป็นอิสระซึ่งกันและกัน การพิสูจน์รูปแบบของการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขในกรณีที่ยอมให้พารามิเตอร์ในส่วนต่างๆไม่เป็นอิสระกันสามารถกระทำได้ แต่จะมีความยุ่งยากซับซ้อนมากขึ้น การใช้เทคนิคการสุ่มตัวอย่างแบบกิบส์แต่เพียงอย่างเดียวอาจไม่เพียงพอที่จะแก้ไขปัญหาดังกล่าว ในบางขั้นตอนอาจต้องใช้อัลกอริทึม Metropolis-Hasting เข้ามาช่วยสำหรับการสุ่มตัวอย่างแบบหลายตัวแปร (multivariate) จากที่กล่าวมาข้างต้นผู้วิจัยมองว่าการพัฒนาวิธีการประมาณค่าพารามิเตอร์ให้รองรับความไม่เป็นอิสระกันระหว่างพารามิเตอร์ในส่วนต่างๆเป็นเรื่องดี เพราะจะทำให้ได้วิธีการที่มีประสิทธิภาพมากยิ่งขึ้น แต่จะมีความยุ่งยากและต้องใช้เวลาในการประมาณค่ามากขึ้นตามไปด้วย

3) การกำหนดข้อสมมติเบื้องต้นให้พารามิเตอร์ในโมเดลเป็นอิสระซึ่งกันและกันอาจมีข้อเสียดังที่ได้กล่าวไว้ในการอภิปรายข้างต้น แต่มีข้อดีประการหนึ่งหากพิจารณาในมุมมองของการประยุกต์ใช้ กล่าวคือการนำอัลกอริทึมดังกล่าวไปใช้งานเพื่อประมาณค่าพารามิเตอร์ภายใต้โมเดลที่เป็นโมเดลย่อยของโมเดลที่ใช้ในการวิจัยนี้ ผู้วิจัยสามารถตัดเฉพาะขั้นตอนที่จะใช้งานไปได้โดยที่ไม่จำเป็นต้องคำนึงถึงความสัมพันธ์กันระหว่างพารามิเตอร์ เนื่องได้จากสมมติให้เป็น

อิสระซึ่งกันและกันแล้ว ทำให้การใช้งานมีความสะดวกและง่ายมากกว่าการสมมติให้พารามิเตอร์ในส่วนต่างๆของโมเดลมีความสัมพันธ์กัน

4) ในการศึกษาผู้วิจัยเลือกใช้การแจกแจงความน่าจะเป็นก่อนหน้าแบบไม่ให้สารสนเทศ (non-informative prior distribution) ซึ่งเป็นการแจกแจงความน่าจะเป็นก่อนหน้าที่ไม่มีอิทธิพลต่อการแจกแจงความน่าจะเป็นภายหลัง ดังนั้นสารสนเทศที่ใช้ในการประมาณค่าพารามิเตอร์ในโมเดลจะมาจากข้อมูลเชิงประจักษ์ทั้งหมด ทั้งนี้เพื่อเป็นการควบคุมปัจจัยแทรกซ้อนในงานวิจัยทำให้ผู้วิจัยสามารถเปรียบเทียบความสามารถในการประมาณค่าพารามิเตอร์ของวิธีการประมาณค่าแบบเบส์กับโปรแกรม Mplus ได้ แต่ในทางปฏิบัติผู้วิเคราะห์สามารถเพิ่มสารสนเทศให้แก่การแจกแจงความน่าจะเป็นก่อนหน้าได้ตามที่ต้องการ โดยปรับเปลี่ยนค่าของพารามิเตอร์ในการแจกแจงความน่าจะเป็นก่อนหน้าให้เป็นไปตามลักษณะที่ต้องการ ซึ่งผลการวิเคราะห์ที่ได้อาจมีความแตกต่างจากเดิมเพราะได้รับอิทธิพลจากการแจกแจงความน่าจะเป็นก่อนหน้ามากขึ้น

5) ในการศึกษานี้ผู้วิจัยกำหนดให้การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ตัวผกผัน (inverse) ของความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่มและความแปรปรวนของโมเดลในระดับที่หนึ่ง เป็นการแจกแจงความน่าจะเป็นแบบแกมมา โดยกำหนดให้มีพารามิเตอร์รูปร่าง (shape parameter: α) เท่ากับ 0.001 และพารามิเตอร์สเกล (scale parameter: β) เท่ากับ 0.001 การกำหนดในข้างต้นเป็นการกำหนดให้การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ความแปรปรวนดังกล่าวเป็นการแจกแจงแบบไม่ให้สารสนเทศ (non-informative prior) Gelman (2006) แนะนำว่าการกำหนดการแจกแจงความน่าจะเป็นดังกล่าวอาจส่งอิทธิพลต่อผลการวิเคราะห์ทำให้คลาดเคลื่อนจากความเป็นจริง โดยเฉพาะในกรณีที่ค่าของพารามิเตอร์ความแปรปรวนที่ต้องการประมาณนั้นมีค่าที่ใกล้ 0 กรณีดังกล่าว Gelman ว่าการแจกแจงความน่าจะเป็นแบบสม่ำเสมอ (uniform prior distribution) จะมีความเหมาะสมมากกว่า

6) เกณฑ์การพิจารณาที่ผู้วิจัยเลือกใช้ในการศึกษาส่วนของข้อมูลจำลองคือเกณฑ์ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (mean squares error: MSE) ซึ่งสามารถพิสูจน์ในทางคณิตศาสตร์ได้ว่า $MSE = bias^2 + var(\hat{\theta})$ เมื่อ $\hat{\theta}$ คือค่าประมาณพารามิเตอร์ หมายความว่า

เกณฑ์การพิจารณาดังกล่าวใช้ในการพิจารณาความแม่นยำและประสิทธิภาพของค่าประมาณพารามิเตอร์ในภาพรวม การใช้เกณฑ์ในข้างต้นแต่เพียงอย่างเดียวทำให้เกิดคำถามถึงความสามารถในการอนุมานด้านอื่นๆเพราะเกณฑ์ดังกล่าวไม่สามารถตอบคำถามได้โดยตรง

7) โปรแกรม R เป็นโปรแกรมที่มีประสิทธิภาพในการคำนวณที่สูงแต่อย่างไรก็ตาม โปรแกรมดังกล่าวเป็นโปรแกรมที่มีข้อด้อยข้อหนึ่งหรือหากกระบวนการในการคำนวณมีการวนรอบจำนวนมาก จะทำให้การประมวลผลต้องใช้เวลามากกว่าปกติ อัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ที่ผู้วิจัยพัฒนาขึ้นในงานวิจัยนี้จะเห็นว่ามีการวนรอบจำนวนมากในหลายขั้นตอน ทำให้การประมวลผลจำเป็นต้องใช้เวลาจำนวนมากจึงเป็นสาเหตุทำให้ การศึกษาด้วยข้อมูลจำลองมีสถานการณ์จำลองที่ไม่หลากหลายและครอบคลุมเท่าที่ควร ปัจจัยหนึ่งที่ในงานวิจัยนี้ไม่ได้ศึกษาแต่ผู้วิจัยคาดว่าจะมีความสำคัญและส่งผลกระทบต่อความสามารถของวิธีการประมาณคือปัจจัยระดับของความผันแปรระหว่างกลุ่ม โดยผู้วิจัยคาดว่าในกรณีที่ความผันแปรระหว่างกลุ่มมีน้อยวิธีการประมาณค่าแบบเบส์และวิธีการประมาณค่าจากโปรแกรม Mplus น่าจะให้ผลการประมาณค่าที่ใกล้เคียงกัน แต่เมื่อระดับความผันแปรระหว่างกลุ่มมีมากขึ้นวิธีการประมาณค่าแบบเบส์จะมีความสามารถที่สูงกว่าวิธีการประมาณค่าจากโปรแกรม Mplus

8) การศึกษาด้วยวิธีการจำลองแบบมอนติคาร์โล เนื่องจากข้อจำกัดของโปรแกรม R ดังที่กล่าวไว้ในข้างต้น ผู้วิจัยจึงกำหนดให้มีจำนวนรอบของการกระทำซ้ำ (replicate) จำนวน 100 รอบต่อ 1 สถานการณ์ ซึ่งถือว่าเป็นจำนวนที่อาจน้อยเกินไป ทำให้ค่าเกณฑ์การพิจารณา MSE ที่ใช้อาจยังมีความไม่นิ่งพอ อย่างไรก็ตามเพื่อพิจารณาผลการวิเคราะห์ที่ได้พบว่ามีแนวโน้มของค่า MSE เป็นไปตามที่คาดไว้

9) งานวิจัยนี้ยังขาดการพัฒนาเครื่องมือสำหรับการทดสอบสมมติฐานหรือเปรียบเทียบโมเดลซึ่งเป็นส่วนที่มีความสำคัญในเชิงปฏิบัติ สถิติตัวหนึ่งที่สำคัญและควรพัฒนาต่อคือสถิติ Bayes factor ซึ่งสามารถนำมาตอบคำถามในเชิงปฏิบัติที่ผู้วิจัยอาจต้องการทราบว่ามีความผันแปรระหว่างกลุ่มในโมเดลการวัดหรือโมเดลสมการโครงสร้างพหุระดับจริงหรือไม่ หรือพารามิเตอร์บางตัวควรกำหนดให้เป็นพารามิเตอร์แบบคงที่แทนที่จะเป็นแบบสุ่มหรือไม่ นอกจากนี้สถิติ Bayes factor ยังมีความสามารถในการเปรียบเทียบโมเดลแบบไม่ติดกลุ่ม (non-nested model) อีกด้วย

<u>อภิปรายในส่วนขององค์ความรู้ที่ได้รับ</u>

1) การวิจัยในส่วนของการศึกษาด้วยวิธีการจำลองจะพบว่ามีขอบเขตการวิจัยที่ค่อนข้างจำกัด ทั้งในส่วนของวิธีการประมาณค่าพารามิเตอร์ที่ได้พัฒนาขึ้น และในส่วนขององค์ความรู้ที่ได้จากการศึกษาเปรียบเทียบด้วยวิธีการจำลอง ส่งผลให้ขอบเขตของการนำผลการวิจัยที่ได้ไปใช้งานยังไม่มีความทั่วไปพอ อย่างไรก็ตามวิธีการประมาณค่าที่พัฒนาขึ้นยังสามารถพัฒนาต่อยอดให้มีความทั่วไปมากยิ่งขึ้นได้ โดยได้เสนอแนวคิดเอาไว้ในข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

2) จากผลการศึกษาด้วยวิธีการจำลองจะพบว่าวิธีการประมาณค่าแบบเบส์เป็นวิธีที่ดีกว่าในทุกกรณี สาเหตุเนื่องจากการประมาณค่าด้วยวิธีการประมาณแบบเบส์เป็นการประมาณค่าภายใต้โมเดลสมการโครงสร้างพหุระดับที่มีโมเดลการวัดเป็นแบบพารามิเตอร์สุ่ม ในขณะที่การประมาณค่าด้วยวิธีการประมาณแบบภาวะความควรจะเป็นสูงสุดแบบ restricted เป็นการประมาณค่าภายใต้โมเดลสมการโครงสร้างพหุระดับที่มีโมเดลการวัดเป็นแบบพารามิเตอร์คงที่ จึงเป็นธรรมชาติวิธีการประมาณแบบเบส์จึงย่อมให้ผลการประมาณค่าพารามิเตอร์ที่มีความถูกต้องมากกว่าเนื่องจากค่าพารามิเตอร์ที่ได้มากจากโมเดลที่มีความละเอียดและซับซ้อนมากกว่า สาเหตุที่วิธีการประมาณค่าพารามิเตอร์ทั้งสองจำเป็นต้องประมาณค่าภายใต้โมเดลที่แตกต่างกัน เนื่องด้วยข้อจำกัดของการประมาณค่าพารามิเตอร์แบบภาวะความควรจะเป็นสูงสุดแบบ restricted ทำให้ไม่สามารถประมาณค่าพารามิเตอร์ภายใต้โมเดลเดียวกับวิธีการประมาณค่าแบบเบส์ได้

3) เมื่อพิจารณาผลการวิจัยที่ได้จากการจำลองจะพบว่า เมื่อจำนวนตัวอย่างมีขนาดเล็กวิธีการประมาณค่าพารามิเตอร์แบบเบส์จะดีกว่าวิธีการประมาณค่าพารามิเตอร์แบบภาวะความควรจะเป็นสูงสุดแบบ restricted อย่างเห็นได้ชัด แต่เมื่อจำนวนตัวอย่างเพิ่มมากขึ้นวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted มีแนวโน้มที่จะได้ค่าประมาณที่ใกล้เคียงกับวิธีการประมาณค่าแบบเบส์มากยิ่งขึ้น ยกเว้นในกรณีที่ระดับของค่าเฉลี่ยความเที่ยงรวมมีค่าเท่ากับ 0.3 ผลการวิจัยที่ได้นี้สอดคล้องกับคุณสมบัติของวิธีการทั้งสองกล่าวคือ วิธีการประมาณค่าพารามิเตอร์แบบภาวะความควรจะเป็นสูงสุดแบบ restricted เป็นวิธีการประมาณค่าพารามิเตอร์แบบดั้งเดิมที่ขึ้นกับทฤษฎีการลู่เข้าเชิงกำกับ (asymptotic theory) กล่าวคือ

คุณสมบัติที่ดีของตัวประมาณดังกล่าวจะเกิดขึ้นเมื่อจำนวนตัวอย่างมีจำนวนมากเพียงพอ ในขณะที่วิธีการประมาณค่าแบบเบส์ไม่ได้ขึ้นกับทฤษฎีลู่เข้าเชิงกำกับดังกล่าว วิธีการประมาณค่าแบบเบส์จึงเป็นวิธีการที่ควรให้ผลการประมาณที่ดีกว่าวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted เมื่อขนาดตัวอย่างเล็ก และประสิทธิภาพของวิธีการทั้งสองควรลู่เข้าหากันเมื่อจำนวนตัวอย่างมีขนาดใหญ่เพิ่มขึ้น อย่างไรก็ตาม ณ ระดับค่าเฉลี่ยความเที่ยงรวมเท่ากับ 0.3 พบว่าพฤติกรรมของตัวประมาณทั้งสองไม่ได้เป็นไปตามทฤษฎีดังกล่าว ดังจะเห็นว่ายิ่งเพิ่มจำนวนตัวอย่างให้มากขึ้น เปอร์เซ็นต์ความแตกต่างของค่า MSE (พิจารณาจาก RDMSE) จะยิ่งมีค่าที่เพิ่มขึ้น สาเหตุเนื่องจาก ณ ระดับความเที่ยงดังกล่าวเป็นระดับความเที่ยงที่ต่ำมาก การเพิ่มจำนวนตัวอย่างจึงไม่น่าที่จะช่วยให้ประสิทธิภาพของวิธีการประมาณค่าแบบดั้งเดิมดีขึ้นได้อย่างชัดเจน

4) เมื่อพิจารณาผลการวิจัยในส่วนของการวิเคราะห์ด้วยข้อมูลทุติยภูมิจะพบว่าผลการวิเคราะห์ที่ได้จากวิธีการประมาณค่าแบบเบส์และวิธีการประมาณค่าแบบภาวะความควรจะเป็นสูงสุดแบบ restricted มีความแตกต่างกันอย่างเด่นชัดเพียงส่วนของอิทธิพลคงที่ที่ส่งจากตัวแปรผลผลสัมฤทธิ์ทางการเรียน (ACH_W) ที่ส่งผลต่อสุขภาวะของนักเรียน (SWB_w) โดยวิธีการประมาณค่าแบบเบส์ให้ค่าประมาณเท่ากับ 0.113 และมีนัยสำคัญเชิงสถิติที่ระดับ 0.01 ในขณะที่วิธีการประมาณค่าพารามิเตอร์แบบภาวะความควรจะเป็นสูงสุดแบบ restricted ให้ค่าประมาณเท่ากับ -0.025 และไม่มีนัยสำคัญเชิงสถิติ ในขณะที่พารามิเตอร์ในส่วนอื่นๆมีความแตกต่างกันเล็กน้อย ซึ่งสาเหตุอาจเนื่องมาจากการที่ข้อมูลชุดนี้มีความผันแปรระหว่างกลุ่มในระดับที่ไม่มากทำให้ผลการวิเคราะห์ระหว่างโมเดลการวัดที่มีสัมประสิทธิ์ความถดถอยแบบคงที่กับสัมประสิทธิ์ความถดถอยแบบสุ่มมีผลการวิเคราะห์ที่ไม่แตกต่างกันมาก

ข้อเสนอแนะในการนำไปใช้

1) อัลกอริทึมการประมาณค่าพารามิเตอร์ที่พัฒนาขึ้นในการวิจัยนี้พัฒนาขึ้นเพื่อใช้ในการประมาณค่าพารามิเตอร์ในโมเดลสมการโครงสร้างพหุระดับที่โมเดลการวัดมีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์แบบสุ่ม อย่างไรก็ตามอัลกอริทึมดังกล่าวสามารถนำไปใช้เพื่อวิเคราะห์โมเดลย่อยได้ เช่น โมเดลเชิง

เส้นพหุระดับ โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน หรือโมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ ได้โดยการลดขั้นตอนของอัลกอริทึมหรือกำหนดหรือปล่อยค่าพารามิเตอร์ให้เหมาะสมกับโมเดลที่ต้องการวิเคราะห์ ผู้วิจัยได้นำเสนอตัวอย่างการนำอัลกอริทึมที่พัฒนาขึ้นไปใช้ในโมเดลย่อยต่างๆดังนี้

ตัวอย่างการปรับปรุงอัลกอริทึมที่พัฒนาขึ้นเพื่อใช้สำหรับประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ

ในการวิเคราะห์องค์ประกอบเชิงสำรวจโดยใช้สถิติแบบเบส์ สามารถกระทำได้โดยลดรูปโมเดลเหลือเพียงโมเดลการวัดดังสมการ

$$\underline{y}_{ij} = \underline{\mu} + \Lambda \cdot \underline{\xi}_{ij} + \underline{\varepsilon}_{ij}$$

โดยที่ $\underline{y}_{ij}$ คือเวกเตอร์ของตัวแปรสังเกตได้ขนาด px1, $\underline{\mu}$ คือเวกเตอร์ของพารามิเตอร์จุดตัดแกนขนาด px1, Λ คือเมทริกซ์ของน้ำหนักองค์ประกอบขนาด pxq, $\underline{\xi}_{ij}$ คือเวกเตอร์ของตัวแปรแฝงขนาด qx1 และ $\underline{\varepsilon}_{ij}$ คือเวกเตอร์ของความคลาดเคลื่อนจากการวัดขนาด px1 ข้อสมมติของโมเดลมีสามประการได้แก่ 1) $\underline{\xi}_{ij} \sim N(0, I)$ และ 2) $\underline{\varepsilon}_{ij} \sim N(0, \Psi_{\varepsilon})$ โดยที่ Ψ_{ε} เป็นเมทริกซ์แนวทแยงมุม และ 3) $\underline{\xi}_{ij}$ และ $\underline{\varepsilon}_{ij}$ เป็นอิสระซึ่งกันและกัน

โมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจจะมีความแตกต่างจากโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันกล่าวคือการวิเคราะห์องค์ประกอบเชิงสำรวจจะไม่มีการกำหนดเงื่อนไขให้กับค่าพารามิเตอร์เมทริกซ์น้ำหนักองค์ประกอบ ดังนั้นในการประมาณค่าพารามิเตอร์ในโมเดลข้างต้นโดยอาศัยวิธีการประมาณค่าแบบเบส์สามารถกระทำได้จากอัลกอริทึมที่พัฒนาขึ้นโดยทำการปรับปรุงให้เหลือเฉพาะส่วนของโมเดลการวัดจำนวน 5 ขั้นตอนดังนี้

1. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\mu}$ จาก

$$\underline{\mu} \;\; | \left\{\underline{y}_{ij}\right\}, \kappa, \Sigma_{\mu}, \{\Lambda\}, \left\{\underline{\xi}_{ij}\right\}, \{\Psi_{\epsilon}\} \sim Normal\left(\hat{\kappa}, \hat{\Sigma}_{\mu}\right)$$

