บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันสถิติวิเคราะห์ชั้นสูงได้มีการพัฒนาไปอย่างรวดเร็ว โดยเฉพาะอย่างยิ่งเทคนิคการวิเคราะห์องค์ประกอบ (Factor Analysis) ซึ่งเป็นเทคนิคที่นักสถิติทั้งหลายถือว่าเป็นยอดของสถิติวิเคราะห์ (Kerlinger & Lee, 2000, p. 825) และเป็นสถิติวิเคราะห์ชั้นสูงที่มีผู้พัฒนาโปรแกรมสำเร็จรูปหลายโปรแกรม ทำให้การวิเคราะห์องค์ประกอบกระทำได้สะดวกและรวดเร็วขึ้น จึงมีผู้นำเทคนิคนี้ไปใช้ในการทำวิจัยต่าง ๆ มากขึ้น

การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) จัดเป็นส่วนหนึ่งของโมเดลสมการเชิงโครงสร้าง (Structural Equation Modeling: SEM) (Flora & Curran, 2004, p. 466) ที่ใช้ในการตรวจสอบเพื่อยืนยันความแม่นยำของโมเดลการวัด (Measurement Model) ที่ประกอบด้วยชุดตัวชี้วัด (Indicators) หรือชุดข้อคำถาม (Items) หรือตัวแปรสังเกตได้ (Observed Variables) ของตัวแปรแฝง (Latent Variables) หรือองค์ประกอบ (Factor) เป็นการตรวจสอบความสอดคล้องระหว่างโมเดลสมมติฐานตามทฤษฎีกับข้อมูลเชิงประจักษ์ (Brown, 2006, p. 1) เป็นเทคนิคที่ได้รับความนิยมอย่างแพร่หลายในการวิจัยทางสังคมศาสตร์ พฤติกรรมศาสตร์ การศึกษา จิตวิทยา ธุรกิจ และสาขาต่าง ๆ (Curran, West, & Finch, 1996, p. 16; Lei & Lomax, 2005, p. 1; Finney & DiStefano, 2006, p. 269; Brown, 2006, p. 12)

การนำเทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันมาประยุกต์ใช้ในงานวิจัยนั้น สิ่งที่นักวิจัยต้องให้ความสำคัญเป็นอันดับแรก คือ การตรวจสอบข้อตกลงเบื้องต้น (Assumption) ของข้อมูลที่นำมาใช้วิเคราะห์ เนื่องจากการละเมิดข้อตกลงเบื้องต้นนั้นอาจทำให้ค่าประมาณพารามิเตอร์ของโมเดลเกิดความเอนเอียง (Bias) หรือส่งผลต่อดัชนีวัดความกลมกลืนต่าง ๆ รวมถึงการทดสอบสมมติฐานมีความไม่น่าเชื่อถือ และนำไปสู่การอธิบายผลที่ไม่ถูกต้องได้ (Finney & DiStefano, 2006, p. 269; Lei & Lomax, 2005, p. 2; Vermunt & Magidson, 2004, p. 1) โดยข้อตกลงเบื้องต้นที่สำคัญในการอนุมานทางสถิติที่เกี่ยวข้องกับข้อมูลเชิงพหุทั่วไป คือ ทฤษฎีการแจกแจงแบบโค้งปกติ (Normal Theory: NT) คือ ข้อมูลต้องมาจากประชากรที่มีการแจกแจงแบบปกติพหุตัวแปร (Multivariate Normality Distribution) ข้อมูลมีลักษณะเป็นแบบต่อเนื่อง และกลุ่มตัวอย่างมีขนาดใหญ่เพียงพอ แต่บ่อยครั้งที่พบว่าข้อมูลในงานวิจัยต่าง ๆ ไม่ได้เป็นไปตามข้อตกลงเบื้องต้นดังกล่าว ดังนั้น การประมาณค่าพารามิเตอร์ที่จะให้ค่าประมาณของพารามิเตอร์ (Parameter Estimates) มีค่าใกล้เคียงกับค่าจริงของพารามิเตอร์ (True Parameters) มากที่สุดนั้น

ผู้วิจัยควรเลือกใช้วิธีการประมาณค่าพารามิเตอร์ และขนาดกลุ่มตัวอย่างที่เหมาะสมสอดคล้องกับธรรมชาติและลักษณะของข้อมูลและเพื่อให้ผลลัพธ์จากการประมาณค่ามีความน่าเชื่อถือมากที่สุด

โปรแกรมสำเร็จรูปที่ใช้วิเคราะห์โมเดลสมการเชิงโครงสร้าง เช่น LISREL (Jöreskog & Sörbom, 1993 cited in Curran, West, & Finch, 1996, p. 17) EQS (Bentler, 1989 cited in Curran, West & Finch, 1996, p. 17) PROC CALIS (SAS Institute, Inc cited in Curran, West & Finch, 1996, p. 17) RAMONA (Brown, Mels & Coward, 1994 cited in Curran, West & Finch, 1996, p. 17) หรือ Mplus (Muthén & Muthén, 2007) ได้มีการพัฒนาเทคนิคการประมาณค่าพารามิเตอร์การวิเคราะห์ข้อมูลที่หลากหลายวิธี ซึ่งแต่ละวิธีการมีข้อตกลงเบื้องต้นแตกต่างกัน มีความอ่อนไหว (Sensitive) และความแกร่ง (Robustness) ต่อการละเมิดข้อตกลงเบื้องต้นในเงื่อนไขที่แตกต่างกัน

