

ปกิณกะ

Miscellany

บทนำของการวิเคราะห์องค์ประกอบ

เสฏฐพงศ์ เลิศสกุลบรรลือ พ.บ.*

ระพีพงศ์ สุพรรณไชยมัตย์ พ.บ., Ph.D.**

* กองระบาดวิทยา กรมควบคุมโรค

** สำนักงานพัฒนานโยบายสุขภาพระหว่างประเทศ

การวิเคราะห์องค์ประกอบ (factor analysis) เป็นวิธีทางสถิติที่ใช้วิเคราะห์ตัวแปรหลายตัว เพื่อสรุปหรือลดจำนวนตัวแปร โดยเป็นเทคนิคในการจับกลุ่มตัวแปรที่มีความสัมพันธ์กันให้อยู่ในกลุ่มเดียวกัน นำไปสู่การตรวจสอบความตรงตามโครงสร้าง (construct validity) ทั้งนี้สามารถใช้วิธีนี้ในการพัฒนาทฤษฎีใหม่หรือทดสอบทฤษฎีเดิมได้

การวิเคราะห์องค์ประกอบ แบ่งเป็น 2 ประเภท คือ การวิเคราะห์องค์ประกอบเชิงสำรวจ (exploratory factor analysis) และการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) โดยการวิเคราะห์องค์ประกอบเชิงสำรวจ ใช้เมื่อผู้ศึกษายังไม่ทราบหรือยังไม่มีแนวคิดเรื่องโครงสร้างความสัมพันธ์ของตัวแปรอย่างชัดเจน มักใช้เพื่อรวบรวมข้อมูลและพัฒนาแนวคิดหรือทฤษฎีใหม่ ส่วนการวิเคราะห์องค์ประกอบเชิงยืนยันใช้เมื่อผู้ศึกษาทราบหรือคาดการณ์ความสัมพันธ์ของตัวแปรได้แล้ว โดยใช้เพื่อตรวจสอบหรือยืนยันว่าตัวแปรที่สนใจมีความสัมพันธ์กับองค์ประกอบที่แฝงอยู่ตรงตามโครงสร้างตามที่คาดการณ์ไว้ นอกจากนี้ การวิเคราะห์องค์ประกอบยังช่วยลดจำนวนตัวแปรโดยรวมตัวแปรที่สัมพันธ์กันให้อยู่ในองค์ประกอบเดียว องค์ประกอบเหล่านี้สามารถสร้างค่าน้ำหนักองค์ประกอบ (factor score) เพื่อใช้เป็นตัวแปรในการวิเคราะห์ทางสถิติต่อไป

เช่น การวิเคราะห์เชิงพรรณนา ความถดถอย ความแปรปรวน หรือโมเดลสมการโครงสร้างที่ซับซ้อนขึ้น อีกทั้งยังช่วยแก้ปัญหาภาวะร่วมเส้นตรงเชิงพหุ (multicollinearity) โดยรวมตัวแปรอิสระที่สัมพันธ์กันและนำองค์ประกอบใหม่มาใช้ในการวิเคราะห์ข้อมูล

ข้อตกลงเบื้องต้นของการใช้สถิติการวิเคราะห์องค์ประกอบ⁽¹⁻³⁾

1) ตัวแปรที่ใช้ในการวิเคราะห์องค์ประกอบควรเป็นตัวแปรต่อเนื่อง หรืออยู่ในมาตราระดับช่วง (interval scale) และมาตราอัตราส่วน (ratio scale) เนื่องจากตัวแปรเหล่านี้ควรมีความสัมพันธ์กัน อย่างไรก็ตาม การใช้ตัวแปรมาตราอันดับ (ordinal scale) ในการวิเคราะห์องค์ประกอบก็ได้รับความนิยม เนื่องจากตัวแปรลำดับแม้จะมีไม่ก็ระดับก็มักถูกมองว่าเป็นตัวแปรต่อเนื่องได้โดยไม่กระทบต่อความแม่นยำของผลลัพธ์ ทำให้สามารถใช้วิธีการประมาณค่ามาตรฐานได้อย่างเหมาะสม⁽⁴⁾