โดยที่ $\hat{\Sigma}_{\mu} = \left[\Sigma_{\mu}^{-1} + \sum_{j=1}^{J} n_j \Psi_{\epsilon}^{-1}\right]^{-1}$

และ $\hat{\kappa} = \hat{\Sigma}_{\mu}\left[\Sigma_{\mu}^{-1}\kappa + \sum_{j=1}^{J} \Psi_{\epsilon}^{-1} \sum_{i=1}^{n_j}\left(\underline{y}_{ij} - \Lambda \;\; \underline{\xi}_{ij}\right)\right]$

2. ในกรณีการสุ่มตัวอย่างพารามิเตอร์น้ำหนักองค์ประกอบ เนื่องจากการวิเคราะห์องค์ประกอบเชิงสำรวจไม่ได้มีการกำหนดเงื่อนไขข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบ ดังนั้นเมื่อเปรียบเทียบกับอัลกอริทึมที่พัฒนาขึ้นในงานวิจัยนี้จะสามารถลดขั้นตอนของการสร้างตัวแปรสังเกตที่ปรับได้เนื่องจากการกำหนดข้อจำกัดในเมทริกซ์น้ำหนักองค์ประกอบได้ดังนี้ สำหรับหน่วยที่ j และตัวแปรสังเกตได้ที่ k สุ่มตัวอย่างพารามิเตอร์ $\underline{\Lambda}_k$ จาก

$$\underline{\Lambda}_k | \left\{\underline{y}_{ijk}\right\}, \left\{\underline{\xi}_{ijk}\right\}, \psi_{\epsilon k} \sim Normal(\underline{\hat{g}}_k, \widehat{H}_k)$$

เมื่อ $\underline{\xi}_{ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ , $\underline{y}_{ijk}$ เป็นค่าตัวแปรสังเกตได้

โดยที่ $\widehat{H}_k = \left[H_k^{-1} + \psi_{\epsilon k}^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ijk} \underline{\xi}_{ijk}^T\right]^{-1}$

และ $\underline{\hat{g}}_k = \widehat{H}_k \left[H_k^{-1} \underline{g}_k + \psi_{\epsilon k}^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ijk} \underline{y}_{ijk}\right]$

3. สำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j สุ่มตัวอย่างตัวแปรแฝง $\underline{\omega}_{ij}$ จาก

$$\underline{\xi}_{ij} | \underline{\theta}\ , \Sigma_{\omega j}, \underline{\mu}\ , \Lambda\ , \Psi_\epsilon, \underline{y}_{ij} \sim Normal\left(\underline{\hat{\theta}}_{ij}, \hat{\Sigma}_\xi\right)$$

โดยที่ $\hat{\Sigma}_\xi = \left[\Phi^{-1} + \Lambda^T\ \Psi_\epsilon^{-1} \Lambda\ \right]^{-1}$

และ $\underline{\hat{\theta}}_{ij} = \hat{\Sigma}_\omega \Lambda^T\ \Psi_\epsilon^{-1} \left(\underline{y}_{ij} - \underline{\mu}\ \right)$

4. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Φ^{-1} จาก

$$\Phi^{-1} | \left\{\underline{\xi}_{ij}\right\}, \underline{\nu}_j, \rho_\phi, R_\phi \sim Wishart\left(\hat{\rho}_\phi, \hat{R}_\phi\right)$$

โดยที่ $\hat{\rho}_\phi = n + \rho_\phi$

และ $\hat{R}_\phi = \left[R_\phi^{-1} + \sum_{j=1}^{J} \sum_{i=1}^{n_j} \underline{\xi}_{ij} \underline{\xi}_{ij}^T\right]^{-1}$

5. สำหรับตัวแปรสังเกตได้ที่ k และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\psi_{\epsilon k}^{-1}$ จาก

$$\psi_{\epsilon k}^{-1} | \mu_k, \underline{\Lambda}_k, \left\{\underline{\xi}_{ij}\right\}, \left\{\underline{y}_{ij}\right\}, \alpha_k, \beta_k \sim Gamma\left(\hat{\alpha}_{\epsilon k}, \hat{\beta}_{\epsilon k}\right)$$

โดยที่ $\hat{\alpha}_{\epsilon k} = \frac{n_j}{2} + \alpha_k$

และ $\hat{\beta}_{\epsilon k} = \beta_k + \frac{1}{2} \sum_{i=1}^{n_j} \left(y_{ij} - \mu_k - \underline{\Lambda}_k^T \underline{\xi}_{ij}\right)^T \left(y_{ij} - \mu_k - \underline{\Lambda}_k^T \underline{\xi}_{ij}\right)$

ตัวอย่างการปรับปรุงอัลกอริทึมที่พัฒนาขึ้นเพื่อใช้สำหรับประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจแบบพหุระดับที่มีพารามิเตอร์น้ำหนักองค์ประกอบแบบสุ่มและความแปรปรวนของความคลาดเคลื่อนจากการวัดแบบสุ่ม

ในการวิเคราะห์องค์ประกอบเชิงสำรวจแบบพหุระดับโดยใช้สถิติแบบเบส์ สามารถกระทำได้โดยลดรูปโมเดลเหลือเพียงโมเดลการวัดดังสมการ

$$\underline{y}_{ij} = \underline{\mu} + \Lambda_j \cdot \underline{\xi}_{ij} + \underline{\varepsilon}_{ij}$$

โดยที่ $\underline{y}_{ij}$ คือเวกเตอร์ของตัวแปรสังเกตได้ขนาด px1, $\underline{\mu}$ คือเวกเตอร์ของพารามิเตอร์จุดตัดแกนขนาด px1, Λ_j คือเมทริกซ์ของน้ำหนักองค์ประกอบขนาด pxq, $\underline{\xi}_{ij}$ คือเวกเตอร์ของตัวแปรแฝงขนาด qx1 และ $\underline{\varepsilon}_{ij}$ คือเวกเตอร์ของความคลาดเคลื่อนจากการวัดขนาด px1 ข้อสมมติของโมเดลมีสามประการได้แก่ 1) $\underline{\xi}_{ij} \sim N(0, I)$ และ 2) $\underline{\varepsilon}_{ij} \sim N(\nu_j, \Psi_{\varepsilon j})$ โดยที่ $\Psi_{\varepsilon j}$ เป็นเมทริกซ์แนวทแยงมุม และ 3) $\underline{\xi}_{ij}$ และ $\underline{\varepsilon}_{ij}$ เป็นอิสระซึ่งกันและกัน

การประมาณค่าพารามิเตอร์ในกรณีนี้คล้ายคลึงกับการประมาณค่าการวิเคราะห์องค์ประกอบเชิงสำรวจแบบดั้งเดิมแต่ต้องมีการคำนึงถึงความเป็นพหุระดับของข้อมูลเพิ่มเติมดังนี้

1. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\mu}$ จาก

   $$\underline{\mu} \mid \{\underline{y}_{ij}\}, \kappa, \Sigma_\mu, \{\Lambda_j\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\} \sim Normal(\hat{\kappa}, \hat{\Sigma}_\mu)$$

   โดยที่ $\hat{\Sigma}_\mu = \left[\Sigma_\mu^{-1} + \sum_{j=1}^{J} n_j \Psi_{\epsilon j}^{-1}\right]^{-1}$

   และ $\hat{\kappa} = \hat{\Sigma}_\mu \left[\Sigma_\mu^{-1}\kappa + \sum_{j=1}^{J} \Psi_{\epsilon j}^{-1} \sum_{i=1}^{n_j} \left(\underline{y}_{ij} - \Lambda_j \underline{\omega}_{ij}\right)\right]$

2. ในกรณีการสุ่มตัวอย่างพารามิเตอร์น้ำหนักองค์ประกอบ เนื่องจากการวิเคราะห์องค์ประกอบเชิงสำรวจไม่ได้มีการกำหนดเงื่อนไขข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบ ดังนั้นเมื่อเปรียบเทียบกับอัลกอริทึมที่พัฒนาขึ้นในงานวิจัยนี้จะสามารถลดขั้นตอนของการสร้างตัวแปรสังเกตที่ปรับได้เนื่องจากการกำหนดข้อจำกัดในเมทริกซ์น้ำหนักองค์ประกอบได้ดังนี้ สำหรับหน่วยที่ j และตัวแปรสังเกตได้ที่ k สุ่มตัวอย่างพารามิเตอร์ $\underline{\Lambda}_{kj}$ จาก

   $$\underline{\Lambda}_{kj} \mid \{\underline{y}_{ijk}\}, \{\underline{\xi}_{ijk}\}, \psi_{\epsilon kj} \sim Normal(\underline{\hat{g}}_{kj}, \hat{H}_{kj})$$

เมื่อ $\underline{\xi}_{ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ_j, $\underline{y}_{ijk}$ เป็นค่าตัวแปรสังเกตได้ที่ปรับค่าที่เนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบในแถวที่ k พารามิเตอร์เมทริกซ์ Λ_j

โดยที่ $\hat{H}_{kj} = \left[H_{kj}^{-1} + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ijk}\underline{\xi}_{ijk}^T\right]^{-1}$

และ $\underline{\hat{g}}_{kj} = \hat{H}_{kj}\left[H_{kj}^{-1}\underline{g}_{kj} + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ijk}\underline{y}_{ijk}\right]$

3. สำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j สุ่มตัวอย่างตัวแปรแฝง $\underline{\xi}_{ij}$ จาก

$$\underline{\xi}_{ij}|\underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}\ \ , \Lambda_j, \Psi_{\epsilon j}, \underline{y}_{ij} \sim Normal(\underline{\hat{\theta}}_{ij}, \hat{\Sigma}_{\omega j})$$

โดยที่ $\hat{\Sigma}_{\xi j} = \left[\Phi^{-1} + \Lambda_j^T \Psi_{\epsilon j}^{-1} \Lambda_j\right]^{-1}$

และ $\underline{\hat{\theta}}_{ij} = \hat{\Sigma}_{\xi j}\Lambda_j^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}\ \ \right)$

4. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ค่าเฉลี่ยของตัวแปรแฝง $\underline{\nu}_j$ จาก

$$\underline{\nu}_j|\left\{\underline{\xi}_{ij}\right\}, \Phi, \Delta \sim Normal(\underline{\hat{\nu}}_j, \hat{\Delta})$$

โดยที่ $\hat{\Delta} = \left[\Delta^{-1} + n_j\Phi^{-1}\right]^{-1}$ และ $\underline{\hat{\nu}}_j = \hat{\Delta}\left[\Phi^{-1}\sum_{i=1}^{n_j}\underline{\xi}_{ij}\right]$

5. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Φ^{-1} จาก

$$\Phi^{-1}|\left\{\underline{\xi}_{ij}\right\}, \underline{\nu}_j, \rho_\phi, R_\phi \sim Wishart(\hat{\rho}_\phi, \hat{R}_\phi)$$

โดยที่ $\hat{\rho}_\phi = n + \rho_\phi$

และ $\hat{R}_\phi = \left[R_\phi^{-1} + \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\underline{\xi}_{ij} - \underline{\nu}_j\right)\left(\underline{\xi}_{ij} - \underline{\nu}_j\right)^T\right]^{-1}$

6. สำหรับตัวแปรสังเกตได้ที่ k และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\psi_{\epsilon kj}^{-1}$ จาก

$$\psi_{\epsilon kj}^{-1}|\mu_k, \underline{\Lambda}_k, \{\underline{\omega}_{ij}\}, \left\{\underline{y}_{ij}\right\}, \alpha_{kj}, \beta_{kj} \sim Gamma(\hat{\alpha}_{\epsilon kj}, \hat{\beta}_{\epsilon kj})$$

โดยที่ $\hat{\alpha}_{\epsilon kj} = \frac{n_j}{2} + \alpha_{kj}$

และ $\hat{\beta}_{\epsilon kj} = \beta_{kj} + \frac{1}{2}\sum_{i=1}^{n_j}\left(y_{ij} - \mu_k - \underline{\Lambda}_k^T\underline{\omega}_{ij}\right)^T\left(y_{ij} - \mu_k - \underline{\Lambda}_k^T\underline{\omega}_{ij}\right)$

<u>ตัวอย่างการปรับปรุงอัลกอริทึมที่พัฒนาขึ้นเพื่อใช้สำหรับประมาณค่าพารามิเตอร์ในโมเดลเชิงเส้นพหุระดับ</u>

ในกรณีนี้โมเดลการวิเคราะห์จะไม่มีโมเดลการวัด เนื่องจากโมเดลดังกล่าวไม่ได้มีการคำนึงถึงความคลาดเคลื่อนจากการวัดที่เกิดขึ้น โมเดลจึงลดรูปจึงเป็นดังต่อไปนี้

$$\eta_{ij} = \beta_j^T \underline{\Xi}_{ij} + \delta_{ij} = \underline{\gamma}^T \underline{\Xi}_{ij} + \underline{u}_j^T \underline{\Xi}_{ij} + \delta_{ij}$$

เมื่อ $\underline{\Xi}_{ij} = \begin{pmatrix} 1 & \underline{\xi}_{ij} \end{pmatrix}^T$ และ $\beta_j = \underline{\gamma} + \underline{u}_j$,

$\underline{\epsilon}_{ij} \sim N(\underline{0}, \Psi_{\epsilon j})$ โดยที่ $\Psi_{\epsilon j} = diag(\psi_{\epsilon kj})$,

และ $\underline{u}_j \sim N\left(\underline{0}, \Sigma_u\right)$

อัลกอริทึมสำหรับประมาณค่าพารามิเตอร์ในโมเดลนี้ สามารถปรับปรุงจากอัลกอริทึมหลักโดยการตัดส่วนที่เป็นการประมาณค่าในโมเดลการวัดออกไปทำให้ได้อัลกอริทึมที่จะใช้งานจริงเพียง 4 ขั้นตอนดังต่อไปนี้

1. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\gamma}$ จาก

$$\underline{\gamma} | \{\underline{\omega}_{ij}\}, \underline{u}_j, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{\gamma}}\ , \hat{\Sigma}_\gamma\right)$$

โดยที่ $\hat{\Sigma}_\gamma = \left[\Sigma_{\gamma 0}^{-1} + \frac{1}{\sigma_\delta^2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \underline{\Xi}_{ij} \underline{\Xi}_{ij}^T\right]^{-1}$

และ$\underline{\hat{\gamma}} = \hat{\Sigma}_\gamma \left[\Sigma_{\gamma 0}^{-1} \underline{\gamma} + \frac{1}{\sigma_\delta^2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \underline{\Xi}_{ij} (\eta_{ij} - \underline{\Xi}_{ij}^T \underline{u}_j)\right]$ เมื่อ $\underline{\Xi}_{ij} = \begin{pmatrix} 1 & \underline{\xi}_{ij} \end{pmatrix}^T$

2. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{u}_j$ จาก

$$\underline{u}_j | \{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{u}}_j, \hat{\Sigma}_{u\mathrm{j}}\right)$$

โดยที่ $\hat{\Sigma}_{u\mathrm{j}} = \left[\frac{1}{\sigma_\delta^2} \sum_{i=1}^{n_j} \underline{\Xi}_{ij} \underline{\Xi}_{ij}^T + \Sigma_u^{-1}\right]^{-1}$

และ $\underline{\hat{u}}_j = \frac{1}{\sigma_\delta^2} \hat{\Sigma}_{u\mathrm{j}} \left[\sum_{i=1}^{n_j} \underline{\Xi}_{ij}^T \left(\eta_{ij} - \underline{\Xi}_{ij}^T \underline{\gamma}\right)\right]$

3. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Σ_u^{-1} จาก

$$\Sigma_u^{-1} | \underline{u}_j, \rho_u, R_u \sim Wishart\left(\hat{\rho}_u, \hat{R}_u\right)$$

โดยที่ $\hat{\rho}_u = J + \rho_u$ และ $\hat{R}_u = \left[R_u^{-1} + \sum_{j=1}^{J} \underline{u}_j^T \underline{u}_j\right]^{-1}$

4. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $(\sigma_\delta^2)^{-1}$ จาก

$$(\sigma_\delta^2)^{-1} | \{\underline{\omega}_{ij}\}, \left\{\underline{\beta_j}\right\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta \sim Gamma\left(\hat{\alpha}_\delta, \hat{\beta}_\delta\right)$$

โดยที่ $\hat{\alpha}_\delta = \frac{\sum_{j=1}^{J} n_j}{2} + \alpha_\delta$ และ $\hat{\beta}_\delta = \beta_\delta + \frac{1}{2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(\eta_{ij} - \underline{\Xi}_{ij}^T \underline{\gamma} - \underline{\Xi}_{ij}^T \underline{u_j}\right)^2$

2) ในเชิงปฏิบัติการเลือกการแจกแจงความน่าจะเป็นก่อนหน้าที่ไม่เหมาะสมอาจส่งอิทธิพลให้ผลการวิเคราะห์มีความคลาดเคลื่อน ผู้วิเคราะห์อาจจำเป็นต้องพิจารณาความแกร่งของ

ผลการวิเคราะห์ที่มีต่อการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในหลากหลายรูปแบบเพื่อยืนยันผลการวิเคราะห์ที่ถูกต้อง เรียกกระบวนการนี้ว่า การวิเคราะห์ความไว (sensitivity analysis) (Bornn, Doucet, & Gottardo, 2010))

3) ในการใช้งานอัลกอริทึมดังกล่าวในทางปฏิบัติสำหรับนักวิจัยที่ไม่ใช่นักสถิติจะพบว่ามีความยุ่งยากมากเนื่องจากต้องทำความเข้าใจกับสูตรทางคณิตศาสตร์และต้องเขียนโปรแกรมเพื่อใช้ในการประมาณค่าพารามิเตอร์เอง เพื่อหลีกเลี่ยงความยุ่งยากดังกล่าวโปรแกรม Winbugs อาจเป็นทางเลือกหนึ่งในการใช้งานจริง เนื่องด้วยโปรแกรมดังกล่าวเป็นโปรแกรมที่ถูกพัฒนาขึ้นสำหรับการวิเคราะห์ข้อมูลด้วยสถิติแบบเบส์ด้วยเทคนิคลูกโซ่มาร์คอฟมอนติคาร์โล นอกจากนี้ยังสามารถใช้งานร่วมกันกับโปรแกรม R ผู้วิเคราะห์สามารถจัดการข้อมูลในโปรแกรม R และส่งข้อมูลดังกล่าวผ่านทาง package R2WinBugs เข้าสู่โปรแกรม Winbugs เพื่อทำการประมาณค่าพารามิเตอร์ในโมเดลที่ต้องการ นอกจากนี้ผู้วิเคราะห์ยังสามารถนำผลการวิเคราะห์ที่ได้จากโปรแกรม Winbugs กลับเข้าสู่โปรแกรม R เพื่อทำการวิเคราะห์ในส่วนที่ต้องการต่อไปได้อีกด้วย

4) การนำอัลกอริทึมที่ผู้วิจัยพัฒนาขึ้นไปใช้ในทางปฏิบัติ หากผู้วิจัยต้องการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ในโมเดล ผู้วิจัยอาจเลือกใช้ช่วงความน่าเชื่อถือ (credible interval) ซึ่งคำนวณจากค่าควอนไทล์เป็นเครื่องมือในการทดสอบสมมติฐานแบบง่ายได้ อย่างไรก็ตามวิธีการดังกล่าวไม่ใช่วิธีการที่ดีที่สุดในการทดสอบสมมติฐานแบบเบส์

5) ในกรณีที่ผู้วิจัยต้องการประมาณค่าพารามิเตอร์ในโมเดลการวัดที่พารามิเตอร์จุดตัดแกนเป็นพารามิเตอร์แบบสุ่ม สามารถกระทำได้โดยการนิยามโมเดลและปรับเปลี่ยนรายละเอียดของอัลกอริทึมที่ใช้ในการประมาณค่า รายละเอียดของวิธีการดังกล่าวผู้วิจัยได้นำเสนอไว้ในภาคผนวก ข