วิธีการประมาณค่าพารามิเตอร์สำหรับโมเดลสมการเชิงโครงสร้าง ที่นักวิจัยทั่วไปนิยมใช้มากที่สุด ได้แก่ วิธีการ Maximum Likelihood (ML) ซึ่งโดยทั่วไปจะถูกตั้งไว้เป็นค่ามาตรฐานเบื้องต้น (Standard Default) ของโปรแกรมวิเคราะห์หลายโปรแกรม (Flora & Curran, 2004, p. 466; Kline, 2005, p. 178; Curran, West, & Finch, 1996, p. 17; Finney & DiStefano, 2006, p. 271) ซึ่งมีข้อตกลงเบื้องต้น คือ 1) หน่วยตัวอย่างที่สังเกตต้องเป็นอิสระต่อกัน 2) กลุ่มตัวอย่างมีขนาดใหญ่เพียงพอ 3) มีการกำหนดคุณลักษณะจำเพาะของโมเดลอย่างถูกต้องตรงกับโครงสร้างจริงของประชากร 4) ข้อมูลจากตัวแปรสังเกตได้มีการแจกแจงแบบปกติพหุ และ 5) ตัวแปรมีลักษณะเป็นข้อมูลแบบต่อเนื่อง (Finney & DiStefano, 2006, p. 271) ถ้าข้อมูลมีลักษณะเป็นไปตามข้อตกลงเบื้องต้นดังกล่าวจะมีผลทำให้การประมาณค่าพารามิเตอร์มีความเหมาะสมและมีความน่าเชื่อถือ คือ ปราศจากความเอนเอียง (Unbiasedness) มีประสิทธิภาพ (Efficiency) และมีความคงเส้นคงวา (Consistency) (Muthén & Kaplan, 1985, p. 172; West, Finch, & Curran, 1995, pp. 56 – 58; Finch, West, & Mackinnon, 1997, pp. 87 – 88; Finney & DiStefano, 2006, p. 271)

ในกรณีที่ข้อมูลที่ได้จากการเก็บรวบรวมไม่เป็นไปตามข้อตกลงเบื้องต้นดังกล่าวข้างต้น ซึ่ง Gierl & Mulvenon (1995 cited in Finney & DiStefano, 2006, p. 270) กล่าวว่า นักวิจัยส่วนใหญ่ไม่ได้ทำการทดสอบข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงของข้อมูลแต่มักจะสันนิษฐาน (Assume) ว่าข้อมูลมีลักษณะการแจกแจงแบบโค้งปกติ ซึ่งบางครั้งหรือบางตัวแปรอาจจะมีลักษณะการแจกแจงไม่เป็นโค้งปกติ (Non – normality) หรือข้อมูลที่ได้จากการเก็บรวบรวมมีลักษณะเป็นข้อมูลแบบจำแนกกลุ่ม (Categorical Data) โดยเฉพาะที่ใช้มาตรวัดแบบเรียงอันดับ (Ordinal Scale) เช่น Likert Type Scale ซึ่งมีธรรมชาติลักษณะเป็นข้อมูลที่ไม่ต่อเนื่อง (Discrete Data) ไม่สามารถนิยามการแจกแจงเป็นโค้งปกติได้ (Flora & Curran, 2004, p. 466; Forero, Olivares, &

Pujol, 2009, p. 625; Muthën & Kaplan, 1992, p. 19; Finney & DiStefano, 2006, p. 271; Muthën, 1984 cited in Myers, Ahn, & Jin, 2011, p. 414) ผลลัพธ์จากการประมาณค่าพารามิเตอร์ ภายใต้เงื่อนไขที่ละเมิดข้อตกลงเบื้องต้นด้วยวิธีความเป็นไปได้สูงสุด (Maximum Likelihood: ML) นั้น จะเกิดปัญหาด้านไม่เที่ยงตรงและความเอนเอียงของผลการวิเคราะห์ค่าประมาณของพารามิเตอร์ ค่าน้ำหนักองค์ประกอบ (Factor Loading) ค่าคลาดเคลื่อนมาตรฐาน (Standard Error) สถิติทดสอบ Chi – square ค่าความเชื่อมั่นของตัวแปรแฝง (Construct Reliability: ρ_c) ค่าความแปรปรวนที่สกัดได้ (Variance Extracted: ρ_v) และดัชนีวัดความสอดคล้องกลมกลืนต่าง ๆ (Schumacker & Lomax, 2010, p. 61; Trierweiler, 2009, p. 1; Kline, 2005, p. 178; Flora & Curran, 2004, p. 466; Babakus, Ferguson, & Jöreskog, 1987, pp. 73 – 74; Green et al., 1997, p. 108; Hutchinson & Olmos, 1998, pp. 344 – 364)

จากข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงไม่เป็นโค้งปกติ และมีลักษณะจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) ที่เป็นข้อจำกัดของวิธีการ Maximum Likelihood ดังที่กล่าวมาข้างต้น วิธีหนึ่งที่นักวิจัยส่วนใหญ่จัดการกับข้อมูลที่มีลักษณะดังกล่าว คือ การเปลี่ยนรูปข้อมูล (Data Transforming) ด้วยวิธีการต่าง ๆ เพื่อให้แต่ละตัวแปรมีการแจกแจงเป็นโค้งปกติ (Univariate Normal Distribution) และข้อมูลมีลักษณะต่อเนื่อง (Continuous Data) แล้วจึงดำเนินการประมาณค่าพารามิเตอร์ด้วยวิธีการ Maximum Likelihood (Schumacker & Lomax, 2010, p. 28) แต่ยังไม่มีความชัดเจนเกี่ยวกับการแปลผลการวิเคราะห์หลังจากการเปลี่ยนรูปข้อมูล ซึ่งทำให้ธรรมชาติของข้อมูลเปลี่ยนไป ดังนั้น จึงได้มีการศึกษาและพัฒนาวิธีการประมาณค่าพารามิเตอร์ที่สอดคล้องกับลักษณะและธรรมชาติของข้อมูลมากขึ้น เช่น วิธี Robust Weight Least Square (Mean and Variance Adjusted Chi – square) (WLSMV) เป็นวิธีการที่พัฒนาโดย Muthén et al. (1997 cited in Trierweiler, 2009, p. 3) และประยุกต์การคำนวณในโปรแกรม Mplus ในปี ค.ศ. 2007 (Muthén & Muthén, 2007) ซึ่งเป็นวิธีการที่มีความแกร่งต่อการละเมิดข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลที่มีลักษณะจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) และมีการแจกแจงไม่เป็นโค้งปกติ (Non – normality) ได้อย่างมีประสิทธิภาพ (Flora & Curran, 2004, pp. 470 – 471) โดยที่ไม่ต้องดำเนินการเปลี่ยนรูปข้อมูล และวิธีการประมาณค่าอีกวิธีหนึ่ง คือ วิธีการ Bayesian เป็นวิธีการประมาณค่าพารามิเตอร์ที่เริ่มได้รับความนิยมมากขึ้นในการประมาณค่าโมเดลสมการเชิงโครงสร้าง (Muthén & Asparouhov, 2011, p. 5) ซึ่ง Muthén & Muthén (1998 – 2010 cited in Muthén & Asparouhov, 2011, p. 5) ได้พัฒนาวิธีการ Bayesian ประยุกต์การคำนวณในโปรแกรม Mplus ด้วยเช่นกัน ซึ่งมีการวิเคราะห์ที่ง่าย สะดวก และมีจุดเด่นดังนี้ คือ 1) สามารถศึกษาค่าประมาณของพารามิเตอร์ได้หลากหลาย ทุกรูปแบบการแจกแจงของตัวแปร ค่าประมาณพารามิเตอร์และสถิติทดสอบมีความแกร่งต่อการแจกแจงที่ไม่เป็นโค้งปกติ 2) มีความแกร่ง