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1) ในงานวิจัยนี้เลือกใช้การแจกแจงความน่าจะเป็นก่อนหน้าสำหรับพารามิเตอร์น้ำหนักองค์ประกอบ และความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นการแจกแจงความน่าจะเป็นแบบปกติและการแจกแจงความน่าจะเป็นแบบแกมมา ตามลำดับ มีงานวิจัยได้

ทำการศึกษาว่าการเลือกใช้การแจกแจงความน่าจะเป็นดังกล่าวเป็นการแจกแจงความน่าจะเป็นก่อนหน้าอาจทำให้ลูกโซ่มาร์คอฟที่สร้างขึ้นจากอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์นั้นลู่เข้าสู่การแจกแจงความน่าจะเป็นภายหลังที่ต้องการได้ช้า (Ghosh และ Dunson, 2008; Gelman 2006) การวิจัยต่อไปจึงอาจพัฒนาอัลกอริทึมโดยการกำหนดการแจกแจงความน่าจะเป็นก่อนหน้าตัวอื่นๆ ที่มีคุณสมบัติที่ดีกว่าตัวปัจจุบัน เช่น t distribution, folded-t distribution หรือการแจกแจงความน่าจะเป็นในวงศ์ของเอกซ์โพเนนเชียลตัวอื่นๆ เป็นต้น

2) อัลกอริทึมที่พัฒนาขึ้นสามารถประมาณค่าในโมเดลซึ่งเป็นโมเดลย่อยโมเดลหนึ่งของโมเดลสมการโครงสร้างพหุระดับในกรณีทั่วไป ดังนั้นในการวิจัยครั้งต่อไปอาจพัฒนาอัลกอริทึมให้สามารถใช้ได้กับโมเดลในกรณีทั่วไปมากขึ้น เช่น ในกรณีที่มีตัวแปรตามหลายตัวหรือวิเคราะห์ในโมเดลที่มีตัวแปรปรับหรือตัวแปรส่งผ่านได้โดยไม่ยากนัก ซึ่งอาจกระทำได้โดยใช้แนวคิดของการประมาณค่าพารามิเตอร์แบบเบส์ในโมเดลสมการโครงสร้างตามแนวทางงานวิจัยของ Lee และ Song (2004, 2003 และ 2001) หรือ Dunson, Palomo และ Bollen (2005) และโมเดลเชิงเส้นพหุระดับตามแนวทางของ Browne (1998) ในข้อเสนอแนะส่วนนี้ผู้วิจัยได้นำเสนอตัวอย่างแนวคิดเบื้องต้นสำหรับการพัฒนาอัลกอริทึมเพื่อประมาณค่าในโมเดลที่มีความเป็นกรณีทั่วไปมากขึ้นในบางโมเดลดังต่อไปนี้

<u>แนวคิดในการขยายอัลกอริทึมในกรณีที่ต้องการวิเคราะห์ตัวแปรปรับ</u>

<u>โมเดลการวัด</u>

$$\underline{y}_{ij} = \underline{\mu} + \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij}$$

<u>โมเดลสมการโครงสร้าง</u>

$$\eta_{ij} = \underline{\Xi}_{ij}^T W_j \underline{\gamma} + \underline{\Xi}_{ij}^T \underline{u}_j + \delta_{ij}$$

เมื่อ W_j คือ design matrix ของตัวแปรอิสระในระดับที่ 2

จากโมเดลในข้างต้นการปรับเปลี่ยนอัลกอริทึมเดิมที่ได้พัฒนาไว้กระทำเพียง 3 ขั้นตอนดังนี้

1. ขั้นตอนการสุ่มตัวอย่างพารามิเตอร์ $\underline{\gamma}$ จาก

$$\underline{\gamma} | \{\underline{\omega}_{ij}\}, W_j, \underline{u}_j, \sigma_\delta^2, \Sigma_u \sim Normal\left(\underline{\hat{\gamma}}\ \ , \hat{\Sigma}_\gamma\right)$$

โดยที่ $\hat{\Sigma}_{\gamma} = \left[\Sigma_{\gamma 0}^{-1} + \frac{1}{\sigma_{\delta}^{2}} \sum_{j=1}^{J} \sum_{i=1}^{n_j} W_j^T \underline{\Xi}_{ij} \underline{\Xi}_{ij}^T W_j\right]^{-1}$

และ$\underline{\hat{\gamma}} \quad = \hat{\Sigma}_{\gamma} \left[\Sigma_{\gamma 0}^{-1} \underline{\gamma} + \frac{1}{\sigma_{\delta}^{2}} \sum_{j=1}^{J} W_j \sum_{i=1}^{n_j} \underline{\Xi}_{ij} (\eta_{ij} - \underline{\Xi}_{ij}^T \underline{u}_j)\right]$

2. ขั้นตอนการสุ่มตัวอย่างเศษเหลือในระดับที่สอง $\underline{u}_j$ จาก

$$\underline{u}_j | \{\underline{\omega}_{ij}\}, W_j, \underline{\gamma}, \sigma_{\delta}^{2}, \Sigma_u \sim Normal(\underline{\hat{u}}_j, \hat{\Sigma}_{u\mathrm{j}})$$

โดยที่ $\hat{\Sigma}_{u\mathrm{j}} = \left[\frac{1}{\sigma_{\delta}^{2}} \sum_{i=1}^{n_j} \underline{\Xi}_{ij} \underline{\Xi}_{ij}^T + \Sigma_u^{-1}\right]^{-1}$

และ $\underline{\hat{u}}_j = \frac{1}{\sigma_{\delta}^{2}} \hat{\Sigma}_{u\mathrm{j}} \left[\sum_{i=1}^{n_j} \underline{\Xi}_{ij}^T \left(\eta_{ij} - \underline{\Xi}_{ij}^T W_j \underline{\gamma}\right)\right]$

3. ขั้นตอนการสุ่มตัวอย่างพารามิเตอร์ $(\sigma_{\delta}^{2})^{-1}$ จาก

$$(\sigma_{\delta}^{2})^{-1} | \{\underline{\omega}_{ij}\}, \underline{\gamma}, \{\underline{u}_j\}, \alpha_{\delta}, \beta_{\delta} \sim Gamma(\hat{\alpha}_{\delta}, \hat{\beta}_{\delta})$$

โดยที่ $\hat{\alpha}_{\delta} = \frac{\sum_{j=1}^{J} n_j}{2} + \alpha_{\delta}$

และ $\hat{\beta}_{\delta} = \beta_{\delta} + \frac{1}{2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(\eta_{ij} - \underline{\Xi}_{ij}^T W_j \underline{\gamma} - \underline{\Xi}_{ij}^T \underline{u_j}\right)^2$

<u>แนวคิดในการขยายอัลกอริทึมในกรณีที่ต้องการวิเคราะห์ตัวแปรตามแฝงหลายตัวแปร</u>

<u>โมเดลการวัด</u>

$$\underline{y}_{ij} = \underline{\mu} + \Lambda_j \underline{\omega}_{ij} + \underline{\epsilon}_{ij}$$

<u>โมเดลสมการโครงสร้าง</u>

$$\underline{\eta}_{ij} = \Pi_j \underline{\eta}_{ij} + \Gamma_j \underline{\xi}_{ij} + \underline{\delta}_{ij}$$

โดยที่ Π_j และ Γ_j คือเมทริกซ์ของสัมประสิทธิ์ความถดถอยแบบสุ่มขนาด $q_1 \times q_1$ และ $q_1 \times q_2$ ตามลำดับ, $\underline{\epsilon}_{ij} \sim N(\underline{0}, \Psi_{\epsilon j})$ และ $\underline{\delta}_{ij} \sim N(\underline{0}, \Psi_{\delta})$ เมื่อ $\Psi_{\epsilon j} = diag(\psi_{\epsilon k})$ และ $\Psi_{\delta} = diag(\psi_{\delta h})$

จากโมเดลสมการโครงสร้างในข้างต้นจะสามารถพิสูจน์คุณสมบัติของโมเดลได้ดังต่อไปนี้

จากโมเดลสมการโครงสร้าง $\underline{\eta}_{ij} = \Pi_j \underline{\eta}_{ij} + \Gamma_j \underline{\xi}_{ij} + \underline{\delta}_{ij}$

$$\underline{\eta}_{ij} - \Pi_j \underline{\eta}_{ij} = \Gamma_j \underline{\xi}_{ij} + \underline{\delta}_{ij}$$

$$(I - \Pi_j) \underline{\eta}_{ij} = \Gamma_j \underline{\xi}_{ij} + \underline{\delta}_{ij}$$

$$\underline{\eta}_{ij} = (I - \Pi_j)^{-1}\left(\Gamma_j\underline{\xi}_{ij} + \underline{\delta}_{ij}\right)$$

จากการปรับรูปของโมเดลในข้างต้นทำให้ได้ว่า

$$Var\left(\underline{\eta}_{ij}\middle|\Pi_j, \Gamma_j\right) = (I - \Pi_j)^{-1}\left[\Gamma_j\Phi\Gamma_j^T + \Psi_\delta\right](I - \Pi_j)^{-T}$$

และ $Cov\left(\underline{\eta}_{ij}, \underline{\xi}_{ij}\right) = (I - \Pi_j)^{-1}\Gamma_j\underline{\xi}_{ij}$

ในการประมาณค่าพารามิเตอร์ในโมเดลในส่วนของโมเดลสมการโครงสร้างสามารถกระทำได้โดยใช้แนวคิดเดียวกับการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบดังต่อไปนี้

จาก $\underline{\eta}_{ij} = \Pi_j\underline{\eta}_{ij} + \Gamma_j\underline{\xi}_{ij} + \delta_{ij} = \Lambda_{\omega j}\underline{\omega}_{ij} + \underline{\delta}_{ij}$ โดยที่ $\Lambda_{\omega j} = \left(\underline{\eta}_{ij}, \underline{\xi}_{ij}\right)^T$

จะเห็นว่าโมเดลในสมการข้างต้นเป็นโมเดลที่ยังไม่สามารถระบุได้ ดังนั้นในการประมาณค่าพารามิเตอร์จำเป็นที่จะต้องถูกต้อง หลังจากทำการกำหนดเงื่อนไขให้กับพารามิเตอร์ได้อย่างเหมาะสมแล้ว การประมาณค่าพารามิเตอร์ในโมเดลสามารถกระทำได้โดยอาศัยแนวคิดเดียวกับการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบ

<u>แนวคิดในการขยายอัลกอริทึมในกรณีที่โมเดลการวัดมีจุดตัดแกนเป็นพารามิเตอร์สุ่ม</u>

ในงานวิจัยทางสังคมศาสตร์ จิตวิทยา หรือทางการศึกษา ในบางกรณีผู้วิจัยอาจสนใจให้พารามิเตอร์จุดตัดแกนมีความผันแปรระหว่างกลุ่ม และกำหนดโมเดลในระดับที่สองให้กับพารามิเตอร์จุดตัดแกน ทำให้สามารถรวมคะแนนพารามิเตอร์จุดตัดแกนขึ้นไปเป็นตัวแปรแฝงในระดับที่สองได้ รายละเอียดของโมเดลการวัดเป็นดังนี้

<u>โมเดลการวัดระดับที่หนึ่ง</u>

$$\underline{y}_{ij} = \underline{\mu}_j + \Lambda\underline{\omega}_{ij} + \underline{\epsilon}_{ij}$$

<u>โมเดลการวัดระดับที่สอง</u>

$$\underline{\mu}_j = \underline{\kappa} + \Lambda_b\underline{\varphi}_j + \underline{e}_j$$

ข้อสมมติเบื้องต้นของโมเดลกำหนดให้ $\underline{\epsilon}_{ij} \sim N(0, \Psi_{\epsilon j})$, $\underline{\omega}_{ij} \sim N(\underline{0}, \Sigma_{\omega j})$ โดยที่

$$\Sigma_{\omega j} = \begin{bmatrix} \underline{\beta}_j\Phi_j\underline{\beta}_j^T + \sigma_\delta^2 & \underline{\beta}_j^T\Phi_j \\ \underline{\beta}_j\Phi_j & \Phi_j \end{bmatrix}$$

$\underline{e}_j \sim N(\underline{0}, \Psi_b)$ โดยที่ $\Psi_{b\epsilon}$ เป็นเมทริกซ์ความแปรปรวนร่วมทแยงมุมของความคลาดเคลื่อนจากการวัดในระดับที่สอง และ $\underline{\varphi}_j \sim N(\underline{0}, \Phi_b)$ จากข้อสมมติดังกล่าวจะได้ว่า $\underline{\mu}_j \sim N(\underline{\alpha}, \Sigma_b)$ โดยที่ $\Sigma_b = \Lambda_b \Phi_b \Lambda_b^T + \Psi_b$

จากสมการที่ (1) จะสามารถเขียนสมการรวมได้ดังนี้

$$\underline{y}_{ij} = \underline{\kappa} + \Lambda_b \underline{\varphi}_j + \Lambda \underline{\omega}_{ij} + \underline{e}_j + \underline{\epsilon}_{ij}$$

จากข้อสมมติในข้างต้นจะได้ว่า

$$E\left(\underline{y}_{ij}\right) = \underline{\kappa}$$

$$Cov\left(\underline{y}_{ij} | \Phi_j, \Psi_{\epsilon j}\right) = \Lambda_b \Phi_b \Lambda_b^T + \Lambda \Sigma_{\omega j} \Lambda^T + \Psi_b + \Psi_{\epsilon j}$$

การประมาณค่าพารามิเตอร์ในโมเดลข้างต้นผู้วิจัยได้พัฒนาและแสดงไว้ในภาคผนวก ข แล้ว

3) ข้อสมมติข้อหนึ่งของวิธีการที่ผู้วิจัยพัฒนาขึ้นคือ ตัวแปรสังเกตได้ที่นำมาวิเคราะห์จำเป็นที่จะต้องเป็นตัวแปรประเภทต่อเนื่อง (continuous observed variables) ซึ่งโดยทั่วไปข้อมูลในงานวิจัยทางสังคมศาสตร์นั้นมักทำการวัดโดยอาศัยแบบวัดทำให้สเกลการวัดตัวแปรสังเกตได้แต่ละตัวนั้นเป็นแบบไม่ต่อเนื่อง (discrete observed variables) เช่น การใช้สเกลการวัดของ Likert ที่คำตอบของคำถามในแต่ละข้อนั้นอยู่ในรูปของสเกลอันดับ ไม่เห็นด้วยมากที่สุด (1), ไม่เห็นด้วย (2), ไม่มีความคิดเห็น (3), เห็นด้วย (4) หรือ เห็นด้วยมากที่สุด (5) การวิเคราะห์โดยใช้ข้อมูลลักษณะดังกล่าวจะทำให้ผลการวิเคราะห์มีความบิดเบือนจากความเป็นจริง (Lee และ Song, 2004; Lee, Poon และ Bentler, 1990 a และ 1990 b) จากปัญหาในข้างต้นในการวิจัยในอนาคตอาจขยายอัลกอริทึมที่ได้พัฒนาไว้แล้วในงานวิจัยนี้ให้รองรับการวิเคราะห์ในกรณีที่ข้อมูลของตัวแปรสังเกตได้เป็นตัวแปรไม่ต่อเนื่อง โดยอาจพัฒนาโดยใช้แนวทางการระบุค่า threshold (threshold specifications) หรือการใช้โมเดลดการตอบสนองข้อสอบ (item response model) เป็นต้น

4) ผู้วิจัยได้ใช้โปรแกรม R เป็นเครื่องมือในการประมาณค่าพารามิเตอร์และพบว่าเป็นโปรแกรมที่มีความสามารถสูงและมีคำสั่งที่เข้าใจง่าย แต่ใช้เวลาในการประมวลผลมากโดยเฉพาะกับคำสั่งที่มีการวนรอบจำนวนมากดังที่ใช้ในงานวิจัยนี้ สาเหตุเพราะโปรแกรม R เป็นโปรแกรมภาษาแบบ interpreter ซึ่งจะแปลผลคำสั่งเป็นรายบรรทัด ดังนั้นในการศึกษาด้วยวิธีการจำลองที่มีการวนรอบจำนวนมากจึงควรใช้โปรแกรมภาษาแบบ complier ซึ่งจะทำให้การ

ประมวลผลรวดเร็วขึ้นอย่างมาก ในปัจจุบันโปรแกรม R สามารถทำงานร่วมกับโปรแกรมภาษา complier เช่น ภาษาซี (C) หรือ ภาษาฟอแทรน (Fortran) ซึ่งผู้วิจัยสามารถจำลองข้อมูลจากในโปรแกรม R จากนั้นส่งออกไปประมวลผลกับโปรแกรมดังกล่าว แล้วจึงดึงข้อมูลการวิเคราะห์กลับเข้ามาเก็บไว้ในโปรแกรม R เพื่อวิเคราะห์ในขั้นตอนอื่นๆต่อไป ซึ่งจะช่วยประหยัดเวลาได้เป็นจำนวนมาก (Java, Gaile และ Manly, 2007; Peng และ Leeuw, 2002)

5) โปรแกรม Mplus version 7.0 ซึ่งเป็น version ที่จะทำการวางจำหน่ายล่าสุดนี้ มีความสามารถในการวิเคราะห์โดยใช้สถิติแบบเบส์ ซึ่งทำให้สามารถวิเคราะห์โมเดลเดียวกันกับงานวิจัยนี้ได้ ในการศึกษาครั้งต่อไปจึงควรศึกษาเปรียบเทียบความสามารถของวิธีการประมาณค่าแบบเบส์ที่ผู้วิจัยพัฒนาขึ้นและวิธีการประมาณค่าแบบเบส์จากโปรแกรม Mplus version 7.0 ในแต่ละสถานการณ์จำลอง

6) ในกรณีที่ผู้วิจัยต้องการเปรียบเทียบความเหมาะสมระหว่างโมเดลคู่แข่ง ตัวสถิติที่มีความสำคัญที่สามารถใช้ได้ในสถิติแบบเบส์คือตัวสถิติ Bayes factor ซึ่งได้เขียนรายละเอียดเอาไว้ในหน้า 62 ของงานวิจัย อย่างไรก็ตามการคำนวณค่า Bayes factor สำหรับโมเดลการวิเคราะห์ที่มีความซับซ้อนดังเช่นในงานวิจัยนี้สามารถกระทำได้ยากมาก ทำให้การใช้งานในเชิงปฏิบัติทำได้ลำบาก อย่างไรก็ตามโปรแกรม Mplus version 7.0 มีฟังก์ชันย่อยทำให้สามารถคำนวณค่า Bayes factor เพื่อเปรียบเทียบโมเดลคู่แข่งขัน (candidate model) ในงานวิจัยครั้งต่อไปจึงเป็นที่น่าสนใจที่จะทำการศึกษาเกี่ยวกับตัวสถิติ Bayes factor นี้ ทั้งในเชิงทฤษฎี และในเชิงปฏิบัติ โดยใช้โปรแกรม Mplus version 7.0 เป็นเครื่องมือสำหรับการศึกษา

รายการอ้างอิง

<u>ภาษาไทย</u>

กฤษณะ เนียมมณี. (2544). *ทฤษฎีความน่าจะเป็น* (พิมพ์ครั้งที่ 2). กรุงเทพมหานคร: ห้างหุ้นส่วนจำกัดพิทักษ์การพิมพ์.

เกียรติเทพ ตั้งสันติถาวร. (2548). การเปรียบเทียบการประมาณค่าพารามิเตอร์ด้วยวิธีเบส์กับวิธีวิเคราะห์ความถดถอยแบบสองขั้นน้อยสุด. *วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย* .

ธีระพร วีระถาวร. (2539). *ความน่าจะเป็นกับการประยุกต์* (พิมพ์ครั้งที่ 2). กรุงเทพมหานคร: บริษัทวิทยพัฒน์ จำกัด.

ธีระพร วีระถาวร. (2541). *ตัวแบบเชิงเส้นและการประยุกต์* (พิมพ์ครั้งที่ 1). กรุงเทพมหานคร: บริษัทวิทยพัฒน์ จำกัด.

นงลักษณ์ วิรัชชัย และ สมหวัง พิธิยานุวัฒน์. (2543). ธรรมชาติของศาสตร์ทางการศึกษาและวิธีวิทยาการวิจัยการศึกษา. *วิธีวิทยาการวิจัย 13(2)* , 34-72.

ปรางค์ทิพย์ รัชตะปิติ. (2550). การเปรียบเทียบการประมาณค่าพารามิเตอร์ด้วยวิธีเบส์กับวิธีวิเคราะห์ความถดถอยพหุคูณแบบกำลังสองน้อยสุดสองขั้น. *วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย* .

พงศ์วัชร ฟองกันทา. (2551). ปัจจัยด้านนักเรียนและครูที่ส่งผลต่อความต้องการจำเป็นด้านคุณภาพนักเรียนการวิเคราะห์พหุระดับ. *วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย* .

มณฑิรา ดวงสาพล. (2550). การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุดร่วมกับวิธีดั้งเดิมเมื่อเปิดความคลาดเคลื่อนจากการวัดค่าตัวแปรอิสระ. *วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย* .

มานพ วราภักดิ์. (2550). *การจำลอง* (พิมพ์ครั้งที่ 1). กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

วีรพา ฐานะปรัชญ์. (2542). การวิเคราะห์เชิงเบส์สำหรับตัวแบบการถดถอยเชิงเส้นเชิงเดียว. *วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย* .

วุฒิพงษ์ เดโชดมพันธ์. (2546). การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณเมื่อเกิดความคลาดเคลื่อนในตัวแปรอิสระ. *วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย* .

ศิริชัย กาญจนวาสี. (2550). *การวิเคราะห์พหุระดับ* (พิมพ์ครั้งที่ 4). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2550). *ทฤษฎีการวัดแบบดั้งเดิม* (พิมพ์ครั้งที่ 4). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สิวะโชติ ศรีสุทธิยากร. (2550). การจำลองตัวแบบความถดถอยเชิงลำดับชั้น เมื่อการแจกแจงความคลาดเคลื่อนสุ่มไม่ได้มีการแจกแจงแบบปกติ. *วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย* .

<u>ภาษาต่างประเทศ</u>

Albert, J. (2009). *Bayesian Computation with R* (2 ed.). London: Springer+Business Media, LLC.

Ansari, A., & Jedidi, K. (2000). Bayesian factor analysis for multilevel binary observations. *Psychometrika , 65*, 475-498.

Ansari, A., Jedidi, K., & Dube, L. (2002). Heterogeneous factor analysis models: A Bayesian approach. *Psychometrika , 67*, 49-78.

Asparouhov, T., & Muthen, B. (2010, August). Bayesian Analysis of Latent Variable Models using Mplus.

Aspatouhov, T., & Muthen, B. (2011, May 27). *New Modeling Methods for Multilevel Data.* Retrieved October 5, 2011, from http://www.statmodel.com/download/2011%20Tihomir%20talk.pdf

Asparouhov, T., & Muthen, B. (2012, May 24). *General random effect latent variable modeling: Random subjects, items, contexts, and oarameters.* Retrieved July 13, 2012, from http://csm.lshtm.ac.uk/files/2011/02/Tihomir-Asparouhov-24-05-2012.pdf

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis.* New York: Springer-Verlag.

Boomsma, A. (1982). *The robustness of LISREL against small sample sizes in factor analysis model. In K.G. Jorkog and H. Wold (eds), System under Indirect Observation: Causality, Structural, Prediction.* Amsterdam: North-Holland.

Box, G. E., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis.* Reading MA: Addison-Wesley.

Browne, W. (1998). *Applying MCMC methods to multi-level models*. Ph.D. dissertation Department of Mathematical Sciences, University of Bath .

Browne, W., & Draper, D. (2000). Implementation and performance issue in the bayesian fitting of multilevel model. *Computational Statistics , 15*, 391-420.

Browne, W., Goldstein, H., Woodhouse, G., & Yang, M. (2001). An MCMC algorithm for adjusting for errors in variables in random slopes multilevel models. *Multilevel modeling newsletter , 13*, 4-9.

Chaimongkol, S., Huffer, F., & Kamata, A. (2006). A Bayesian approach for fitting a random effect differential item functioning accross group units. *Thailand Statistician , 4*, 27-41.

Cowles, M. (2004). Review of WinBUGS 1.4. *The American Statiscian , 58(4)*, 330-336.

Demspter, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society , 39*, 1-38.

Dunson, D. B., Palomo, J., & Bollen, K. (2005, July 27). *Bayesian structural equation modeling*. Retrieved Apr 13, 2009, from http://beta.samsi.info/sites/default/files/tr2005-05.pdf

E., D. (2004). *Robustness of hierarchical linear model parameter estimate under violations of second-level residual homoskedasticity and independence asumptions*. Ph.D. dissertation, Department of Education Psychology and Learning Systems, The Florida State Universoty, USA .

Ferrao, M. E., Leckie, G., & Goldstein, H. (2008). *Modelling measurement errors in random coefficient multilevel models: an application to measuring school effectiveness.* Centre for Multilevel Modelling, University of Bristol, UK .

Fox, J. -P. (2005). Multilevel IRT using dichotomous and polytomous response data. *British Journal of Mathematical and Statistical Psychology , 58*, 145-172.

Fox, J. -P., & Glas. (2001). Bayesian estimation of multilevel IRT model using Gibbs sampling. *Psychometrika , 66*, 259-286.

Fox, J. -P., & Glas. (2001). Bayesian modelling of measurement error in predictor variables using item response theory. *Psychometrika , 68*, 169-191.

Fuller, W. (1987). *Measurement error models.* New York: John Wiley & Sons, Inc.

Gamerman, D., & Lopes, H. F. (2006). *Markov Chain Monte Carlo.* London: Chapman & Hall/CRC.

Gelman, A. (2002). *Prior distribution.* Chiccheter: John Wiley & Sons, Ltd.

Gelman, A. (2006). Prior distribution for variance parameters in hierarchical models. *Bayesian Analysis , 1(3)*, 515-533.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis.* UK: Chapman & Hall.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian Statistics , 4*, 169-193.

Gill, J. (2008). *Bayesian methods: A social and behavioral sciences approach.* London: Chapman & Hall/CRC.

Goldstein, H. (1995). Hierarchical data modelling in social sciences. *Journal of Educational and Behavioral Statistics , 20(2)*, 201-204.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika , 73*, 43-56.

Goldstein, H. (1995). *Multilevel statistical models.* Newyork-Toronto: John Wiley & Sons, Inc.

Goldstein, H. (1989). Restricted unbiased iterative generalized least squares. *Biometrika , 76*, 622-623.

Goldstein, H. (2006). Subjective Bayesian analysis: Principle and practice. *Bayesian Analysis , 1(3)*, 403-420.

Goldstein, H., Kounali, D., & Robinson, A. (2008). Modelling measurement errors and category misclassifications in multilevel models. *Statistical Modelling , 8(3)*, 243-261.

Iversen, G. (1989). *Bayesian statistical inference.* Sage University papers.

Joreskog, K. (1970). A general method for analysis of covariance structures. *Biometrika , 57*, 239-251.

Lee, S. Y. (1981). A Bayesian approach to confirmatory factor analysis. *Psychometrika , 46*, 153-160.

Lee, S. Y., & Shi, J. Q. (2000). Joint Bayesian analysis of factor scores ans structural parameters in the factor analysis model. *Annals of the Institute of Statistical mathematics , 52*, 733-736.

Lee, S. Y., & Song, X. Y. (2004). Evaluation of the Bayesian and maximum likelihood approachs in analyzing structural equation model with small sample sizes. *Multivariate Behavioral Research , 39*, 653-686.

Martin, A. D., Quinn, K. M., & Park, J. H. (2011). MCMCpack: Markov chain Monte Carlo in R. *Journal of Statistical Software*, 1-21.

Martin, A. D., Quinn, K. M., & Park, J. H. (2012, June 14). *Markov chain Monte Carlo (MCMC) Package.* Retrieved 7 5, 2012, from The R Project for Statistical Computing: http://cran.r-project.org/web/packages/MCMCpack/MCMCpack.pdf

Martin, A. D., Quinn, K. M., & Park, J. H. (2012, June 14). *Package 'MCMCpack'.* Retrieved Aug 12, 2012, from R project: http://cran.r-project.org/web/packages/MCMCpack/MCMCpack.pdf

Mass, C., & Hox, J. J. (2001). *Robustness of multilevel parameter estimates against small sample sizes.* Department of Methodology and Statistics, Utrecht University .

Muthen, B., & Asparouhov, T. (2012). *New developments in Mplus Version 7: Part 1.* Retrieved August 28, 2012, from Mplus: http://www.statmodel.com/download/handouts/MuthenV7Part1.pdf

Muthen, B., & Asparouhov, T. (2012, August). *New developments in Mplus Version 7: Part2.* Retrieved August 28, 2012, from Mplus: http://www.statmodel.com/download/handouts/MuthenV7Part2.pdf

Muthen, B., & Asparouhov, T. (2012, August). *New developments in Mplus Version 7: Part3.* Retrieved August 28, 2012, from Mplus: http://www.statmodel.com/download/handouts/MuthenV7Part3.pdf

Muthén, L.K., and Muthén, B.O. (2010). *Mplus user's guide. 6th ed*. Los Angeles, CA: Muthén and Muthén.

Nounou, M., Bakshi, V., Goil, P., & Shen, X. (2001). *Process modeling by bayesian latent variable regression.* The Ohio State University, Columbus .

Paris, Q. (2004). *Robust estimators of errors-in-variables models Part I.* Department of Agricultural and Resource Economics University of California, Davis .

Peng, R. D., & Leeuw, J. d. (2002, August 28). *An introduction to the .C interface to R.* Retrieved March 5, 2009, from Johns Hopkins Bloomberg School of Public Health: http://www.biostat.jhsph.edu/~rpeng/docs/interface.pdf

Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D., & Almond, R. (2012, July 24). *Package 'coda'.* Retrieved July 30, 2012, from R project: http://cran.r-project.org/web/packages/coda/coda.pdf

R Development Core Team. (2004). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria .

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models applications and data analysis methods.* (2, Ed.) London: Sage Publications, Inc.

Ripley, B., Venables, B., & Hornik, K. (2012, 08 01). *Support Functions and Datasets for Venables and Riple's MASS.* Retrieved 08 13, 2012, from The R Project for Statistical Computing: http://cran.r-project.org/web/packages/MASS/MASS.pdf

Scineines, R., Hoijtink, H., & Boomsma, A. (1999). Bayesian esimation and testing of structural equation models. *Psychometrika , 64*, 37-52.

Snijder, T., & Bosker, R. (1999). *Multilevel analysis.* London: Sage Publications, Inc.

Song, X. Y., & Lee, S. Y. (2001). Bayesian estimation and test for factor analysis model with continuous and polytomous data in several population. *British Journal of Mathematical and Statistical Psychology , 54*, 237-263.

Srisuttiyakorn, S., & Kanjanawasee, S. (2009). Multilevel parameter estimates when level-one independent variables measure with error. *IMES.* Bangkok.

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software , 12(3)*, 1-16.

Terhorst, L. (n.d.). *A comparison of estimation methods when interaction is omitted from a multilevel model.* Ph.D. dissertation University of Pittsburgh, USA .

Ulrich, P. (2008). *Bayesian inference for latent variable models.* Cambridge,UK: University of Cambridge

Woodhouse, G. M. (1996). Adjusting for measurement error in multilevel analysis. *Journal of the Royal Statistical Society, A.*, 201-12.

ภาคผนวก

ภาคผนวก ก

1. คำสั่งที่ใช้ในศึกษาด้วยวิธีการจำลองแบบมอนติคาร์โลด้วยโปรแกรม R

1.1 คำสั่งสร้างข้อมูลตัวแปรแฝง

1.2 คำสั่งสร้างข้อมูลตัวแปรสังเกตได้

1.3 คำสั่งอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ (gibbs-sampling algorithm) กรณีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์สุ่ม

1.4 คำสั่งสรุปค่าพารามิเตอร์

2. คำสั่งที่ใช้ในศึกษาด้วยวิธีการจำลองแบบมอนติคาร์โลด้วยโปรแกรม Mplus

3. คำสั่งที่ใช้วิเคราะห์ข้อมูลจริงด้วยโปรแกรม Mplus

1. คำสั่งที่ใช้ในศึกษาด้วยวิธีการจำลองแบบมอนติคาร์โลด้วยโปรแกรม R

1.1 คำสั่งสร้างข้อมูลตัวแปรแฝง

```
generate.latents<-function(gamma.true,sigma2d.true,sigmaU.true,phi.true,nuj.true,n,J,q1,q2,q)
{
ksi<-matrix(nrow=n*J,ncol=q2)
ksij<-array(dim=c(n,q2,J))
eta<-matrix(nrow=n*J,ncol=q1)
etaj<-matrix(nrow=n,ncol=J)
dj<-matrix(nrow=n,ncol=J)
uj.true<-mvrnorm(J,mu=rep(0,q2),Sigma=sigmaU.true)

for (j in 1:J)
{
ksij[,,j]<-mvrnorm(n,mu=nuj.true[j,],Sigma=phi.true)
ksi[(1+(j-1)*n):(j*n),]<-ksij[,,j]
dj[,j]<-rnorm(n,0,sqrt(sigma2d.true))
etaj[,j]<-ksij[,,j]%*%as.matrix((gamma.true+uj.true[j,]),nrow=q2,ncol=1)+dj[,j]
eta[(1+(j-1)*n):(j*n),]<-etaj[,j]
}
output<-list(etaj,eta,ksij,ksi,uj.true,dj)
names(output)<-c("etaj.true","eta.true","ksij.true","ksi.true","uj.true","dj.true")
return(output)
}
```

1.2 คำสั่งสร้างข้อมูลตัวแปรสังเกตได้

```
generate.observed<-function(etaj.true,ksij.true,alpha.true,lambdaj.true,sigma2.epsj.true,n,J,p1,p2,p)
{
yj<-array(dim=c(p,n,J))
epsj<-array(dim=c(p,n,J))
```

```
wj<-array(dim=c(q,n,J))
for (j in 1:J)
{
for (i in 1:n)
{
epsj[,i,j]<-mvrnorm(1,mu=rep(0,p),Sigma=diag(sigma2.epsj.true[,j]))
wj[,i,j]<-c(etaj.true[i,j],ksij.true[i,,j])
yj[,i,j]<-alpha.true+lambdaj.true[,,j]%*%wj[,i,j]+epsj[,i,j]
}
}
output<-list(yj,epsj,wj)
names(output)<-c("yj","epsj","wj")
return(output)
}
```

1.3 คำสั่งอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ (gibbs-sampling algorithm) กรณีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์สุ่ม

```
random.loadings.mcmc<-function(iter,prior,initial,yj,lam.free,lam.fixvalue,iteration,n,J,q,q1,q2,p)
{
iteration<-iter
#measurement model parameters
alpha.iter<-matrix(nrow=iteration,ncol=p)
lambdaj.iter<-array(dim=c(p,q,J,iteration))
latent.iter<-array(dim=c(q,n,J,iteration))
nuj.iter<-array(dim=c(J,q2,iteration))
phi.iter<-array(dim=c(q2,q2,iteration))
sigma2.epsj.iter<-array(dim=c(p,J,iteration))
```

```
#strutural model parameters
gamma.iter<-matrix(nrow=iteration,ncol=q2)
sigma2d.iter<-matrix(nrow=iteration,ncol=1)
sigmaU.iter<-array(dim=c(q2,q2,iteration))

#prior
mu.alpha<-prior$mu.alpha
cov.alpha<-prior$cov.alpha
Hk<-prior$Hk
gk<-prior$gk
rho.phi<-prior$rho.phi
R.phi<-prior$R.phi
ak<-prior$ak
bk<-prior$bk
mu.gamma<-prior$mu.gamma
cov.gamma<-prior$cov.gamma
rho.u<-prior$rho.u
R.u<-prior$R.u
ad<-prior$a.d
bd<-prior$b.d
delta<-prior$delta

#initial
alpha<-initial$alpha
lambdaj<-initial$lambdaj
sigma2.epsj<-initial$sigma2.epsj
nuj<-initial$nuj
phi<-initial$phi
gamma<-initial$gamma
```

```
sigma2d<-initial$sigma2d
sigmaU<-initial$sigmaU
uj<-initial$uj
ksij<-initial$ksij
etaj<-initial$etaj
wj<-initial$wj
ksi<-initial$ksi
eta<-initial$eta
betaj<-gamma+uj

rep<-0
for (m in 1:iteration)
{
rep<-rep+1
cat("iteration=",rep,"\n")

#1. sampling alpha
sum1<-0
sum2<-0
for (j in 1:J)
{
sum1<-sum1+n*solve(diag(sigma2.epsj[,j]))
for (i in 1:n)
{
sum2<-sum2+solve(diag(sigma2.epsj[,j]))%*%(yj[,i,j]-lambdaj[,,j]%*%wj[,i,j])
}
}
cov.alpha.pos<-solve(solve(cov.alpha)+sum1)
mu.alpha.pos<-cov.alpha.pos%*%(cov.alpha%*%mu.alpha+sum2)
```

```
alpha<-mvrnorm(1,mu.alpha.pos,cov.alpha.pos)

#2. sampling factor loadings matrix
for (j in 1:J)
{
for (k in 1:p)
{
w.fix.kj<-colSums((1-lam.free[k,])*lambdaj[k,,j]*wj[,,j])
w.free.kj<-colSums((lam.free[k,])*wj[,,j])
y.adj.kj<-yj[k,,j]-w.fix.kj-alpha[k]
Hk.pos<-solve(solve(Hk)+solve(sigma2.epsj[k,j])*t(w.free.kj)%*%w.free.kj)
gk.pos<-Hk.pos%*%(solve(Hk)%*%gk+solve(sigma2.epsj[k,j])*t(w.free.kj)%*%y.adj.kj)
lambdaj[k,,j]<-mvrnorm(1,mu=gk.pos,Sigma=Hk.pos)*lam.free[k,]
}
lambdaj[,,j]<-lambdaj[,,j]+lam.fixvalue
}

#3. sampling latent variables
cov.wj<-matrix(nrow=q,ncol=q)
mu.wj<-matrix(nrow=q,ncol=1)
for (j in 1:J)
{
cov.wj[1,1]<-t(betaj[j,])%*%phi%*%(betaj[j,])+sigma2d
cov.wj[1,2:q]<-t(betaj[j,])%*%phi
cov.wj[2:q,1]<-phi%*%betaj[j,]
cov.wj[2:q,2:q]<-phi
mu.wj[1,1]<-t(betaj[j,])%*%nuj[j,]
mu.wj[2:q,1]<-nuj[j,]
cov.wij.pos<-solve(solve(cov.wj)+t(lambdaj[,,j])%*%solve(diag(sigma2.epsj[,j]))%*%lambdaj[,,j])
```

```
for (i in 1:n)
{
mu.wij.pos<-
cov.wij.pos%*%(solve(cov.wj)%*%mu.wj+t(lambdaj[,,j])%*%solve(diag(sigma2.epsj[,j]))%*%(yj[,i,j]-
alpha))
wj[,i,j]<-mvrnorm(1,mu=mu.wij.pos,Sigma=cov.wij.pos)
}
etaj[,j]<-wj[1,,j]
ksij[,,j]<-t(wj[2:q,,j]) # !!! ระวังถ้าไม่ใส่ transpose จะเรียงข้อมูลผิด
}

#4. sampling mean of exogenous latents
for (j in 1:J)
{
cov.nuj.pos<-solve(solve(delta)+n*solve(phi))
mu.nuj.pos<-cov.nuj.pos%*%solve(phi)%*%colSums(as.matrix(ksij[,,j],nrow=n,ncol=q2))
nuj[j,]<-mvrnorm(1,mu=mu.nuj.pos,Sigma=cov.nuj.pos)
}

#5. sampling inverse of phi
rho.pos<-rho.phi+n*J
sum<-0
for (j in 1:J)
{
for (i in 1:n)
{
sum<-sum+(ksij[i,,j]-nuj[j,])%*%t(ksij[i,,j]-nuj[j,])
}
}
```

```
R.pos<-solve(solve(R.phi)+sum)
phi<-solve(rwish(rho.pos,R.pos))

#6. sampling measurement error variances
for (j in 1:J)
{
ak.posj<-n/2+ak
for (k in 1:p)
{
sum<-t(yj[k,,j]-alpha[k]-t(wj[,,j])%*%lambdaj[k,,j])%*%(yj[k,,j]-alpha[k]-t(wj[,,j])%*%lambdaj[k,,j])
bk.posj<-(0.5*sum+bk)
sigma2.epsj[k,j]<-solve(rgamma(1,ak.posj,bk.posj))
}
}

#7. sampling level-1 random regression coefficients
for (j in 1:J)
{
cov.uj.pos<-solve(solve(sigmaU)+t(ksij[,,j])%*%ksij[,,j]/sigma2d)
mu.uj.pos<-cov.uj.pos%*%(solve(sigmaU)%*%gamma+t(ksij[,,j])%*%etaj[,j]/sigma2d)
betaj[j,]<-mvrnorm(1,mu=mu.uj.pos,Sigma=cov.uj.pos)
}
gamma<-colMeans(betaj) #fix effects parameters
for (j in 1:J)
{
uj[j,]<-betaj[j,]-gamma #level-2 residual
}
```

```
#8. sampling level-2 covariance of residuals
rho.u.pos<-rho.u+J
R.u.pos<-solve(solve(R.u)+t(uj)%*%uj)
sigmaU<-solve(rwish(rho.u.pos,R.u.pos))

#9. sampling level-1 residual variance
ad.pos<-n*J/2+ad
sum<-0
for (j in 1:J)
{
sum<-sum+t(etaj[,j]-ksij[,,j]%*%as.matrix(gamma+uj[j,],nrow=q2,ncol=1))%*%(etaj[,j]-
ksij[,,j]%*%as.matrix(gamma+uj[j,],nrow=q2,ncol=1))
}
bd.pos<-(bd+0.5*sum)
sigma2d<-rgamma(1,ad.pos,bd.pos)^(-1)
#collect mcmc chain
#measurement model
alpha.iter[m,]<-alpha
lambdaj.iter[,,,m]<-lambdaj
latent.iter[,,,m]<-wj
nuj.iter[,,m]<-nuj
phi.iter[,,m]<-phi
sigma2.epsj.iter[,,m]<-sigma2.epsj
#structural model
gamma.iter[m,]<-gamma
sigma2d.iter[m,]<-sigma2d
sigmaU.iter[,,m]<-sigmaU
} #end of MCMC iteration
```

```
output<-
list(alpha.iter,lambdaj.iter,latent.iter,nuj.iter,phi.iter,sigma2.epsj.iter,gamma.iter,sigma2d.iter,sigmaU.i
ter)
names(output)<-c("alpha","lambdaj","latent","nuj","phi","sigma2.epsj","gamma","sigma2d","sigmaU")
return(output)
} #end of function
```