ต่อเงื่อนไขที่กลุ่มตัวอย่างขนาดเล็ก 3) ระยะเวลาในการคำนวณด้วยคอมพิวเตอร์มีความรวดเร็วกว่าวิธีการอื่น ๆ และ 4) สามารถวิเคราะห์ข้อมูลสำหรับโมเดลแบบใหม่ ๆ ได้ดี เช่น โมเดลที่มีจำนวนพารามิเตอร์มาก ๆ (Muthén & Asparouhov, 2011, p. 5)

จากประเด็นที่ศึกษาค้นคว้ามาข้างต้นนำไปสู่ข้อสงสัยว่าภายใต้บริบทของการทำวิจัยในปัจจุบันที่นักวิจัยส่วนหนึ่งมักทำการวัดประเมินตัวแปรโดยใช้เครื่องมือวัดที่สร้างโดยบนฐานคิดของมาตรวัดแบบเรียงอันดับ (Ordinal Scale) ตามแนวทางของลิเคอร์ท (Likert Type Scale) ซึ่งมีธรรมชาติเป็นข้อมูลที่ไม่ต่อเนื่อง (Discrete Data) มีลักษณะเป็นข้อมูลจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) และมีการแจกแจงไม่เป็นโค้งปกติ (Non – normality Distribution) แล้วใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันไปใช้เพื่อตรวจสอบคุณภาพของโมเดลการวัด ภายใต้การสมมติว่าข้อมูลที่วัดได้นั้นมีลักษณะเป็นข้อมูลต่อเนื่อง (Continuous Data) แล้วใช้วิธีการประมาณค่าพารามิเตอร์ด้วยวิธี Maximum Likelihood หรือทำการเปลี่ยนรูปข้อมูลให้มีการแจกแจงปกติโดยชุดคำสั่งของโปรแกรมสำเร็จรูปก่อนทำการวิเคราะห์นั้น จะให้ผลการวิเคราะห์ข้อมูลที่มีความเที่ยงตรงเพียงใดและมีค่าการเอนเอียงไปจากค่าพารามิเตอร์ที่แท้จริงเพียงใด เมื่อเปรียบเทียบกับการใช้วิธี Robust Weight Least Square (Mean and Variance Adjusted Chi – square) หรือ WLSMV ที่ได้รับการพัฒนาขึ้นมาอย่างสอดคล้องกับลักษณะและธรรมชาติของข้อมูลจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) และมีการแจกแจงไม่เป็นโค้งปกติ (Non – normality Distribution) และผลการวิเคราะห์จะแตกต่างไปจากการใช้วิธี Bayesian ที่กำลังเริ่มได้รับความนิยมมากขึ้นในการประมาณค่าพารามิเตอร์ของโมเดลหรือไม่ โดยเงื่อนไขในการตรวจสอบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ดังกล่าวข้างต้นจะดำเนินการภายใต้เงื่อนไขของขนาดกลุ่มตัวอย่างที่มีขนาดแตกต่างกัน 4 ขนาด ที่คำนวณบนฐานคิดของการกำหนดขนาดของกลุ่มตัวอย่างสำหรับการวิจัยที่ต่างกัน 4 แนวทาง ซึ่งจะทำให้เกิดกรณีเงื่อนไขของการทดสอบรวม 16 เงื่อนไข และการศึกษาในครั้งนี้จะดำเนินการบนกรณีของการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับแรก (The First Order Confirmatory Factor Analysis) เท่านั้น

คำถามของการวิจัย

ภายใต้เงื่อนไขด้านลักษณะข้อมูลประชากรที่เป็นข้อมูลไม่ต่อเนื่อง (Discrete Data) มีลักษณะจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) ที่มีการแจกแจงไม่เป็นโค้งปกติ (Non – normality Distribution) และขนาดกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ที่แตกต่างกันนั้น วิธีการประมาณค่าพารามิเตอร์ทั้ง 4 แนวทาง จะให้ผลการวิเคราะห์ที่มีความเที่ยงตรงและความเอนเอียงแตกต่างกันหรือไม่ และวิธีการใดที่มีประสิทธิภาพมากที่สุด ภายใต้เงื่อนไขใดบ้าง

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์แบบ ML ในกรณีที่ไม่มีการเปลี่ยนรูปข้อมูลและกรณีที่มีการเปลี่ยนรูปข้อมูลให้มีการแจกแจงเป็นโค้งปกติ กับวิธีการประมาณค่าพารามิเตอร์แบบ WLSMV และวิธีการ Bayesian ในกรณีที่ข้อมูลการวัดอยู่ในมาตรวัดจำแนกกลุ่มแบบเรียงอันดับที่มีการแจกแจงไม่เป็นโค้งปกติ ภายใต้เงื่อนไขของขนาดกลุ่มตัวอย่างที่แตกต่างกัน 4 ขนาด ได้แก่ ขนาด 80, 160, 370 และ 623 หน่วยตัวอย่าง