1.4 คำสั่งสรุปค่าพารามิเตอร์

```
summary<-function(mcmc.output,burnin,thin,iteration)
{
remain.chain<-seq(burnin+1,iteration,thin)
pos.mean.alpha<-colMeans(mcmc.output$alpha[remain.chain,])
pos.sd.alpha<-apply(mcmc.output$alpha[remain.chain,],2,sd)
pos.mean.lambdaj<-array(dim=c(p,q,J))
pos.sd.lambdaj<-array(dim=c(p,q,J))
for (j in 1:J)
{
for (k in 1:p)
{
pos.mean.lambdaj[k,,j]<-rowMeans(mcmc.output$lambdaj[k,,j,remain.chain])
pos.sd.lambdaj[k,,j]<-apply(t(mcmc.output$lambdaj[k,,j,remain.chain]),2,sd)
}
}
pos.mean.latent<-matrix(nrow=n*J,ncol=q)
pos.sd.latent<-matrix(nrow=n*J,ncol=q)
for (j in 1:J)
{
for (i in 1:n)
```

```
{
pos.mean.latent[i+(j-1)*n,]<-colMeans(t(mcmc.output$latent[,i,j,remain.chain]))
pos.sd.latent[i+(j-1)*n,]<-apply(t(mcmc.output$latent[,i,j,remain.chain]),2,sd)
}
}
pos.mean.nuj<-matrix(nrow=J,ncol=q2)
pos.sd.nuj<-matrix(nrow=J,ncol=q2)
for (j in 1:J)
{
pos.mean.nuj[j,]<-
rowMeans(matrix(mcmc.output$nuj[j,,remain.chain],nrow=q2,ncol=length(remain.chain)))
pos.sd.nuj[j,]<-
apply(t(matrix(mcmc.output$nuj[j,,remain.chain],nrow=q2,ncol=length(remain.chain))),2,sd)
}
pos.mean.phi<-matrix(nrow=q2,ncol=q2)
pos.sd.phi<-matrix(nrow=q2,ncol=q2)
for (k in 1:q2)
{
for (s in 1:q2)
{
pos.mean.phi[k,s]<-
mean(matrix(mcmc.output$phi[k,s,remain.chain],nrow=1,ncol=length(remain.chain)))
pos.sd.phi[k,s]<-
apply(t(matrix(mcmc.output$phi[k,s,remain.chain],nrow=1,ncol=length(remain.chain))),2,sd)
}
}
pos.mean.sigma2.epsj<-matrix(nrow=p,ncol=J)
pos.sd.sigma2.epsj<-matrix(nrow=p,ncol=J)
for (j in 1:J)
```

```
{
pos.mean.sigma2.epsj[,j]<-rowMeans(mcmc.output$sigma2.epsj[,j,remain.chain])
pos.sd.sigma2.epsj[,j]<-apply(t(mcmc.output$sigma2.epsj[,j,remain.chain]),2,sd)
}
pos.mean.gamma<-
colMeans(matrix(mcmc.output$gamma[remain.chain,],nrow=remain.chain,ncol=q2))
pos.sd.gamma<-apply(mcmc.output$gamma[remain.chain,],2,sd)
pos.mean.sigma2d<-mean(mcmc.output$sigma2d[remain.chain])
pos.sd.sigma2d<-sd(mcmc.output$sigma2d[remain.chain])
pos.mean.sigmaU<-matrix(nrow=q2,ncol=q2)
pos.sd.sigmaU<-matrix(nrow=q2,ncol=q2)
for (k in 1:q2)
{
for (s in 1:q2)
{
pos.mean.sigmaU[k,s]<-
mean(matrix(mcmc.output$sigmaU[k,s,remain.chain],nrow=1,ncol=length(remain.chain)))
pos.sd.sigmaU[k,s]<-
apply(t(matrix(mcmc.output$sigmaU[k,s,remain.chain],nrow=1,ncol=length(remain.chain))),2,sd)
}
}
output<-
list(pos.mean.alpha,pos.sd.alpha,pos.mean.lambdaj,pos.sd.lambdaj,pos.mean.latent,pos.sd.latent,p
os.mean.nuj,pos.sd.nuj,pos.mean.phi,pos.sd.phi,pos.mean.sigma2.epsj,pos.sd.sigma2.epsj,pos.me
an.gamma,pos.sd.gamma,pos.mean.sigma2d,pos.sd.sigma2d,pos.mean.sigmaU,pos.sd.sigmaU)
names(output)<-
c("mean.alpha","sd.alpha","mean.lambdaj","sd.lambdaj","mean.latent","sd.latent","mean.nuj","sd.nuj","m
ean.phi","sd.phi","mean.sigma2.epsj","sd.sigma2.epsj","mean.gamma","sd.gamma","mean.sigma2d","s
d.sigma2d","mean.sigmaU","sd.sigmaU")
```

```
return(output)
}#end of summary function
```

2. คำสั่งที่ใช้ในศึกษาด้วยวิธีการจำลองแบบมอนติคาร์โลด้วยโปรแกรม Mplus

```
title: Monte carlo simulation for two-levels random coeffocients model with CFA
montecarlo:
   names are y1-y6;
   nobservations = 1500;
   nreps = 20;
   ncsizes = 1;
   csizes = 50(30);
seed=2854795;
model population:
   %within%
   eta by y1@1 y2*.8 y3*.6;
   ksi1 by y4@1 y5*.8 y6*.6;
   beta1j | eta on ksi1;
   [ksi1*0];
   ksi1*1;
   eta*.8;
    y1-y6*0.6666666667;
   %between%
   [beta1j*1];
   beta1j*0.2;
analysis: type=twolevel random;
      algorithm = integration;
model:
```

```
 %within%
 eta by y1-y3;
 ksi1 by y4-y6;
 beta1j | eta on ksi1;
 %between%
output: tech8 tech9;
```

3. คำสั่งที่ใช้วิเคราะห์ข้อมูลจริงด้วยโปรแกรม Mplus

```
TITLE: model
DATA: FILE IS "C:\Users\Siwachoat\Desktop\empiricaldata.dat";
variable: name are gpa sw1-sw6 sb1-sb4 clus;
cluster is clus;
within is sw1-sw6 gpa sb1-sb4;

analysis: type is twolevel random;
algorithm=integration;
estimator is MLR;

model:
%within%
SWB by sw1* sw2 sw3 sw4 sw5@1 sw6;
SIB by sb1* sb2@1 sb3 sb4;
s1 | SWB on gpa;
SWB on SIB;

%between%

output: sampstat standardized;
```

ภาคผนวก ข

อัลกอริทึมการสุ่มตัวอย่างแบบกิบสำหรับประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์องค์ประกอบที่พารามิเตอร์จุดตัดแกนและเมทริกซ์ความแปรปรวนร่วมของตัวแปรแฝงมีความผันแปรระหว่างกลุ่ม

ในงานวิจัยทางสังคมศาสตร์ จิตวิทยา หรือทางการศึกษา ในบางกรณีผู้วิจัยอาจสนใจให้พารามิเตอร์จุดตัดแกนมีความผันแปรระหว่างกลุ่ม และกำหนดโมเดลในระดับที่สองให้กับพารามิเตอร์จุดตัดแกน ทำให้สามารถรวมคะแนนพารามิเตอร์จุดตัดแกนขึ้นไปเป็นตัวแปรแฝงในระดับที่สองได้ รายละเอียดของโมเดลการวัดเป็นดังนี้

<u>โมเดลการวัดระดับที่หนึ่ง</u>

$$\underline{y}_{ij} = \underline{\mu}_j + \Lambda\underline{\omega}_{ij} + \underline{\epsilon}_{ij} \qquad (1\ ก)$$

<u>โมเดลการวัดระดับที่สอง</u>

$$\underline{\mu}_j = \underline{\kappa} + \Lambda_b\underline{\varphi}_j + \underline{e}_j \qquad (1\ ข)$$

ข้อสมมติเบื้องต้นของโมเดลกำหนดให้ $\underline{\epsilon}_{ij} \sim N\left(0, \Psi_{\epsilon j}\right)$, $\underline{\omega}_{ij} \sim N\left(\underline{0}, \Sigma_{\omega j}\right)$ โดยที่

$$\Sigma_{\omega j} = \begin{bmatrix} \underline{\beta}_j \Phi_j \underline{\beta}_j^T + \sigma_\delta^2 & \underline{\beta}_j^T \Phi_j \\ \underline{\beta}_j \Phi_j & \Phi_j \end{bmatrix} \qquad (2)$$

$\underline{e}_j \sim N\left(\underline{0}, \Psi_b\right)$ โดยที่ $\Psi_{b\epsilon}$ เป็นเมทริกซ์ความแปรปรวนร่วมทแยงมุมของความคลาดเคลื่อนจากการวัดในระดับที่สอง และ $\underline{\varphi}_j \sim N\left(\underline{0}, \Phi_b\right)$ จากข้อสมมติดังกล่าวจะได้ว่า $\underline{\mu}_j \sim N\left(\underline{\alpha}, \Sigma_b\right)$ โดยที่ $\Sigma_b = \Lambda_b \Phi_b \Lambda_b^T + \Psi_b$

จากสมการที่ (1) จะสามารถเขียนสมการรวมได้ดังนี้

$$\underline{y}_{ij} = \underline{\kappa} + \Lambda_b\underline{\varphi}_j + \Lambda\underline{\omega}_{ij} + \underline{e}_j + \underline{\epsilon}_{ij} \qquad (3)$$

จากข้อสมมติในข้างต้นจะได้ว่า

$$E\left(\underline{y}_{ij}\right) = \underline{\kappa} \qquad (4)$$

$$Cov\left(\underline{y}_{ij} | \Phi_j, \Psi_{\epsilon j}\right) = \Lambda_b \Phi_b \Lambda_b^T + \Lambda \Sigma_{\omega j} \Lambda^T + \Psi_b + \Psi_{\epsilon j} \qquad (5)$$

<u>การแจกแจงความน่าจะเป็นก่อนหน้า (prior distribution)</u>

การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ในโมเดลกำหนดดังต่อไปนี้

1. $\underline{\mu}_j \sim Normal\left(\underline{\kappa}, \Sigma_b\right)$
2. $\underline{\kappa} \sim Normal(m, C)$
3. $\underline{\Lambda}_{bk} \sim Normal\left(\underline{g}_{bk}, H_{bk}\right)$ โดยที่ $\underline{\Lambda}_{bk}$ คือเวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์น้ำหนักองค์ประกอบ Λ_b ที่เป็น free parameters
4. $\Phi_b^{-1} \sim Wishart(\rho_b, R_b)$
5. $\psi_{b\epsilon k}^{-1} \sim Gamma(\alpha_{bk}, \beta_{bk})$ โดยที่ $\psi_{b\epsilon k}$ คือสมาชิกในแนวทแยงมุมของเมทริกซ์ $\Psi_{b\epsilon}$

6. $\underline{\Lambda}_k \sim Normal\left(\underline{g}_k, H_k\right)$ โดยที่ $\underline{\Lambda}_k$ คือเวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์น้ำหนักองค์ประกอบ Λ ที่เป็น free parameters
7. $\Phi_j^{-1} \sim Wishart(\rho_{\phi j}, R_{\phi j})$
8. $\underline{\beta_j} \sim Normal\left(\underline{\gamma}, \Sigma_u\right)$
9. $\Sigma_u^{-1} \sim Wishart(\rho_u, R_u)$
10. $\left(\sigma_\delta^2\right)^{-1} \sim Gamma(\alpha_\delta, \beta_\delta)$

<u>อัลกอริทึมสำหรับประมาณค่าพารามิเตอร์ (Estimation Algorithm)</u>

ในทำนองเดียวกับวิธีการในกรณีแรก อัลกอริทึมที่ใช้ในการประมาณค่าพารามิเตอร์ในโมเดลประกอบไปด้วย 12 ขั้นตอนรายละเอียดเป็นดังต่อไปนี้

1. สำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j สุ่มพารามิเตอร์ $\underline{\kappa}$ จาก

$$\underline{\kappa} | \underline{m}, C, \underline{\mu}_j, \Sigma_b \sim N\left(\underline{\widehat{m}}, \hat{C}\right)$$

โดยที่ $\hat{C} = [C^{-1} + J\Sigma_b^{-1}]^{-1}$, $\underline{\widehat{m}} = \hat{C}\left[C^{-1}\underline{m} + \Sigma_b^{-1}\sum_{j=1}^{J}\underline{\mu}_j\right]$

และ $\Sigma_b = \Lambda_b \Phi_b \Lambda_b^T + \Psi_b$

2. สำหรับหน่วยตัวอย่างที่ j สุ่มพารามิเตอร์ $\underline{\mu_j}$ จาก

$$\underline{\mu}_j \mid \left\{\underline{y}_{ij}\right\}, \underline{\kappa}, \Sigma_b, \Lambda, \Phi_j, \Psi_{\epsilon j} \sim Normal\left(\underline{\hat{\kappa}}_j, \hat{\Sigma}_j\right)$$

โดยที่ $\hat{\Sigma}_j = \left[\Sigma_b^{-1} + n_j \Sigma_j^{-1}\right]^{-1}$, $\underline{\hat{\kappa}}_j = \hat{\Sigma}_j\left[\Sigma_b^{-1}\underline{\kappa} + \Sigma_j^{-1}\sum_{i=1}^{n_j}\underline{y}_{ij}\right]$

และ $\Sigma_j = \Lambda \Sigma_{\omega j} \Lambda^T + \Psi_{\epsilon j}$

3. สำหรับหน่วยที่ j สุ่มพารามิเตอร์ $\underline{\Lambda}_{bk}$ จาก

$$\underline{\Lambda}_{bk} | \underline{g}_{bk}, H_{bk}, \{\tilde{\mu}_{jk}\}, \left\{\underline{\varphi}_{-jk}\right\}, \psi_{bk} \sim Normal\left(\underline{\hat{g}}_{bk}, \hat{H}_{bk}\right)$$

เมื่อ $\underline{\Lambda}_{bk}$ คือ เวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์ Λ_b เฉพาะที่เป็น free parameters, $\underline{\tilde{\varphi}}_{jk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงในระดับที่สองของกลุ่มที่ j ที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น fix parameters ในแถวที่ k ของเมทริกซ์ Λ_b, $\underline{\varphi}_{-jk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงในระดับที่สองของกลุ่มที่ j ที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น free parameters ในแถวที่ k ของเมทริกซ์ Λ_b (กล่าวคือสอดคล้องกับเวกเตอร์ $\underline{\Lambda}_{bk}$) และ $\tilde{\mu}_{jk}$ คือ เวกเตอร์ของพารามิเตอร์จุดตัดของกลุ่มที่ j

และมีการปรับค่าเนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบใน แถวที่ k ของเมทริกซ์ $\underline{\Lambda}_b$ โดยที่ $\tilde{\mu}_{jk} = \mu_{jk} - \underline{1}^T \underline{\tilde{\varphi}}_{jk} - \kappa_k$

โดยที่ $\hat{H}_{bk} = \left[H_{bk}^{-1} + \frac{1}{\psi_{bk}} \sum_{j=1}^{J} \underline{\varphi}_{-jk} \underline{\varphi}_{-jk}^{T} \right]^{-1}$

และ $\underline{\hat{g}}_{bk} = \hat{H}_{bk} \left[H_{bk}^{-1} \underline{g}_{bk} + \frac{1}{\psi_{bk}} \underline{\Lambda}_{bk}^{T} \sum_{j=1}^{J} \underline{\varphi}_{-jk} \tilde{\mu}_{jk} \right]$

4. สำหรับหน่วยที่ j สุ่มตัวอย่าง $\underline{\varphi}_j$ จาก

$$\underline{\varphi}_j | \Phi_b, \underline{\mu}_j, \underline{\kappa}, \Lambda_b, \Psi_b \sim Normal\left(\underline{\hat{\varphi}}_j, \hat{\Phi}_b \right)$$

โดยที่ $\hat{\Phi}_b = [\Phi_b^{-1} + \Lambda_b^T \Psi_b^{-1} \Lambda_b]^{-1}$

และ $\underline{\hat{\varphi}}_j = \hat{\Phi}_b \Lambda_b^T \Psi_b^{-1} \left(\underline{\mu}_j - \underline{\kappa} \right)$

5. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Φ_b^{-1} จาก

$$\Phi_b^{-1} | \left\{ \underline{\hat{\varphi}}_j \right\} \sim Wishart(\hat{\rho}_b, \hat{R}_b)$$

โดยที่ $\hat{\rho}_b = J + \rho_b$ และ $\hat{R}_b = R_b^{-1} + \sum_{j=1}^{J} \underline{\hat{\varphi}}_j \; \underline{\hat{\varphi}}_j^T$

6. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ ψ_{bk}^{-1} จาก

$$\psi_{bk}^{-1} | \alpha_{bk}, \beta_{bk}, \underline{\kappa}, \Lambda_b, \left\{ \underline{\varphi}_j \right\}, \left\{ \underline{\mu}_j \right\} \sim Gamma(\hat{\alpha}_{bk}, \hat{\beta}_{bk})$$