ขอบเขตการวิจัย

1. ขอบเขตด้านประชากรและกลุ่มตัวอย่าง

1.1 ประชากร

ข้อมูลประชากรที่ใช้ในการศึกษาในครั้งนี้ เป็นข้อมูลประชากรเทียมที่สร้างขึ้นโดยอ้างอิงและประยุกต์จากงานวิจัยเชิงสำรวจ ที่มีลักษณะข้อมูลจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) ที่ใช้มาตรวัด Likert Type Scale 5 ระดับ จำนวน 10,000 ชุดข้อมูล และมีลักษณะการแจกแจงในแต่ละตัวแปร (Univariate Distribution) ไม่เป็นโค้งปกติ คือ เบ้ทางซ้าย (Negative Skewness) ทุกตัวแปร และโมเดลที่ใช้เป็นต้นแบบในการศึกษา ผู้วิจัยดำเนินการสร้างโมเดลการวัด หรือโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับแรก (First Order Confirmatory Factor Analysis) ที่ประกอบด้วย 2 องค์ประกอบ องค์ประกอบละ 4 ตัวแปรสังเกตได้ รวมทั้งสิ้น 8 ตัวแปรสังเกตได้

1.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่างได้มาจากสุ่มจากข้อมูลประชากร ด้วยขนาดกลุ่มตัวอย่างที่แตกต่างกัน 4 ขนาด โดยทำการศึกษาในแต่ละเงื่อนไข จำนวน 200 ชุดการทดลองซ้ำ (Replications)

2. ขอบเขตด้านตัวแปร

2.1 ตัวแปรอิสระ (Independent Variable) มี 2 ตัวแปร คือ วิธีการประมาณค่าพารามิเตอร์ และขนาดของกลุ่มตัวอย่าง

2.1.1 วิธีการประมาณค่าพารามิเตอร์ 4 วิธี คือ

2.1.1.1 วิธีการประมาณค่าพารามิเตอร์แบบ ML (Maximum Likelihood)

2.1.1.2 วิธีการประมาณค่าพารามิเตอร์แบบ ML (Maximum Likelihood) หลังการเปลี่ยนรูปข้อมูลให้มีการแจกแจงเป็นโค้งปกติ

2.1.1.3 วิธีการประมาณค่าพารามิเตอร์แบบ WLSMV (Robust Weighted Least Square – Mean and Variance Adjusted χ^2)

2.1.1.4 วิธีการประมาณค่าพารามิเตอร์แบบ Bayesian

2.1.2 ขนาดกลุ่มตัวอย่าง แบ่งเป็น 4 กลุ่ม คือ

2.1.2.1 ขนาด 80

2.1.2.2 ขนาด 160

2.1.2.3 ขนาด 370

2.1.2.4 ขนาด 623

2.2 ตัวแปรตาม (Dependent Variable) ได้แก่ ความเที่ยงตรงของการประมาณค่าพารามิเตอร์ ประกอบด้วยผลการวิเคราะห์

2.2.1 ค่าความเอนเอียงของค่าประมาณพารามิเตอร์ (Parameter Estimates Bias)

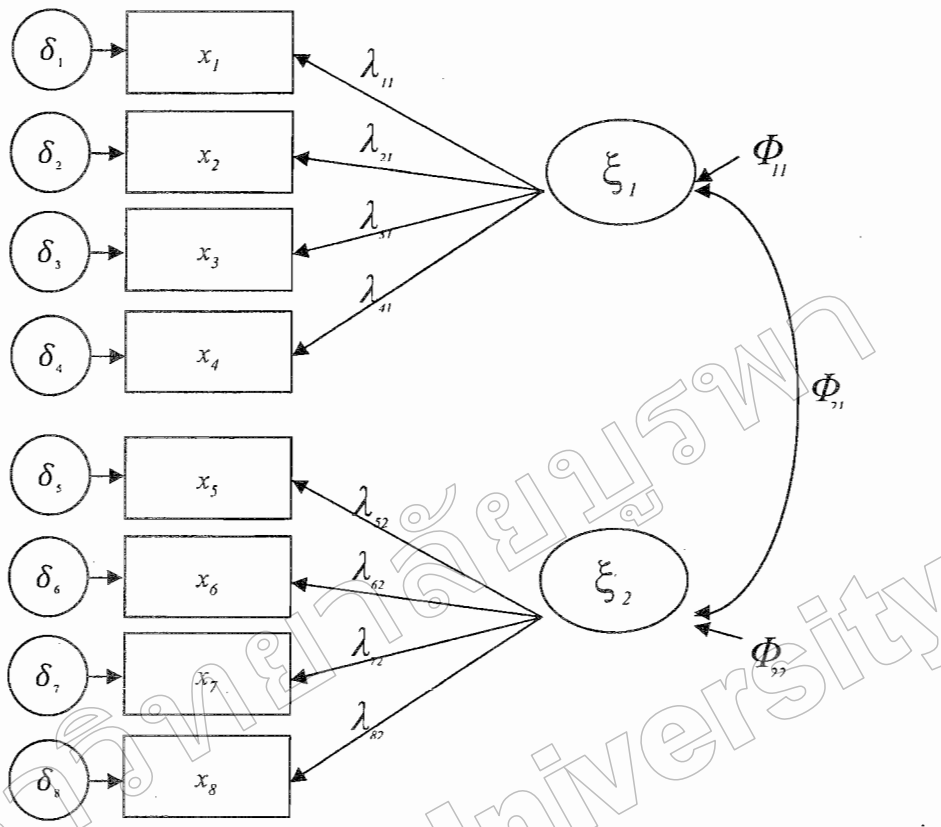
2.2.2 ค่าความเอนเอียงของค่าคลาดเคลื่อนมาตรฐาน (Standard Error Bias)

2.2.3 ค่าความเชื่อมั่นของตัวแปรแฝง (Construct Reliability: ρ_c) และค่าเฉลี่ยของความแปรปรวนที่ถูกสกัดได้ (Average Variance Extracted: ρ_v)

2.2.4 ร้อยละของโมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์ (Percent of Model Fit) โดยพิจารณาจากดัชนีวัดความสอดคล้องกลมกลืนของโมเดล (Model Fit Indices)

กรอบแนวคิดในการวิจัย

การใช้เทคนิค CFA นั้นมีความสำคัญมากในการวิจัยทุกสาขาวิชา โดยเฉพาะในกระบวนการพัฒนามาตรวัดเพื่อตรวจสอบโครงสร้างของตัวแปร ขั้นตอนการทดสอบและพัฒนาเครื่องมือการวิจัย เช่น แบบสอบถาม เป็นต้น (Brown, 2006, p.1) ดังนั้น ผู้วิจัยจึงเลือกศึกษาโมเดล CFA โดยใช้ข้อมูลประชากรเทียมที่สร้างขึ้นจากการประยุกต์ใช้ผลการเก็บรวบรวมจากการสำรวจข้อมูลจริง ซึ่งเพื่อให้ได้ข้อมูลที่สอดคล้องกับความเป็นจริงมากที่สุด และการจัดกลุ่มองค์ประกอบของตัวแปรสังเกตได้ผู้วิจัยใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) ซึ่งประกอบด้วย 2 องค์ประกอบ องค์ประกอบละ 4 ตัวแปรสังเกตได้ ซึ่งแสดงโมเดลได้ดังภาพที่ 1 – 1



ภาพที่ 1 – 1 โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) ที่ใช้เป็นโมเดลการวิจัย

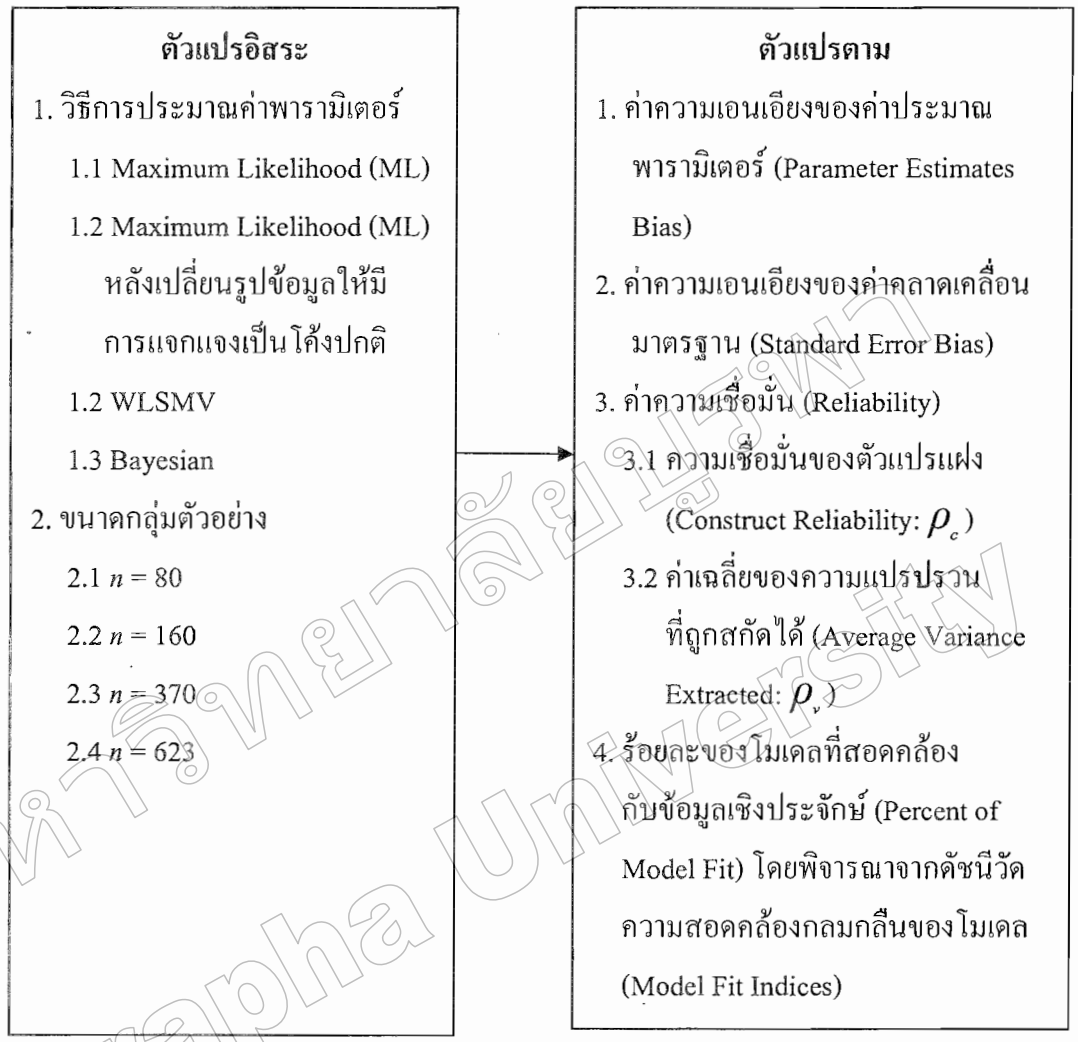
หลักการวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) คือ ตรวจสอบความกลมกลืนระหว่างโมเดลสมมติฐานกับข้อมูลเชิงประจักษ์ การเปรียบเทียบใช้เมตริกซ์ความแปรปรวน – ความแปรปรวนร่วมเป็นตัวเกณฑ์ในการเปรียบเทียบ โดยนำเมตริกซ์ความแปรปรวน – ความแปรปรวนร่วมที่คำนวณได้จากกลุ่มตัวอย่างที่เป็นข้อมูลเชิงประจักษ์ (แทนด้วยสัญลักษณ์ S) มาเทียบกับเมตริกซ์ความแปรปรวน – ความแปรปรวนร่วมที่ถูกสร้างขึ้นจากพารามิเตอร์ที่ประมาณค่าได้จากโมเดลสมมติฐานวิจัย (แทนด้วยสัญลักษณ์ $\hat{\Sigma}$) ถ้าเมตริกซ์ทั้งสองมีค่าใกล้เคียงกัน หมายความว่า โมเดลสมมติฐานวิจัยมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ เนื่องจากเมตริกซ์ $\hat{\Sigma}$ ถูกสร้างจากค่าประมาณของพารามิเตอร์ ดังนั้นการประมาณค่าพารามิเตอร์จึงใช้หลักการวิเคราะห์เปรียบเทียบความกลมกลืนระหว่างเมตริกซ์ดังกล่าวเป็นเงื่อนไขในการประมาณค่าพารามิเตอร์ จุดมุ่งหมายของการประมาณค่าพารามิเตอร์ คือ การหาค่าพารามิเตอร์ที่จะทำให้เมตริกซ์ S และ $\hat{\Sigma}$ มีค่าใกล้เคียงกันมากที่สุด ซึ่งวิธีการประมาณค่าพารามิเตอร์มีหลากหลายวิธีและถูกประยุกต์เพื่ออำนวยความสะดวกในการคำนวณในโปรแกรม