โดยที่ $\hat{\alpha}_{bk} = \frac{J}{2} + \alpha_{bk}$

และ $\hat{\beta}_{bk} = \beta_{bk} + \frac{1}{2} \sum_{j=1}^{J} \left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T \underline{\varphi}_j \right)^T \left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T \underline{\varphi}_j \right)$

7. สำหรับหน่วยที่ j และตัวแปรสังเกตได้ที่ k สุ่มตัวอย่างพารามิเตอร์ $\underline{\Lambda}_k$ จาก

$$\underline{\Lambda}_k | \left\{ \underline{\tilde{y}}_{ijk} \right\}, \{ \underline{\omega}_{-ijk} \}, \psi_{\epsilon kj} \sim Normal(\underline{\hat{g}}_k, \hat{H}_k)$$

เมื่อ $\underline{\omega}_{-ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ และเป็น free parameters, $\underline{\tilde{\omega}}_{ijk}$ คือเวกเตอร์ของตัวแปรแฝงที่สอดคล้องกับน้ำหนักองค์ประกอบในแถวที่ k ของพารามิเตอร์เมทริกซ์ Λ และเป็น fix parameters, $\underline{\tilde{y}}_{ijk}$ เป็นค่าตัวแปรสังเกตได้ที่ปรับค่าที่เนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบในแถวที่ k พารามิเตอร์เมทริกซ์ Λ ซึ่ง $\underline{\tilde{y}}_{ijk} = y_{ijk} - \underline{1}^T \underline{\tilde{\omega}}_{ijk} - \mu_k$

โดยที่ $\widehat{H}_k = \left[H_k^{-1} + \sum_{j=1}^{J} \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\omega}_{-ijk}^{T}\right]^{-1}$

และ $\underline{\hat{g}}_k = \widehat{H}_k \left[H_k^{-1} \underline{g}_k + \sum_{j=1}^{J} \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\tilde{y}}_{ijk}\right]$

8. สำหรับหน่วยตัวอย่างที่ I ในหน่วยที่ j สุ่มตัวอย่างตัวแปรแฝง $\underline{\omega}_{ij}$ จาก

$$\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}_j, \Lambda \;\; , \Psi_{\epsilon j}, \underline{y}_{ij} \sim Normal\left(\underline{\hat{\theta}}_{ij}, \hat{\Sigma}_{\omega j}\right)$$

โดยที่ $\hat{\Sigma}_{\omega} = \left[\Sigma_{\omega j}^{-1} + \Lambda^{T} \Psi_{\epsilon j}^{-1} \Lambda \;\; \right]^{-1}$

และ $\underline{\hat{\theta}}_{ij} = \hat{\Sigma}_{\omega} \left[\Sigma_{\omega j}^{-1} \underline{\theta}_j + \Lambda^{T} \Psi_{\epsilon j}^{-1} \left(\underline{y}_{ij} - \underline{\mu}_j\right)\right]$

9. สำหรับหน่วยที่ j สุ่มตัวอย่าง Φ_j^{-1} จาก

$$\Phi_j^{-1} | \left\{\underline{\xi}_{ij}\right\}, \rho_{\phi j}, R_{\phi j} \sim Wishart\left(\hat{\rho}_{\phi j}, \hat{R}_{\phi j}\right)$$

โดยที่ $\hat{\rho}_{\phi j} = n_j + \rho_{\phi j}$

และ $\hat{R}_{\phi j} = \left[R_{\phi j}^{-1} + \sum_{i=1}^{n_j} \underline{\xi}_{ij} \underline{\xi}_{ij}^{T}\right]^{-1}$

11. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $\underline{\beta_j}$ จาก

$$\underline{\beta_j} | \{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_{\delta}^{2}, \Sigma_u \sim Normal\left(\underline{\hat{\beta}_j}, \hat{\Sigma}_{\beta j}\right)$$

โดยที่ $\hat{\Sigma}_{\beta j} = \left[\Sigma_u^{-1} + \frac{1}{\sigma_{\delta}^{2}} \sum_{i=1}^{n_j} \underline{\xi}_{ij} \underline{\xi}_{ij}^{T}\right]^{-1}$

และ $\underline{\hat{\beta}_j} = \hat{\Sigma}_{\beta j} \left[\Sigma_u^{-1} \underline{\gamma} + \frac{1}{\sigma_{\delta}^{2}} \sum_{i=1}^{n_j} \underline{\xi}_{ij} \eta_{ij}\right]$

จากขั้นตอนนี้จะสามารถประมาณค่าพารามิเตอร์อิทธิพลคงที่ได้จาก $\underline{\gamma} = \sum_{j=1}^{J} \underline{\beta_j}$

และเศษเหลือในระดับที่สองได้ $\underline{u}_j$ ได้จาก $\underline{u}_j = \underline{\beta_j} - \underline{\gamma}$

12. สำหรับหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ Σ_u^{-1} จาก

$$\Sigma_u^{-1} | \underline{u}_j, \rho_u, R_u \sim Wishart\left(\hat{\rho}_u, \hat{R}_u\right)$$

โดยที่ $\hat{\rho}_u = J + \rho_u$ และ $\hat{R}_u = \left[R_u^{-1} + \sum_{j=1}^{J} \underline{u}_j^{T} \underline{u}_j\right]^{-1}$

13. สำหรับหน่วยตัวอย่างที่ i และหน่วยที่ j สุ่มตัวอย่างพารามิเตอร์ $(\sigma_{\delta}^{2})^{-1}$ จาก

$$(\sigma_{\delta}^{2})^{-1} | \{\underline{\omega}_{ij}\}, \left\{\underline{\beta_j}\right\}, \{\underline{u}_j\}, \alpha_{\delta}, \beta_{\delta} \sim Gamma\left(\hat{\alpha}_{\delta}, \hat{\beta}_{\delta}\right)$$

โดยที่ $\hat{\alpha}_{\delta} = \frac{\sum_{j=1}^{J} n_j}{2} + \alpha_{\delta}$ และ $\hat{\beta}_{\delta} = \beta_{\delta} + \frac{1}{2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(\eta_{ij} - \underline{\beta_j}^{T} \underline{\xi}_{ij}\right)^{2}$

<u>รายละเอียดของการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขที่ใช้ในอัลกอริทึม</u>

จากการพิสูจน์ในทำนองเดียวกับกรณีแรก การแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ในแต่ละขั้นตอนของอัลกอริทึมการสุ่มตัวอย่างแบบกิบส์ในข้างต้นมีรายละเอียดดังต่อไปนี้

1. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์* $\underline{\kappa}$ *เมื่อกำหนด* $\underline{m}, C, \underline{\mu}_j, \Sigma_b$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\kappa}$ เมื่อกำหนด $\underline{m}, C, \underline{\mu}_j, \Sigma_b$ คือ

$$p\left(\underline{\kappa}|\underline{m}, C, \underline{\mu}_j, \Sigma_b\right) \propto p\left(\underline{\kappa}|\underline{m}, C\right) \times p\left(\left\{\underline{\mu}_j\right\} | \underline{\kappa}, \Sigma_b\right) \tag{6}$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\underline{\kappa}$ คือ

$$p\left(\underline{\kappa}|\underline{m}, C\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\kappa} - \underline{m}\right)^T C^{-1}\left(\underline{\kappa} - \underline{m}\right)\right\} \tag{7}$$

และฟังก์ชันภาวะความควรจะเป็นของ $\underline{\mu}_j$ เมื่อกำหนดพารามิเตอร์ $\underline{\kappa}, \Sigma_b$ คือ

$$\begin{aligned} p\left(\left\{\underline{\mu}_j\right\} | \underline{\kappa}, \Sigma_b\right) &\propto \prod_{j=1}^{J} exp\left\{-\frac{1}{2}\left(\underline{\mu}_j - \underline{\kappa}\right)^T \Sigma_b^{-1}\left(\underline{\mu}_j - \underline{\kappa}\right)\right\} \\ &= exp\left\{-\frac{1}{2}\sum_{j=1}^{J}\left(\underline{\mu}_j - \underline{\kappa}\right)^T \Sigma_b^{-1}\left(\underline{\mu}_j - \underline{\kappa}\right)\right\} \end{aligned} \tag{8}$$

แทนสมการที่ (7) และ (8) ลงในสมการที่ (6) จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\kappa}$ *เมื่อกำหนด* $\underline{m}, C, \underline{\mu}_j, \Sigma_b$ คือ

$$\begin{aligned} &p\left(\underline{\kappa}|\underline{m}, C, \underline{\mu}_j, \Sigma_b\right) \\ &\propto exp\left\{-\frac{1}{2}\left(\underline{\kappa} - \underline{m}\right)^T C^{-1}\left(\underline{\kappa} - \underline{m}\right)\right\} \times exp\left\{-\frac{1}{2}\sum_{j=1}^{J}\left(\underline{\mu}_j - \underline{\kappa}\right)^T \Sigma_b^{-1}\left(\underline{\mu}_j - \underline{\kappa}\right)\right\} \\ &= exp\left\{-\frac{1}{2}\left[\underline{\kappa}^T\left(C^{-1} + J\Sigma_b^{-1}\right)\underline{\kappa} - 2\underline{\kappa}^T\left(C^{-1}\underline{m} + \Sigma_b^{-1}\sum_{j=1}^{J}\underline{\mu}_j\right)\right]\right\} \end{aligned} \tag{9}$$

เปรียบเทียบสมการ (9) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบปกติจะได้ว่า

$$\underline{\kappa}|\underline{m}, C, \underline{\mu}_j, \Sigma_b \sim Normal\left(\hat{\underline{m}}, \hat{C}\right) \tag{10}$$

โดยที่ $\hat{C} = \left[C^{-1} + J\Sigma_b^{-1}\right]^{-1}$, $\hat{\underline{m}} = \hat{C}\left[C^{-1}\underline{m} + \Sigma_b^{-1}\sum_{j=1}^{J}\underline{\mu}_j\right]$

และ $\Sigma_b = \Lambda_b \Phi_b \Lambda_b^T + \Psi_b$

2. *การแจกแจงความน่าจะเป็นของพารามิเตอร์* $\underline{\mu}_j$ *เมื่อกำหนด* $\left\{\underline{y}_{ij}\right\}, \left\{\underline{\omega}_{ij}\right\}, \underline{\kappa}, \Sigma_b, \Lambda, \Phi_j, \Psi_{\epsilon j}$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\mu}_j$ เมื่อกำหนด $\left\{\underline{y}_{ij}\right\}, \{\underline{\omega}_{ij}\}, \underline{\kappa}, \Sigma_b, \Lambda, \Phi_j, \Psi_{\epsilon j}$ คือ

$$p\left(\underline{\mu}_j \mid \left\{\underline{y}_{ij}\right\}, \{\underline{\omega}_{ij}\}, \underline{\kappa}, \Sigma_b, \Lambda, \Phi_j, \Psi_{\epsilon j}\right) \propto p\left(\underline{\mu}_j | \underline{\kappa}, \Sigma_b\right) \times p\left(\left\{\underline{y}_{ij}\right\} | \underline{\mu}_j, \Lambda, \Phi_j, \Psi_{\epsilon j}\right) \quad (11)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\underline{\mu}_j$ คือ

$$p\left(\underline{\mu}_j | \underline{\kappa}, \Sigma_b\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\mu}_j - \underline{\kappa}\right)^T \Sigma_b^{-1}\left(\underline{\mu}_j - \underline{\kappa}\right)\right\} \quad (12)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\left\{\underline{y}_{ij}\right\}$ เมื่อกำหนดพารามิเตอร์ $\underline{\mu}_j, \Lambda, \Phi_j, \Psi_{\epsilon j}$ คือ

$$p\left(\left\{\underline{y}_{ij}\right\} | \underline{\mu}_j, \{\underline{\omega}_{ij}\}, \Lambda, \Phi_j, \Psi_{\epsilon j}\right) \propto \prod_{i=1}^{n_j} p\left(\underline{y}_{ij} | \underline{\mu}_j, \underline{\omega}_{ij}, \Lambda, \Phi_j, \Psi_{\epsilon j}\right)$$
$$= exp\left\{-\frac{1}{2}\sum_{i=1}^{n_j}\left(\underline{y}_{ij} - \underline{\mu}_j - \Lambda\underline{\omega}_{ij}\right)^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}_j - \Lambda\underline{\omega}_{ij}\right)\right\} \quad (13)$$

แทนสมการที่ (12) และสมการที่ (13) ลงในสมการที่ (11) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออก จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ $\underline{\mu}_j$ เมื่อกำหนด $\left\{\underline{y}_{ij}\right\}, \underline{\kappa}, \Sigma_b, \Lambda, \Phi_j, \Psi_{\epsilon j}$ คือ

$$p\left(\underline{\mu}_j \mid \left\{\underline{y}_{ij}\right\}, \underline{\kappa}, \Sigma_b, \Lambda, \Phi_j, \Psi_{\epsilon j}\right)$$
$$\propto exp\left\{-\frac{1}{2}\left(\underline{\mu}_j - \underline{\kappa}\right)^T \Sigma_b^{-1}\left(\underline{\mu}_j - \underline{\kappa}\right)\right\} \times exp\left\{-\frac{1}{2}\sum_{i=1}^{n_j}\left(\underline{y}_{ij} - \underline{\mu}_j\right)^T \Sigma_j^{-1}\left(\underline{y}_{ij} - \underline{\mu}_j\right)\right\}$$
$$= exp\left\{-\frac{1}{2}\left[\underline{\mu}_j^T\left(\Sigma_b^{-1} + n_j\Sigma_j^{-1}\right)\underline{\mu}_j - 2\underline{\mu}_j^T\left(\Sigma_b^{-1}\underline{\kappa} + \Sigma_j^{-1}\sum_{i=1}^{n_j}\underline{y}_{ij}\right)\right]\right\} \quad (14)$$

โดยที่ $\Sigma_j = \Lambda\Sigma_{\omega j}\Lambda^T + \Psi_{\epsilon j}$

เมื่อเปรียบเทียบสมการที่ (14) กับรูปแบบการแจกแจงความน่าจะเป็นแบบปกติจะได้ว่า

$$\underline{\mu}_j \mid \left\{\underline{y}_{ij}\right\}, \underline{\kappa}, \Sigma_b, \Lambda, \Phi_j, \Psi_{\epsilon j} \sim Normal\left(\underline{\hat{\kappa}}_j, \hat{\Sigma}_j\right) \quad (15)$$

โดยที่

$\hat{\Sigma}_j = \left[\Sigma_b^{-1} + n_j\Sigma_j^{-1}\right]^{-1}$, $\underline{\hat{\kappa}}_j = \hat{\Sigma}_j\left[\Sigma_b^{-1}\underline{\kappa} + \Sigma_j^{-1}\sum_{i=1}^{n_j}\underline{y}_{ij}\right]$ และ $\Sigma_j = \Lambda\Sigma_{\omega j}\Lambda^T + \Psi_{\epsilon j}$

3. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ของพารามิเตอร์ Λ_b เมื่อกำหนด $\left\{\underline{\varphi}_j\right\}, \underline{\mu}_j, \Psi_b$

เนื่องจากพารามิเตอร์ Λ_b เป็นเมทริกซ์การสุ่มตัวอย่างพารามิเตอร์จากการแจกแจงของเมทริกซ์โดยตรงนั้นเป็นสิ่งที่มีความยุ่งยาก ในทำนองเดียวกับการพิสูจน์ในกรณีแรกผู้วิจัยเลือกสุ่ม

ตัวอย่างจากเวกเตอร์ของพารามิเตอร์น้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์ Λ_b ทีละเวกเตอร์แทน จากแนวคิดในข้างต้นจึงกำหนดสัญลักษณ์ดังต่อไปนี้

$\underline{\Lambda}_{bk}$ คือ เวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์ Λ_b เฉพาะที่เป็น free parameters

$\underline{\tilde{\varphi}}_{jk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงในระดับที่สองของกลุ่มที่ j ที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น fix parameters ในแถวที่ k ของเมทริกซ์ Λ_b

$\underline{\varphi}_{-jk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงในระดับที่สองของกลุ่มที่ j ที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น free parameters ในแถวที่ k ของเมทริกซ์ Λ_b (กล่าวคือสอดคล้องกับเวกเตอร์ $\underline{\Lambda}_{bk}$)

$\tilde{\mu}_{jk}$ คือ เวกเตอร์ของพารามิเตอร์จุดตัดของกลุ่มที่ j และมีการปรับค่าเนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์ $\underline{\Lambda}_b$ โดยที่

$$\tilde{\mu}_{jk} = \mu_{jk} - \underline{1}^T \underline{\tilde{\varphi}}_{jk} - \kappa_k$$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\Lambda}_{bk}$ เมื่อกำหนด $\tilde{\mu}_{jk}, \underline{\varphi}_{-jk}, \psi_{bk}$ คือ

$$p\left(\underline{\Lambda}_{bk} | \underline{g}_{bk}, H_{bk}, \{\tilde{\mu}_{jk}\}, \left\{\underline{\varphi}_{-jk}\right\}, \psi_{bk}\right) \propto p\left(\underline{\Lambda}_{bk} | \underline{g}_{bk}, H_{bk}\right) \times p\left(\{\tilde{\mu}_{jk}\} | \underline{\Lambda}_{bk}, \left\{\underline{\varphi}_{-jk}\right\}, \psi_{bk}\right) \tag{15}$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\underline{\Lambda}_{bk}$ คือ

$$p\left(\underline{\Lambda}_{bk} | \underline{g}_{bk}, H_{bk}\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\Lambda}_{bk} - \underline{g}_{bk}\right)^T H_{bk}^{-1} \left(\underline{\Lambda}_{bk} - \underline{g}_{bk}\right)\right\} \tag{16}$$

และฟังก์ชันภาวะความควรจะเป็นของ $\tilde{\mu}_{jk}$ เมื่อกำหนด $\underline{\Lambda}_{bk}, \underline{\varphi}_{-jk}, \psi_{bk}$ คือ

$$p\left(\{\tilde{\mu}_{jk}\} | \underline{\Lambda}_{bk}, \left\{\underline{\varphi}_{-jk}\right\}, \psi_{bk}\right) \propto \prod_{j=1}^{J} p\left(\tilde{\mu}_{jk} | \underline{\Lambda}_{bk}, \underline{\varphi}_{-jk}, \psi_{bk}\right)$$
$$= \prod_{j=1}^{J} exp\left\{-\frac{1}{2\psi_{bk}}\left(\tilde{\mu}_{jk} - \underline{\varphi}_{-jk}^T \underline{\Lambda}_{bk}\right)^T \left(\tilde{\mu}_{jk} - \underline{\varphi}_{-jk}^T \underline{\Lambda}_{bk}\right)\right\}$$
$$= exp\left\{-\frac{1}{2\psi_{bk}} \sum_{j=1}^{J} \left(\tilde{\mu}_{jk} - \underline{\varphi}_{-jk}^T \underline{\Lambda}_{bk}\right)^T \left(\tilde{\mu}_{jk} - \underline{\varphi}_{-jk}^T \underline{\Lambda}_{bk}\right)\right\} \tag{17}$$

แทนสมการที่ (16) และสมการที่ (17) ลงในสมการที่ (15) ทำการกระจายพจน์ที่เป็นกำลังสองสมบูรณ์และตัดค่าคงที่ออก จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\Lambda}_{bk}$ เมื่อกำหนด $\underline{g}_{bk}, H_{bk}, \{\tilde{\mu}_{jk}\}, \left\{\underline{\varphi}_{-jk}\right\}, \psi_{bk}$ คือ

$$p\left(\underline{\Lambda}_{bk}|\underline{g}_{bk},H_{bk},\{\tilde{\mu}_{jk}\},\left\{\underline{\varphi}_{-jk}\right\},\psi_{bk}\right)$$
$$\propto exp\left\{-\frac{1}{2}\left[\underline{\Lambda}_{bk}^T\left(H_{bk}^{-1}+\frac{1}{\psi_{bk}}\sum_{j=1}^{J}\underline{\varphi}_{-jk}\,\underline{\varphi}_{-jk}^T\right)\underline{\Lambda}_{bk}-2\underline{\Lambda}_{bk}^T\left(H_{bk}^{-1}\underline{g}_{bk}+\frac{1}{\psi_{bk}}\underline{\Lambda}_{bk}^T\sum_{j=1}^{J}\underline{\varphi}_{-jk}\,\tilde{\mu}_{jk}\right)\right]\right\} \quad (18)$$