สำเร็จรูปต่าง ๆ เช่น LISREL, EQS, PROC CALIS, AMOS, RAMONA หรือ Mplus ซึ่งวิธีการที่ผู้วิจัยใช้ศึกษาเปรียบเทียบในครั้งนี้ ได้แก่ วิธี Maximum Likelihood (ML), วิธี Robust Weighted Least Square – Mean and Variance Adjusted χ^2 (WLSMV) และ วิธี Bayesian โดยใช้โปรแกรม Mplus เวอร์ชัน 6.0 เนื่องจากวิธี ML เป็นวิธีที่ได้รับความนิยมใช้มากที่สุด (Curran, West, & Finch, 1996, p. 17; Finney & DiStefano, 2006, p. 271) แต่มีข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลต้องมีลักษณะเป็นแบบต่อเนื่อง การแจกแจงของข้อมูลที่ต้องเป็นโค้งปกติ กลุ่มตัวอย่างต้องมีขนาดใหญ่ ซึ่งหลายครั้งข้อมูลที่ได้จากการเก็บรวบรวมไม่ได้เป็นไปตามข้อตกลงเบื้องต้น ดังนั้นการฝ่าฝืนข้อตกลงเบื้องต้นดังกล่าวจะส่งผลกระทบต่อผลลัพธ์การประมาณค่าต่าง ๆ หรือไม่ อย่างไร วิธีหนึ่งที่นักวิจัยส่วนใหญ่จัดการกับข้อมูลที่มีลักษณะดังกล่าว คือ การเปลี่ยนรูปข้อมูล (Data Transforming) เพื่อให้แต่ละตัวแปรมีการแจกแจงเป็นโค้งปกติ และมีลักษณะต่อเนื่อง แล้วจึงดำเนินการประมาณค่าพารามิเตอร์ด้วยวิธีการ ML (Schumacker & Lomax, 2010, p. 28) แต่ยังไม่มีความชัดเจนเกี่ยวกับการแปลผลการวิเคราะห์หลังจากการเปลี่ยนรูปข้อมูล ซึ่งทำให้ธรรมชาติของข้อมูลเปลี่ยนไป จึงเป็นประเด็นที่น่าสนใจศึกษา วิธีการประมาณค่าพารามิเตอร์แบบ WLSMV จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง พบว่า เป็นวิธีการที่มีความแกร่งต่อการละเมิดข้อตกลงเบื้องต้นเกี่ยวกับลักษณะข้อมูลจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) การแจกแจงไม่เป็นโค้งปกติ และกลุ่มตัวอย่างขนาดเล็กได้อย่างมีประสิทธิภาพ (Flora & Curran, 2004, pp. 469 – 471; Finney & DiStefano, 2006, pp. 270 – 271; Bandalos, 2006, p. 389) และวิธีการประมาณค่าพารามิเตอร์แบบ Bayesian เป็นวิธีการที่มีความแกร่งต่อการละเมิดข้อตกลงเบื้องต้นเกี่ยวกับลักษณะข้อมูลจำแนกกลุ่ม การแจกแจงไม่เป็นโค้งปกติ และกลุ่มตัวอย่างขนาดเล็กอย่างมีประสิทธิภาพเช่นเดียวกับวิธี WLSMV แต่การวิเคราะห์ง่าย สะดวก ใช้ระยะเวลาในการคำนวณด้วยคอมพิวเตอร์มีความรวดเร็วกว่าวิธีการอื่น ๆ และให้ผลลัพธ์ที่มีความแม่นยำมากกว่า วิธี WLSMV (Asparouhov & Muthén, 2010, p. 45)

สำหรับเงื่อนไขการกำหนดขนาดกลุ่มตัวอย่าง โดยทั่วไปในการวิเคราะห์สถิติประเภทพหุตัวแปร (Multivariate Statistics) มีหลักการกำหนดขนาดกลุ่มตัวอย่าง คือ การเป็นตัวแทนที่ดีของประชากร ในขณะเดียวกันก็ต้องคำนึงถึงเทคนิคในการประมาณค่าและข้อตกลงเบื้องต้น แต่อย่างไรก็ตามในทางปฏิบัตินักวิจัยยังต้องการใช้กลุ่มตัวอย่างที่น้อยที่สุดเท่าที่จะเป็นไปได้ แม้ว่าจะทราบดีว่าการใช้กลุ่มตัวอย่างขนาดใหญ่จะมีโอกาสที่ตัวแปรมีการแจกแจงเป็นโค้งปกติมากกว่ากลุ่มตัวอย่างที่เล็ก และให้ค่าสถิติที่ค่อนข้างคงที่กว่า ผู้วิจัยเลือกศึกษากลุ่มตัวอย่างโดยแยกตามแนวคิด 2 ลักษณะ ดังนี้

1. กลุ่มตัวอย่างขนาดเล็กตามเกณฑ์อัตราส่วนระหว่างจำนวนตัวอย่าง (n) และตัวแปรสังเกตได้ในโมเดล (p) (Subject – to – variables Ratio) คือ n: p =10: 1 และ n: p = 20: 1 (Gagné & Hancock, 2006 cited in Myers, Ahn, & Jin, 2011, p. 412; Schumacker & Lomax, 2010, p.42; Kline, 2005, p. 178) เนื่องจากขนาดของโมเดลหรือจำนวนตัวแปรสังเกตได้ในโมเดลเป็นตัวแปรที่ทำให้ค่าประมาณของพารามิเตอร์มีค่าแตกต่างกัน จากการศึกษาของ Myers, Ahn, & Jin (2011, p. 412) ได้เสนอแนะว่าเมื่อโมเดลมีขนาดใหญ่ขึ้นจำนวนกลุ่มตัวอย่างควรมีขนาดเพิ่มขึ้นด้วย ซึ่งการคำนวณโดยยึดเกณฑ์นี้ มีความสะดวกและง่ายต่อการคำนวณ และสามารถศึกษาในกลุ่มตัวอย่างขนาดเล็กได้ ดังนั้น ในโมเดลสมมติฐานของการวิจัยครั้งนี้มีจำนวนตัวแปรสังเกตได้ทั้งสิ้น 8 ตัวแปร ผู้วิจัยจึงเลือกใช้จำนวนกลุ่มตัวอย่าง n = 80 และ n = 160