เปรียบเทียบสมการที่ (18) กับรูปแบบการแจกแจงความน่าจะเป็นแบบปกติจะได้ว่า

$$\underline{\Lambda}_{bk}|\underline{g}_{bk},H_{bk},\{\tilde{\mu}_{jk}\},\left\{\underline{\varphi}_{-jk}\right\},\psi_{bk}\sim Normal\left(\underline{\hat{g}}_{bk},\widehat{H}_{bk}\right) \quad (19)$$

โดยที่ $\widehat{H}_{bk}=\left[H_{bk}^{-1}+\frac{1}{\psi_{bk}}\sum_{j=1}^{J}\underline{\varphi}_{-jk}\,\underline{\varphi}_{-jk}^T\right]^{-1}$

และ $\underline{\hat{g}}_{bk}=\widehat{H}_{bk}\left[H_{bk}^{-1}\underline{g}_{bk}+\frac{1}{\psi_{bk}}\underline{\Lambda}_{bk}^T\sum_{j=1}^{J}\underline{\varphi}_{-jk}\,\tilde{\mu}_{jk}\right]$

4. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของ $\underline{\varphi}_j$ เมื่อกำหนด$\Phi_b,\underline{\mu}_j,\underline{\kappa},\Lambda_b,\Psi_b$*

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของ $\underline{\varphi}_j$ เมื่อกำหนด$\Phi_b,\underline{\mu}_j,\underline{\kappa},\Lambda_b,\Psi_b$ คือ

$$p\left(\underline{\varphi}_j|\Phi_b,\underline{\mu}_j,\underline{\kappa},\Lambda_b,\Psi_b\right)\propto p\left(\underline{\varphi}_j|\Phi_b\right)\times p\left(\underline{\mu}_j|\underline{\varphi}_j,\underline{\kappa},\Lambda_b,\Psi_b\right) \quad (30)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของ $\underline{\varphi}_j$ คือ

$$p\left(\underline{\varphi}_j|\Phi_b\right)\propto exp\left\{-\frac{1}{2}\underline{\varphi}_j^T\Phi_b^{-1}\underline{\varphi}_j\right\} \quad (31)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\underline{\mu}_j$ เมื่อกำหนด $\underline{\varphi}_j,\underline{\kappa},\Lambda_b,\Psi_b$ คือ

$$p\left(\underline{\mu}_j|\underline{\varphi}_j,\underline{\kappa},\Lambda_b,\psi_{bk}\right)\propto exp\left\{-\frac{1}{2}\left(\underline{\mu}_j-\underline{\kappa}-\Lambda_b\underline{\varphi}_j\right)^T\Psi_b^{-1}\left(\underline{\mu}_j-\underline{\kappa}-\Lambda_b\underline{\varphi}_j\right)\right\} \quad (32)$$

แทนสมการที่ (31) และสมการที่ (32) ลงในสมการที่ (30) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออก จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังของ $\underline{\varphi}_j$ เมื่อกำหนด $\Phi_b,\underline{\mu}_j,\underline{\kappa},\Lambda_b,\Psi_b$ คือ

$$p\left(\underline{\varphi}_j|\Phi_b,\underline{\mu}_j,\underline{\kappa},\Lambda_b,\Psi_b\right)\propto exp\left\{-\frac{1}{2}\left[\underline{\varphi}_j^T(\Phi_b^{-1}+\Lambda_b^T\Psi_b^{-1}\Lambda_b)\underline{\varphi}_j-2\underline{\varphi}_j^T\Lambda_b^T\Psi_b^{-1}\left(\underline{\mu}_j-\underline{\kappa}\right)\right]\right\} \quad (33)$$

เปรียบเทียบสมการที่ (33) กับรูปแบบของการแจกแจงแบบปกติจะได้ว่า

$$\underline{\varphi}_j|\Phi_b,\underline{\mu}_j,\underline{\kappa},\Lambda_b,\Psi_b\sim Normal\left(\underline{\hat{\varphi}}_j,\widehat{\Phi}_b\right) \quad (34)$$

โดยที่ $\widehat{\Phi}_b=[\Phi_b^{-1}+\Lambda_b^T\Psi_b^{-1}\Lambda_b]^{-1}$

และ $\underline{\hat{\varphi}}_j=\widehat{\Phi}_b\Lambda_b^T\Psi_b^{-1}\left(\underline{\mu}_j-\underline{\kappa}\right)$

5. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของ* $\boldsymbol{\Phi_b^{-1}}$ *เมื่อกำหนด* $\left\{\underline{\hat{\varphi}}_j\right\}$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของ $\boldsymbol{\Phi_b^{-1}}$ เมื่อกำหนด $\left\{\underline{\hat{\varphi}}_j\right\}$ คือ

$$p\left(\Phi_b^{-1}|\left\{\underline{\hat{\varphi}}_j\right\}\right) \propto p(\Phi_b^{-1}) \times p\left(\left\{\underline{\hat{\varphi}}_j\right\}|\Phi_b\right) \quad (35)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\boldsymbol{\Phi_b}$ คือ

$$p(\Phi_b^{-1}) \propto |\Phi_b^{-1}|^{\frac{\rho_b-q'-1}{2}} exp\left\{-\frac{1}{2}tr(R_b^{-1}\Phi_b^{-1})\right\} \quad (36)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\left\{\underline{\hat{\varphi}}_j\right\}$ เมื่อกำหนด $\boldsymbol{\Phi_b}$ คือ

$$\begin{aligned} p\left(\left\{\underline{\hat{\varphi}}_j\right\}|\Phi_b\right) &\propto \textstyle\prod_{j=1}^{J} p\left(\underline{\hat{\varphi}}_j|\Phi_b\right) \\ &= \textstyle\prod_{j=1}^{J}|\Phi_b|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\underline{\hat{\varphi}}_j^T\Phi_b^{-1}\underline{\hat{\varphi}}_j\right\} \\ &= |\Phi_b|^{-\frac{J}{2}} exp\left\{-\frac{1}{2}\textstyle\sum_{j=1}^{J}\underline{\hat{\varphi}}_j^T\Phi_b^{-1}\underline{\hat{\varphi}}_j\right\} \end{aligned} \quad (37)$$

แทนสมการที่ (36) และสมการที่ (37) ลงในสมการที่ (35) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ $\boldsymbol{\Phi_b^{-1}}$ เมื่อกำหนด $\left\{\underline{\hat{\varphi}}_j\right\}$ คือ

$$\begin{aligned} p\left(\Phi_b^{-1}|\left\{\underline{\hat{\varphi}}_j\right\}\right) &\propto |\Phi_b^{-1}|^{\frac{J+\rho_b-q'-1}{2}} exp\left\{-\frac{1}{2}\left[tr(R_b^{-1}\Phi_b^{-1}) + \textstyle\sum_{j=1}^{J}\underline{\hat{\varphi}}_j^T\Phi_b^{-1}\underline{\hat{\varphi}}_j\right]\right\} \\ &= |\Phi_b^{-1}|^{\frac{J+\rho_b-q'-1}{2}} exp\left\{-\frac{1}{2}\left[tr(\Phi_b^{-1}R_b^{-1}) + tr\left(\Phi_b^{-1}\textstyle\sum_{j=1}^{J}\underline{\hat{\varphi}}_j\ \underline{\hat{\varphi}}_j^T\right)\right]\right\} \\ &= |\Phi_b^{-1}|^{\frac{J+\rho_b-q'-1}{2}} exp\left\{-\frac{1}{2}tr\left(\Phi_b^{-1}\left[R_b^{-1} + \textstyle\sum_{j=1}^{J}\underline{\hat{\varphi}}_j\ \underline{\hat{\varphi}}_j^T\right]\right)\right\} \end{aligned} \quad (38)$$

เปรียบเทียบสมการที่ (38) กับรูปแบบการแจกแจงความน่าจะเป็นแบบ Wishart จะได้ว่า

$$\Phi_b^{-1}|\left\{\underline{\hat{\varphi}}_j\right\} \sim Wishart(\hat{\rho}_b, \hat{R}_b) \quad (39)$$

โดยที่ $\hat{\rho}_b = J + \rho_b$ และ $\hat{R}_b = R_b^{-1} + \sum_{j=1}^{J}\underline{\hat{\varphi}}_j\ \underline{\hat{\varphi}}_j^T$

6. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์* $\boldsymbol{\Psi_b^{-1}}$ *เมื่อกำหนด* $\alpha_{bk}, \beta_{bk}, \underline{\kappa}, \Lambda_b, \left\{\underline{\varphi}_j\right\}, \left\{\underline{\mu}_j\right\}$

เนื่องจาก Ψ_b เป็นเมทริกซ์ทแยงมุมซึ่งมีสมาชิกในแนวทแยงมุมคือ ψ_{bk} และ ψ_{bk} เป็นอิสระกับ ψ_{bh} โดยที่ $k \neq h$ ดังนั้นการสุ่มตัวอย่างเมทริกซ์ Ψ_b^{-1} สามารถกระทำได้โดยการสุ่มตัวอย่าง ψ_{bk}^{-1} จากการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขทีละตัว จากทฤษฎีของเบส์จะได้ว่า

การแจกแจงความน่าจะเป็นภายหลังของพารามิเตอร์ ψ_{bk}^{-1} เมื่อกำหนดพารามิเตอร์ $\alpha_{bk}, \beta_{bk}, \kappa_k, \underline{\Lambda}_{bk}, \left\{\underline{\varphi}_j\right\}, \left\{\underline{\mu}_j\right\}$ คือ

$$p\left(\psi_{bk}^{-1} | \alpha_{bk}, \beta_{bk}, \kappa_k, \underline{\Lambda}_{bk}, \left\{\underline{\varphi}_j\right\}, \left\{\underline{\mu}_j\right\}\right) \propto p(\psi_{bk}^{-1} | \alpha_{bk}, \beta_{bk}) \times p\left(\left\{\underline{\mu}_j\right\} | \kappa_k, \underline{\Lambda}_{bk}, \left\{\underline{\varphi}_j\right\}, \psi_{bk}\right) \quad (40)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ ψ_{bk}^{-1} คือ

$$p(\psi_{bk}^{-1} | \alpha_{bk}, \beta_{bk}) \propto (\psi_{bk}^{-1})^{\alpha_{bk}-1} exp\left\{-\beta_{bk}\psi_{bk}^{-1}\right\} \quad (41)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\left\{\underline{\mu}_j\right\}$ เมื่อกำหนด $\kappa_k, \underline{\Lambda}_{bk}, \left\{\underline{\varphi}_j\right\}$

$$p\left(\left\{\underline{\mu}_j\right\} | \kappa_k, \underline{\Lambda}_{bk}, \left\{\underline{\varphi}_j\right\}, \psi_{bk}\right) \propto \prod_{j=1}^{J} p\left(\underline{\mu}_j | \kappa_k, \underline{\Lambda}_{bk}, \underline{\varphi}_j, \psi_{bk}\right)$$

$$= \prod_{j=1}^{J} (\psi_{bk})^{-\frac{1}{2}} exp\left\{-\frac{1}{2\psi_{bk}}\left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)^T \left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)\right\}$$

$$= (\psi_{bk})^{-\frac{J}{2}} exp\left\{-\frac{1}{2\psi_{bk}}\sum_{j=1}^{J}\left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)^T \left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)\right\} \quad (42)$$

แทนค่าสมการที่ (41) กับสมการที่ (42) ลงในสมการที่ (40) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออก จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ ψ_{bk}^{-1} เมื่อกำหนด $\alpha_{bk}, \beta_{bk}, \underline{\kappa}, \Lambda_b, \left\{\underline{\varphi}_j\right\}, \left\{\underline{\mu}_j\right\}$ คือ

$$p\left(\psi_{bk}^{-1} | \alpha_{bk}, \beta_{bk}, \underline{\kappa}, \Lambda_b, \left\{\underline{\varphi}_j\right\}, \left\{\underline{\mu}_j\right\}\right)$$

$$\propto (\psi_{bk}^{-1})^{\frac{J}{2}+\alpha_{bk}-1} exp\left\{-\psi_{bk}^{-1}\left[\beta_{bk} + \frac{1}{2}\sum_{j=1}^{J}\left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)^T \left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)\right]\right\}$$

(43)

เมื่อเปรียบเทียบสมการที่ (43) กับรูปแบบการแจกแจงความน่าจะเป็นแบบแกมมาจะได้ว่า

$$\psi_{bk}^{-1} | \alpha_{bk}, \beta_{bk}, \underline{\kappa}, \Lambda_b, \left\{\underline{\varphi}_j\right\}, \left\{\underline{\mu}_j\right\} \sim Gamma\left(\hat{\alpha}_{bk}, \hat{\beta}_{bk}\right) \quad (44)$$

โดยที่ $\hat{\alpha}_{bk} = \frac{J}{2} + \alpha_{bk}$ และ $\hat{\beta}_{bk} = \beta_{bk} + \frac{1}{2}\sum_{j=1}^{J}\left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)^T \left(\underline{\mu}_j - \kappa_k - \underline{\Lambda}_{bk}^T\underline{\varphi}_j\right)$

7. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ Λ เมื่อกำหนด* $\left\{\underline{y}_{ij}\right\}, \{\underline{\omega}_{ij}\}, \{\Psi_{\epsilon j}\}$

ในทำนองเดียวกันกำหนดให้

$\underline{\Lambda}_k$ คือ เวกเตอร์ของน้ำหนักองค์ประกอบในแถวที่ k ของเมทริกซ์ Λ เฉพาะที่เป็น free parameters

$\underline{\widetilde{\omega}}_{ijk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น fix parameters ในแถวที่ k ของเมทริกซ์ Λ

$\underline{\omega}_{-ijk}$ คือ เวกเตอร์ของคะแนนตัวแปรแฝงที่สอดคล้องกับค่าน้ำหนักองค์ประกอบที่เป็น free parameters ในแถวที่ k ของเมทริกซ์ Λ (กล่าวคือสอดคล้องกับเวกเตอร์ $\underline{\Lambda}_k$)

$\underline{\tilde{y}}_{ijk}$ คือ เวกเตอร์ของคะแนนตัวแปรสังเกตได้ที่มีการปรับค่าเนื่องจากการกำหนดข้อจำกัดให้กับพารามิเตอร์น้ำหนักองค์ประกอบในเวกเตอร์ $\underline{\Lambda}_{kj}$ โดยที่

$$\underline{\tilde{y}}_{ijk} = y_{ijk} - \underline{1}^T \underline{\widetilde{\omega}}_{ijk} - \mu_k$$

พิจารณาการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\Lambda}_k$ เมื่อกำหนด $\left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}$ จากทฤษฎีของเบส์จะได้ว่า

$$p\left(\underline{\Lambda}_k | \left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}, \underline{g}_k, H_k\right) \propto$$
$$p\left(\underline{\Lambda}_k | \underline{g}_k, H_k\right) p\left(\left\{\underline{\tilde{y}}_{ijk}\right\} | \underline{\Lambda}_k, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}\right) \quad (45)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ $\underline{\Lambda}_k$ มีการแจกแจงแบบปกติ กล่าวคือ $\underline{\Lambda}_k \sim Normal\left(\underline{g}_k, H_k\right)$ ซึ่งเขียนได้ดังนี้

$$p\left(\underline{\Lambda}_k | \underline{g}_k, H_k\right) \propto exp\{-\tfrac{1}{2}\left(\underline{\Lambda}_k - \underline{g}_k\right)^T H_k^{-1} \left(\underline{\Lambda}_k - \underline{g}_k\right)\} \quad (46)$$

และฟังก์ชันภาวะความควรจะเป็นคือ

$$p\left(\left\{\underline{\tilde{y}}_{ijk}\right\} | \underline{\Lambda}_k, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}\right) \propto \prod_{j=1}^{J} \prod_{i=1}^{n_j} p\left(\underline{\tilde{y}}_{ijk} | \underline{\Lambda}_k, \underline{\omega}_{-ijk}, \psi_{\epsilon kj}\right)$$
$$= \prod_{j=1}^{J} \prod_{i=1}^{n_j} exp\left\{-\frac{1}{2\psi_{\epsilon kj}}\left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T \underline{\Lambda}_k\right)^T \left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T \underline{\Lambda}_k\right)\right\}$$
$$= exp\left\{-\frac{1}{2\psi_{\epsilon kj}} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T \underline{\Lambda}_k\right)^T \left(\underline{\tilde{y}}_{ijk} - \underline{\omega}_{-ijk}^T \underline{\Lambda}_k\right)\right\} \quad (47)$$

แทนสมการที่ (46) และ (47) ลงในสมการที่ (45) จากนั้นทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออก จะได้ $p\left(\underline{\Lambda}_k | \left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj}, \underline{g}_k, H_k\right)$

$$\propto exp\left\{-\tfrac{1}{2}\left[\underline{\Lambda}_k^T \left(H_k^{-1} + \sum_{j=1}^{J} \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\omega}_{-ijk}^T\right) \underline{\Lambda}_k\right] - 2\underline{\Lambda}_k^T \left(H_k^{-1} \underline{g}_k + \right.\right.$$
$$\left.\left.\sum_{j=1}^{J} \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\tilde{y}}_{ijk}\right)\right\} \quad (48)$$

เมื่อเปรียบเทียบสมการในข้างต้นกับรูปแบบการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรจะได้ว่า สำหรับตัวแปรสังเกตได้ที่ k ในหน่วยที่ j

$$\underline{\Lambda}_k | \left\{\underline{\tilde{y}}_{ijk}\right\}, \{\underline{\omega}_{-ijk}\}, \psi_{\epsilon kj} \sim Normal(\underline{\hat{g}}_k, \hat{H}_k) \tag{49}$$

โดยที่ $\hat{H}_k = \left[H_k^{-1} + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\omega}_{-ijk}^T\right]^{-1}$

และ $\underline{\hat{g}}_k = \hat{H}_k \left[H_k^{-1} \underline{g}_k + \psi_{\epsilon kj}^{-1} \sum_{i=1}^{n_j} \underline{\omega}_{-ijk} \underline{\tilde{y}}_{ijk}\right]$

8. *การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรแฝง $\underline{\omega}_{ij}$ เมื่อกำหนดพารามิเตอร์ $\underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}_j, \Lambda, \Psi_{\epsilon j}, \underline{y}_{ij}$*

จากทฤษฎีของเบส์จะได้การแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของตัวแปรแฝง $\underline{\omega}_{ij}$ เมื่อกำหนดพารามิเตอร์ $\underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}, \Lambda, \Psi_{\epsilon j}, \underline{y}_{ij}$ คือ

$$p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}_j, \Lambda, \Psi_{\epsilon j}, \underline{y}_{ij}\right) \propto p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}\right) \times p\left(\underline{y}_{ij} | \underline{\omega}_{ij}, \underline{\mu}_j, \Lambda, \Psi_{\epsilon j}\right) \tag{50}$$

การแจกแจงความน่าจะเป็นก่อนหน้าของตัวแปรแฝง $\underline{\omega}_{ij}$ มีการแจกแจงความน่าจะเป็นแบบปกติ คือ

$$p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\omega}_{ij} - \underline{\theta}_j\right)^T \Sigma_{\omega j}^{-1}\left(\underline{\omega}_{ij} - \underline{\theta}_j\right)\right\} \tag{51}$$