2. วิธีคำนวณขนาดกลุ่มตัวอย่างในการวิเคราะห์องค์ประกอบสำหรับตัวแปรแบบจำแนกกลุ่ม (Categorical Variables) โดยใช้สูตรของ Cochran (1977 cited in Bartlett, Kotrlik, & Higgins, 2001, p. 48) ที่ระดับความเชื่อมั่น 95% และ 99% (α= .05 และ α= .01) เนื่องจากเป็นการคำนวณโดยคำนึงถึงธรรมชาติของข้อมูลที่มีลักษณะจำแนกกลุ่ม พิจารณาถึงจำนวนประชากร และใช้ระดับความเชื่อมั่นมาร่วมคำนวณด้วย ถือเป็นวิธีการที่สอดคล้องกับลักษณะงานวิจัย เพราะได้ขนาดกลุ่มตัวอย่างที่สอดคล้องกับ Rules of Thumb ว่าขนาดกลุ่มตัวอย่างซึ่งเพียงพอที่ทำให้การประมาณค่าสำหรับโมเดล CFA ให้ผลลัพธ์ที่มีความเหมาะสม คือ โมเดลมีความลู่เข้าค่าเดียว (Convergence) ความแม่นยำของค่าสถิติ (Statistical Precision) และความมีประสิทธิภาพของค่าสถิติ (Statistical Power) ควรมีขนาดกลุ่มตัวอย่าง (n) $\geqslant$200 (Marsh, Hau, Balla, & Grayson, 1998 cited in Myers, Ahn, & Jin, 2011, p. 412) จากประชากรเทียมที่ผู้วิจัยสร้างขึ้น จำนวน 10,000 หน่วย จะทำให้ได้จำนวนกลุ่มตัวอย่างอย่างน้อย จำนวน 370 หน่วย และ 623 หน่วย ตามลำดับ จากการศึกษาตัวแปรดังกล่าวผู้วิจัยได้นำเสนอเป็นกรอบแนวคิดในการวิจัยดังภาพที่ 1 – 2



ภาพที่ 1 – 2 กรอบแนวคิดในการวิจัย

สมมติฐานของการวิจัย

1. ภายใต้เงื่อนไขขนาดกลุ่มตัวอย่างเท่ากัน เมื่อใช้วิธีการประมาณค่าพารามิเตอร์ต่างกัน จะให้ผลลัพธ์ของการวิเคราะห์ที่มีความเที่ยงตรงและความเอนเอียงแตกต่างกัน โดยวิธี WLSMV และวิธี Bayesian จะให้ผลลัพธ์ของการวิเคราะห์ที่มีความเที่ยงตรงสูงกว่า/ ความเอนเอียงที่ต่ำกว่าวิธีการ ML ทั้งสองกรณี

2. เมื่อใช้วิธีการประมาณค่าพารามิเตอร์เดียวกัน ภายใต้เงื่อนไขขนาดกลุ่มตัวอย่างที่แตกต่างกัน จะให้ผลลัพธ์ของการวิเคราะห์ที่มีความเที่ยงตรงและความเอนเอียงแตกต่างกัน โดยวิธีการประมาณค่า ทั้ง 4 วิธี จะให้ผลลัพธ์ของการวิเคราะห์ที่มีความเที่ยงตรงสูงขึ้นและความเอนเอียงน้อยลงเมื่อใช้กับกลุ่มตัวอย่างขนาดใหญ่ และผลลัพธ์ของการวิเคราะห์จะมีความเที่ยงตรงน้อยลงและความเอนเอียงเพิ่มสูงขึ้นเมื่อใช้กับกลุ่มตัวอย่างขนาดเล็ก

ข้อตกลงเบื้องต้นของงานวิจัย

1. การศึกษาครั้งนี้ผู้วิจัยใช้ข้อมูลประชากรที่อ้างอิงจากฐานข้อมูลงานวิจัยเชิงสำรวจ โดยคัดเลือกชุดข้อมูลที่มีความสมบูรณ์ จำนวน 10,000 หน่วย และกำหนดเป็นประชากรเทียม ซึ่งถือว่ามีขนาดเพียงพอสำหรับการอ้างอิงไปสู่ประชากรที่แท้จริงได้

2. การศึกษาในครั้งนี้ ดำเนินการบนกรณีของการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับแรก (The First Order Confirmatory Factor Analysis) เท่านั้น ซึ่งโมเดลที่ใช้ในการศึกษาในครั้งนี้ คือ โมเดลการวัด ที่ประกอบด้วย 2 องค์ประกอบ องค์ประกอบละ 4 ตัวแปรสังเกตได้

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ผลการศึกษาครั้งนี้มีความสำคัญต่อการวิจัย โดยเฉพาะขั้นตอนการตรวจสอบความเที่ยงตรงของโมเดลการวัด ของแบบสอบถามที่ใช้เป็นเครื่องมือในการวิจัย

2. ผลการศึกษาครั้งนี้เป็นการเพิ่มพูนความรู้และได้แนวปฏิบัติเกี่ยวกับการวิเคราะห์ข้อมูลที่ใช้โมเดล CFA ซึ่งจะเป็นหลักฐานเชิงประจักษ์เพื่อยืนยันประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์ 4 วิธี คือ วิธี ML, วิธี ML (หลังเปลี่ยนรูปข้อมูลให้มีการแจกแจงเป็นโค้งปกติ), วิธี WLSMV และวิธี Bayesian ว่าวิธีใดมีความเหมาะสมกับข้อมูลที่มีลักษณะจำแนกกลุ่มแบบเรียงอันดับ (Ordered Categorical Data) ที่มีการแจกแจงในระดับแต่ละตัวแปรไม่เป็นโค้งปกติ (Univariate Non – normality Distribution)

3. ผู้วิจัยทางด้านการศึกษา พฤติกรรมศาสตร์ และสังคมศาสตร์ สามารถนำวิธีการวิเคราะห์โมเดล CFA ไปประยุกต์ใช้กับงานวิจัยของตนเอง ซึ่งเป็นองค์ความรู้ที่ค่อนข้างใหม่ของวิทยาการวิจัย ยังไม่มีแนวปฏิบัติที่ชัดเจน ในเรื่องวิธีการประมาณค่าพารามิเตอร์สำหรับข้อมูลที่มีลักษณะจำแนกกลุ่มแบบเรียงอันดับและแจกแจงไม่เป็นโค้งปกติ (Non – normality Ordered Categorical Data) กลุ่มตัวอย่างที่มีขนาดน้อยที่สุด ที่ยังส่งผลให้งานวิจัยมีความน่าเชื่อถือและสอดคล้องกับโครงสร้างของข้อมูลจริงของประชากร

นิยามศัพท์เฉพาะ

การประมาณค่าพารามิเตอร์ (Parameter Estimation) หมายถึง การประมาณค่าพารามิเตอร์ของโมเดล ซึ่งเป็นการวิเคราะห์ข้อมูลจากกลุ่มตัวอย่าง โดยการแก้สมการโครงสร้างเพื่อหาค่าพารามิเตอร์ซึ่งเป็นตัวไม่ทราบค่าในสมการ เป็นการดำเนินการโดยเครื่องคอมพิวเตอร์ ได้จากการใช้ข้อมูลจากกลุ่มตัวอย่าง (ความแปรปรวนและความแปรปรวนร่วมของตัวแปรสังเกตได้) ประมาณค่าพารามิเตอร์ของประชากร

ML หมายถึง วิธีการประมาณค่าพารามิเตอร์แบบ Maximum Likelihood

ML_{tr} หมายถึง วิธีการประมาณค่าพารามิเตอร์แบบ Maximum Likelihood หลังเปลี่ยนรูปข้อมูลให้มีการแจกแจงเป็นโค้งปกติ

WLSMV หมายถึง วิธีการประมาณค่าพารามิเตอร์แบบ Robust Weighted Least Square (Mean and Variance Adjusted χ^2)

Bayesian หมายถึง วิธีการประมาณค่าพารามิเตอร์แบบ Bayesian

โมเดล CFA (Confirmatory Factor Analysis Model) หมายถึง โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน หรือโมเดลการวัด (Measurement Model) ซึ่งเป็นส่วนหนึ่งของโมเดลสมการเชิงโครงสร้าง (Structural Equation Modeling) ซึ่งใช้ในการตรวจสอบเพื่อยืนยันความเที่ยงตรงของโมเดลการวัด (Measurement Model) ที่ประกอบด้วย ชุดตัวแปรสังเกตได้ (Observed Variables) ของตัวแปรแฝง (Latent Variables) เป็นการตรวจสอบความสอดคล้องระหว่างโมเดลสมมติฐานตามทฤษฎีกับข้อมูลเชิงประจักษ์

ความเอนเอียงของค่าประมาณพารามิเตอร์ (Parameter Estimates Bias) หมายถึง ค่าความแตกต่างระหว่างค่าประมาณพารามิเตอร์ที่ได้จากกลุ่มตัวอย่างที่แตกต่างจากค่าพารามิเตอร์ที่แท้จริงจากโมเดลประชากร

ความเอนเอียงของค่าคลาดเคลื่อนมาตรฐาน (Bias of Standard Error) หมายถึง ค่าความแตกต่างระหว่างค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์ที่ได้จากกลุ่มตัวอย่าง ที่แตกต่างจากค่าความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ที่แท้จริงจากโมเดลประชากร

ค่าความเชื่อมั่น (Reliability) หมายถึง ความคงเส้นคงวา หรือความสอดคล้องภายใน (Internal Consistency) ของชุดตัวแปรสังเกตได้ ในแต่ละกลุ่มตัวแปรแฝง เป็นค่าที่แสดงถึงระดับความเป็นตัวแทนตัวแปรแฝงของชุดตัวแปรสังเกตได้ หรือความแม่นยำของโมเดลการวัด

ความเชื่อมั่นของตัวแปรแฝง (Construct Reliability: ρ_c) หมายถึง สัดส่วนความแปรปรวนร่วมกันของตัวแปรสังเกตได้ทั้งหมดในตัวแปรแฝงเดียวกัน หรือเรียกว่า ความตรงในการรวมตัว (Convergent Validity)

ค่าเฉลี่ยของความแปรปรวนที่ถูกสกัดได้ (Average Variance Extracted: ρ_v) หมายถึง ค่าเฉลี่ยของความแปรปรวนของตัวแปรสังเกตได้ที่อธิบายได้ด้วยตัวแปรแฝง

ดัชนีวัดความสอดคล้องกลมกลืนของโมเดล (Model Fit Indices) หมายถึง ดัชนีที่ใช้ในการตรวจสอบความตรงของโมเดลเป็นภาพรวมทั้งโมเดล โดยเป็นค่าดัชนีที่แสดงว่าในภาพรวมโมเดลสมมติฐานตามกรอบแนวคิด มีความสอดคล้องกลมกลืนกับโมเดลของข้อมูลเชิงประจักษ์เพียงใด

การเปลี่ยนรูปข้อมูล (Transforming Data) หมายถึง การดำเนินการปรับเปลี่ยนค่าของข้อมูลที่มีลักษณะการแจกแจงที่ไม่เป็นโค้งปกติให้มีการแจกแจงเป็นโค้งปกติ ด้วยวิธีการทางสถิติ มีวัตถุประสงค์เพื่อแปลงข้อมูลเป็นมาตราใหม่แล้วข้อมูลมีการแจกแจงเป็นโค้งปกติปกติหรือใกล้เคียงกับแบบปกติ ค่าเฉลี่ยและความแปรปรวนของข้อมูลที่แปลงแล้วเป็นอิสระต่อกันจะทำให้ความแปรปรวนมีค่าเท่ากัน