และฟังก์ชันภาวะความควรจะเป็นของ $\underline{y}_{ij}$ เมื่อกำหนด $\underline{\omega}_{ij}, \underline{\mu}, \Lambda_j, \Psi_{\epsilon j}$ คือ

$$p\left(\underline{y}_{ij} | \underline{\omega}_{ij}, \underline{\mu}_j, \Lambda, \Psi_{\epsilon j}\right) \propto exp\left\{-\frac{1}{2}\left(\underline{y}_{ij} - \underline{\mu}_j - \Lambda\ \underline{\omega}_{ij}\right)^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}_j - \Lambda\ \underline{\omega}_{ij}\right)\right\} \tag{52}$$

แทนค่าสมการที่ (51) และ (52) ลงในสมการที่ (50) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออกจากการพิจารณาจะได้ว่า $p\left(\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}_j, \Lambda, \Psi_{\epsilon j}, \underline{y}_{ij}\right)$

$$\propto exp\left\{-\frac{1}{2}\left[\underline{\omega}_{ij}^T\left(\Sigma_{\omega j}^{-1} + \Lambda^T \Psi_{\epsilon j}^{-1} \Lambda\right)\underline{\omega}_{ij} - 2\underline{\omega}_{ij}^T\left(\Sigma_{\omega j}^{-1}\underline{\theta}_j + \Lambda^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}_j\right)\right)\right]\right\} \tag{53}$$

เปรียบเทียบสมการที่ (53) กับรูปแบบการแจกแจงความน่าจะเป็นแบบปกติจะได้ว่า สำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j

$$\underline{\omega}_{ij} | \underline{\theta}_j, \Sigma_{\omega j}, \underline{\mu}_j, \Lambda, \Psi_{\epsilon j}, \underline{y}_{ij} \sim Normal\left(\underline{\hat{\theta}}_{ij}, \hat{\Sigma}_{\omega j}\right) \tag{54}$$

โดยที่ $\hat{\Sigma}_{\omega j} = \left[\Sigma_{\omega j}^{-1} + \Lambda^T \Psi_{\epsilon j}^{-1} \Lambda\right]^{-1}$

และ $\underline{\hat{\theta}}_{ij} = \hat{\Sigma}_{\omega j}\left[\Sigma_{\omega j}^{-1} \underline{\theta}_j + \Lambda^T \Psi_{\epsilon j}^{-1}\left(\underline{y}_{ij} - \underline{\mu}_j\right)\right]$

9. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ Φ_j^{-1} เมื่อกำหนดพารามิเตอร์ $\{\underline{\xi}_{ij}\}, \rho_{\phi j}, R_{\phi j}$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ Φ_j^{-1} เมื่อกำหนด $\{\underline{\xi}_{ij}\}, \rho_{\phi j}, R_{\phi j}$ คือ

$$p\left(\Phi_j^{-1} | \{\underline{\xi}_{ij}\}, \rho_{\phi j}, R_{\phi j}\right) \propto p\left(\Phi_j^{-1} | \rho_{\phi j}, R_{\phi j}\right) \times p\left(\{\underline{\xi}_{ij}\} |, \Phi_j\right) \qquad (55)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ Φ_j^{-1} คือ

$$p\left(\Phi_j^{-1} | \rho_{\phi j}, R_{\phi j}\right) \propto \left|\Phi_j^{-1}\right|^{\frac{\rho_{\phi j} - q_2 - 1}{2}} exp\left\{-\frac{1}{2} tr\left(R_{\phi j}^{-1} \Phi_j^{-1}\right)\right\} \qquad (56)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\{\underline{\xi}_{ij}\}$ เมื่อกำหนด Φ_j คือ

$$p\left(\{\underline{\xi}_{ij}\} | \Phi_j\right) \propto \prod_{i=1}^{n_j} \left|\Phi_j\right|^{-\frac{1}{2}} exp\left\{-\frac{1}{2} \underline{\xi}_{ij}^{\,T} \Phi_j^{-1} \underline{\xi}_{ij}\right\}$$

$$= \left|\Phi_j\right|^{-\frac{n_j}{2}} exp\left\{-\frac{1}{2} \sum_{i=1}^{n_j} \underline{\xi}_{ij}^{\,T} \Phi_j^{-1} \underline{\xi}_{ij}\right\} \qquad (57)$$

แทนสมการที่ (56) และ (57) ลงในสมการที่ (55) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ Φ_j^{-1} เมื่อกำหนด $\{\underline{\xi}_{ij}\}, \rho_{\phi j}, R_{\phi j}$ คือ

$$p\left(\Phi_j^{-1} | \{\underline{\xi}_{ij}\}, \rho_{\phi j}, R_{\phi j}\right) \propto \left|\Phi_j\right|^{-\frac{n_j + \rho_{\phi j} - q_2 - 1}{2}} exp\left\{-\frac{1}{2}\left[tr\left(R_{\phi j}^{-1} \Phi_j^{-1}\right) + \sum_{i=1}^{n_j} \underline{\xi}_{ij}^{\,T} \Phi_j^{-1} \underline{\xi}_{ij}\right]\right\}$$

$$= \left|\Phi_j^{-1}\right|^{-\frac{n_j + \rho_{\phi j} - q_2 - 1}{2}} exp\left\{-\frac{1}{2}\left[tr\left(\Phi_j^{-1} R_{\phi j}^{-1}\right) + tr\left(\Phi_j^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ij} \ \underline{\xi}_{ij}^{T}\right)\right]\right\}$$

$$= \left|\Phi_j^{-1}\right|^{-\frac{n_j + \rho_{\phi j} - q_2 - 1}{2}} exp\left\{-\frac{1}{2} tr\left(\Phi_j^{-1} R_{\phi j}^{-1} + \Phi_j^{-1} \sum_{i=1}^{n_j} \underline{\xi}_{ij} \ \underline{\xi}_{ij}^{T}\right)\right\}$$

$$= \left|\Phi_j^{-1}\right|^{-\frac{n_j + \rho_{\phi j} - q_2 - 1}{2}} exp\left\{-\frac{1}{2} tr\left(\Phi_j^{-1}\left[R_{\phi j}^{-1} + \sum_{i=1}^{n_j} \underline{\xi}_{ij} \ \underline{\xi}_{ij}^{T}\right]\right)\right\} \qquad (58)$$

เปรียบเทียบสมการที่ (58) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบ Wishart จะได้ว่าสำหรับหน่วยตัวอย่างที่ i ในหน่วยที่ j

$$\Phi_j^{-1} | \{\underline{\xi}_{ij}\}, \rho_{\phi j}, R_{\phi j} \sim Wishart\left(\hat{\rho}_{\phi j}, \hat{R}_{\phi j}\right) \qquad (59)$$

โดยที่ $\hat{\rho}_{\phi j} = n_j + \rho_{\phi j}$ และ $\hat{R}_{\phi} = \left[R_{\phi j}^{-1} + \sum_{i=1}^{n_j} \underline{\xi}_{ij} \ \underline{\xi}_{ij}^{T}\right]^{-1}$

10. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\beta_j}$ เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\beta_j}$ เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u$ จะได้ว่า

$$p\left(\underline{\beta_j}|\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u\right) \propto p\left(\underline{\beta_j}|\underline{\gamma}, \Sigma_u\right) \times p\left(\{\underline{\omega}_{ij}\}|\underline{\beta_j}, \sigma_\delta^2\right) \quad (60)$$

โดยที่การแจกแจงความน่าจะเป็นของพารามิเตอร์สุ่ม $\underline{\beta_j}$ คือ

$$p\left(\underline{\beta_j}|\underline{\gamma}, \Sigma_u\right) \propto exp\left\{-\frac{1}{2}\left(\underline{\beta_j} - \underline{\gamma}\right)^T \Sigma_u^{-1}\left(\underline{\beta_j} - \underline{\gamma}\right)\right\} \quad (61)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\{\underline{\omega}_{ij}\}$ เมื่อกำหนด $\underline{\beta_j}, \sigma_\delta^2$ คือ

$$p\left(\{\underline{\omega}_{ij}\}|\underline{\beta_j}, \sigma_\delta^2\right) \propto \prod_{i=1}^{n_j} exp\left\{-\frac{1}{2\sigma_\delta^2}\left(\eta_{ij} - \underline{\xi}_{ij}^T\underline{\beta_j}\right)^T\left(\eta_{ij} - \underline{\xi}_{ij}^T\underline{\beta_j}\right)\right\}$$

$$= exp\left\{-\frac{1}{2\sigma_\delta^2}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\xi}_{ij}^T\underline{\beta_j}\right)^T\left(\eta_{ij} - \underline{\xi}_{ij}^T\underline{\beta_j}\right)\right\} \quad (62)$$

แทนสมการที่ (61) และ (62) ลงในสมการที่ (60) ทำการกระจายพจน์กำลังสองสมบูรณ์และตัดพจน์ที่เป็นค่าคงที่ออกจะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ $\underline{\beta_j}$ เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u$ คือ

$$p\left(\underline{\beta_j}|\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u\right)$$

$$\propto exp\left\{-\frac{1}{2}\left[\underline{\beta_j}^T\left(\Sigma_u^{-1} + \frac{1}{\sigma_\delta^2}\sum_{i=1}^{n_j}\underline{\xi}_{ij}\,\underline{\xi}_{ij}^T\right)\underline{\beta_j} - 2\underline{\beta_j}^T\left(\sum_{i=1}^{n_j}\underline{\xi}_{ij}\,\eta_{ij}\right)\right]\right\} \quad (63)$$

เปรียบเทียบสมการ (4.113) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบปกติหลายตัวแปรจะได้ว่าสำหรับหน่วยที่ j

$$\underline{\beta_j}|\{\underline{\omega}_{ij}\}, \underline{\gamma}, \sigma_\delta^2, \Sigma_u \sim Normal\left(\hat{\underline{\beta_j}}, \hat{\Sigma}_{\beta j}\right) \quad (64)$$

โดยที่ $\hat{\Sigma}_{\beta j} = \left[\Sigma_u^{-1} + \frac{1}{\sigma_\delta^2}\sum_{i=1}^{n_j}\underline{\xi}_{ij}\underline{\xi}_{ij}^T\right]^{-1}$

และ $\hat{\underline{\beta_j}} = \hat{\Sigma}_{\beta j}\left[\Sigma_u^{-1}\underline{\gamma} + \frac{1}{\sigma_\delta^2}\sum_{i=1}^{n_j}\underline{\xi}_{ij}\eta_{ij}\right]$

จากตัวอย่างสุ่มของพารามิเตอร์ $\underline{\beta_j}$ จะได้ว่าสามารถประมาณค่าพารามิเตอร์อิทธิพลคงที่ได้โดย

$$\underline{\gamma} = \sum_{j=1}^{J}\underline{\beta_j} \quad (65)$$

และประมาณค่าเศษเหลือในระดับที่สอง $\underline{u_j}$ ได้โดย

$$\underline{u}_j = \underline{\beta_j} - \underline{\gamma} \quad (66)$$

11. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ Σ_u^{-1} เมื่อกำหนด $\{\underline{u}_j\}, \rho_u, R_u$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ Σ_u^{-1} เมื่อกำหนด $\underline{u}_j, \rho_u, R_u$ คือ

$$p\left(\Sigma_u \middle| \{\underline{u}_j\}, \rho_u, R_u\right) \propto p\left(\Sigma_u^{-1} \middle| \rho_u, R_u\right) \times p\left(\{\underline{u}_j\} \middle| \Sigma_u\right) \quad (67)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ Σ_u^{-1} มีการแจกแจงความน่าจะเป็นแบบ Wishart ที่มีพารามิเตอร์ ρ_u, R_u กล่าวคือ

$$p\left(\Sigma_u^{-1} \middle| \rho_u, R_u\right) \propto |\Sigma_u^{-1}|^{\frac{\rho_u - q_2 - 1}{2}} exp\left\{-\frac{1}{2} tr\left(R_u^{-1} \Sigma_u^{-1}\right)\right\} \quad (68)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\underline{u}_j$ เมื่อกำหนด Σ_u คือ

$$p\left(\{\underline{u}_j\} \middle| \Sigma_u\right) \propto \prod_{j=1}^{J} p\left(\underline{u}_j \middle| \Sigma_u\right) \quad (69)$$

$$= \prod_{j=1}^{J} |\Sigma_u|^{-\frac{1}{2}} exp\left\{-\frac{1}{2} \underline{u}_j^T \Sigma_u^{-1} \underline{u}_j\right\}$$

$$= |\Sigma_u|^{-\frac{J}{2}} exp\left\{-\frac{1}{2} \sum_{j=1}^{J} \underline{u}_j^T \Sigma_u^{-1} \underline{u}_j\right\} \quad (70)$$

แทนสมการที่ (69) และ (70) ลงในสมการที่ (68) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ Σ_u เมื่อกำหนด $\underline{u}_j, \rho_u, R_u$ คือ

$$p\left(\Sigma_u \middle| \{\underline{u}_j\}, \rho_u, R_u\right) \propto |\Sigma_u^{-1}|^{\frac{J + \rho_u - q_2 - 1}{2}} exp\left\{-\frac{1}{2} tr\left(\Sigma_u^{-1}\left[R_u^{-1} + \sum_{j=1}^{J} \underline{u}_j^T \underline{u}_j\right]\right)\right\} \quad (71)$$

เปรียบเทียบสมการที่ (71) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบ Wishart จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ Σ_u^{-1} เมื่อกำหนด$\{\underline{u}_j\}, \hat{\rho}_u, \hat{R}_u$ คือ

$$\Sigma_u^{-1} \middle| \{\underline{u}_j\}, \rho_u, R_u \sim Wishart\left(\hat{\rho}_u, \hat{R}_u\right) \quad (72)$$

โดยที่ $\hat{\rho}_u = J + \rho_u$ และ $\hat{R}_u = \left[R_u^{-1} + \sum_{j=1}^{J} \underline{u}_j^T \underline{u}_j\right]^{-1}$

12. การแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ σ_δ^{-2} เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \{\underline{\beta}_j\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta$

จากทฤษฎีของเบส์จะได้ว่าการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของพารามิเตอร์ σ_δ^{-2} เมื่อกำหนด $\{\underline{\omega}_{ij}\}, \{\underline{\beta}_j\}, \{\underline{u}_j\}, \alpha_\delta, \beta_\delta$ คือ

$$p\left(\sigma_\delta^{-2}|\{\underline{\omega}_{ij}\},\left\{\underline{\beta}_j\right\},\{\underline{u}_j\},\alpha_\delta,\beta_\delta\right) \propto p(\sigma_\delta^{-2}|\alpha_\delta,\beta_\delta) \times p\left(\{\underline{\omega}_{ij}\}|\sigma_\delta^{-2},\left\{\underline{\beta}_j\right\},\{\underline{u}_j\}\right) \quad (73)$$

โดยที่การแจกแจงความน่าจะเป็นก่อนหน้าของพารามิเตอร์ σ_δ^{-2} เมื่อกำหนดพารามิเตอร์ $\alpha_\delta, \beta_\delta$ คือ

$$p(\sigma_\delta^{-2}|\alpha_\delta,\beta_\delta) \propto (\sigma_\delta^{-2})^{\alpha_\delta - 1} exp\{-\beta_\delta \sigma_\delta^{-2}\} \quad (74)$$

และฟังก์ชันภาวะความควรจะเป็นของ $\{\underline{\omega}_{ij}\}$ เมื่อกำหนด $\sigma_\delta^{-2},\left\{\underline{\beta}_j\right\},\{\underline{u}_j\}$ คือ

$$p\left(\{\underline{\omega}_{ij}\}|\sigma_\delta^{-2},\left\{\underline{\beta}_j\right\},\{\underline{u}_j\}\right) \propto \prod_{j=1}^{J}\prod_{i=1}^{n_j} p\left(\underline{\omega}_{ij}|\sigma_\delta^{-2},\underline{\beta}_j,\underline{u}_j\right)$$

$$= \prod_{j=1}^{J}\prod_{i=1}^{n_j}(\sigma_\delta^{-2})^{-\frac{1}{2}} exp\left\{-\frac{1}{2\sigma_\delta^2}\left(\eta_{ij} - \underline{\beta}_j^T\underline{\xi}_{ij}\right)^T\left(\eta_{ij} - \underline{\beta}_j^T\underline{\xi}_{ij}\right)\right\}$$

$$= (\sigma_\delta^2)^{-\frac{\sum_{j=1}^{J} n_j}{2}} exp\left\{-\frac{1}{2\sigma_\delta^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\beta}_j^T\underline{\xi}_{ij}\right)^T\left(\eta_{ij} - \underline{\beta}_j^T\underline{\xi}_{ij}\right)\right\} \quad (75)$$

แทนสมการที่ (74) และ (75) ลงในสมการ (73) จะได้ว่าการแจกแจงความน่าจะเป็นภายหลังแบบมีเงื่อนไขของพารามิเตอร์ σ_δ^{-2} เมื่อกำหนด $\{\underline{\omega}_{ij}\},\left\{\underline{\beta}_j\right\},\{\underline{u}_j\},\alpha_\delta,\beta_\delta$ คือ

$$p\left(\sigma_\delta^{-2}|\{\underline{\omega}_{ij}\},\left\{\underline{\beta}_j\right\},\{\underline{u}_j\},\alpha_\delta,\beta_\delta\right)$$

$$\propto (\sigma_\delta^{-2})^{\frac{\sum_{j=1}^{J} n_j}{2}+\alpha_\delta - 1} exp\left\{-\sigma_\delta^{-2}\left[\beta_\delta + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\beta}_j^T\underline{\xi}_{ij}\right)^T\left(\eta_{ij} - \underline{\beta}_j^T\underline{\xi}_{ij}\right)\right]\right\} \quad (76)$$

เปรียบเทียบสมการที่ (76) กับรูปแบบของการแจกแจงความน่าจะเป็นแบบแกมมา จะได้ว่า

$$\sigma_\delta^{-2}|\{\underline{\omega}_{ij}\},\left\{\underline{\beta}_j\right\},\{\underline{u}_j\},\alpha_\delta,\beta_\delta \sim Gamma(\hat{\alpha}_\delta,\hat{\beta}_\delta) \quad (77)$$

โดยที่ $\hat{\alpha}_\delta = \frac{\sum_{j=1}^{J} n_j}{2} + \alpha_\delta$ และ $\hat{\beta}_\delta = \beta_\delta + \frac{1}{2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\eta_{ij} - \underline{\beta}_j^T\underline{\xi}_{ij}\right)^2$

ประวัติผู้เขียนวิทยานิพนธ์

นายสิวะโชติ ศรีสุทธิยากร เกิดวันเสาร์ที่ 5 มีนาคม พ.ศ. 2526 จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาปริญญาวิทยาศาสตร์บัณฑิต จากจุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2547 และสำเร็จการศึกษาปริญญาสถิติศาสตร์มหาบัณฑิต จาก จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2550 และได้เข้าศึกษาในหลักสูตรปริญญาครุศาสตร์ดุษฎีบัณฑิต คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยในปีการศึกษา 2551

ผลการทางวิชาการระหว่างการศึกษาที่ตีพิมพ์เผยแพร่ และนำเสนอในที่ประชุมวิชาการ

1. Srisuttiyakorn, S., & Kanjanawasee, S. (2009). Multilevel Parameter Estimates when Level-One Independent Variables Measure with Error. *IMES.* Bangkok.
2. สิวะโชติ ศรีสุทธิยากร และ ศิริชัย กาญจนวาสี. (2551). *การกำหนดขนาดตัวอย่างสำหรับการวิจัยเชิงทดลอง ด้วยอำนาจการทดสอบสมมติฐานทางสถิติโดยใช้โปรแกรม SAS และ SPSS*. วารสารวิธีวิทยาการวิจัย ปีที่ 21 ฉบับที่ 2 (พฤษภาคม-สิงหาคม 2551).
3. สิวะโชติ ศรีสุทธิยากร และ ศิริชัย กาญจนวาสี. (2555). (รอตีพิมพ์). *วิธีการประมาณค่าพารามิเตอร์แบบเบส์สำหรับโมเดลสมการโครงสร้างพหุระดับกรณีพารามิเตอร์น้ำหนักองค์ประกอบและความแปรปรวนของความคลาดเคลื่อนจากการวัดเป็นพารามิเตอร์สุ่ม.* วารสารวิธีวิทยาการวิจัย.