2) ตัวแปรที่เลือกมาวิเคราะห์องค์ประกอบควรมีความสัมพันธ์กันในระดับสูง ($r = 0.30$ ถึง 0.70) โดยรูปแบบความสัมพันธ์ระหว่างองค์ประกอบและตัวแปรควรอยู่ในลักษณะเชิงเส้น

3) จำนวนตัวแปรที่ใช้ในการวิเคราะห์องค์ประกอบควรมากกว่า 10 ถึง 15 ตัว และควรมีอย่างน้อย 3 ถึง 5 ตัวแปร ต่อหนึ่งองค์ประกอบ

4) ตัวแปรที่เลือกมาวิเคราะห์องค์ประกอบควรมีค่าทิศทางเดียวกัน ไม่ว่าจะเป็นบวกหรือลบ

5) ขนาดตัวอย่างควรมีจำนวนมากกว่าจำนวนตัวแปร โดยแนะนำให้มิขนาดตัวอย่าง อย่างน้อย 5 ถึง 10 เท่าของจำนวนตัวแปร

6) การแจกแจงของข้อมูลในแต่ละตัวแปรไม่ควรมีความเบ้มากเกินไป และไม่ควรมีค่าผิดปกติ (outlier) ที่ต่ำสุดหรือสูงสุด เพราะอาจทำให้ผลลัพธ์ที่ได้ไม่แม่นยำ ขั้นตอนและความหมายของค่าที่ใช้ในการวิเคราะห์องค์ประกอบ^(1,3)

ขั้นตอนการวิเคราะห์องค์ประกอบมีขั้นตอนการทดสอบดังนี้

1) กำหนดปัญหาการวิจัย ทบทวนองค์ประกอบตัวแปรตามทฤษฎี เก็บรวบรวมข้อมูล และเลือกวิธีการวิเคราะห์องค์ประกอบให้สอดคล้องกับวัตถุประสงค์ของการวิจัย

2) ตรวจสอบความเหมาะสมของข้อมูลเพื่อการวิเคราะห์องค์ประกอบ โดยใช้สถิติ Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) และ ทดสอบความสัมพันธ์ของตัวแปรต่างๆ โดยใช้สถิติ Bartlett's Test of Sphericity

3) ตรวจสอบความสอดคล้องของข้อมูลที่ใช้วิเคราะห์กับข้อตกลง และสร้าง correlation matrix

4) สกัดองค์ประกอบ (extraction factor analysis) โดยวิธีการสกัดองค์ประกอบมีหลายวิธี เช่น principal component analysis, maximum likelihood และ principal axis factoring ตัวอย่างการสกัดองค์ประกอบในการ

วิเคราะห์องค์ประกอบเชิงสำรวจอย่างง่าย แสดงในภาพที่ 1 และตารางที่ 1

5) เลือกวิธีการหมุนแกน (factor rotation) เพื่อเพิ่มความชัดเจนในการวิเคราะห์ความสัมพันธ์ และทำให้องค์ประกอบที่ได้มีความชัดเจนยิ่งขึ้น

6) ตั้งชื่อองค์ประกอบที่ได้จากการวิเคราะห์โดยควรเลือกชื่อที่ครอบคลุมตัวแปรทั้งหมดในองค์ประกอบเดียวกัน หรืออาจตั้งชื่อตามตัวแปรที่มีน้ำหนักองค์ประกอบสูงสุด

7) เลือกค่าน้ำหนักองค์ประกอบ และอาจนำค่าน้ำหนักองค์ประกอบ ไปวิเคราะห์ทางสถิติต่อไป

ความหมายค่าต่างๆ ในการวิเคราะห์องค์ประกอบที่ต้องรู้ ได้แก่

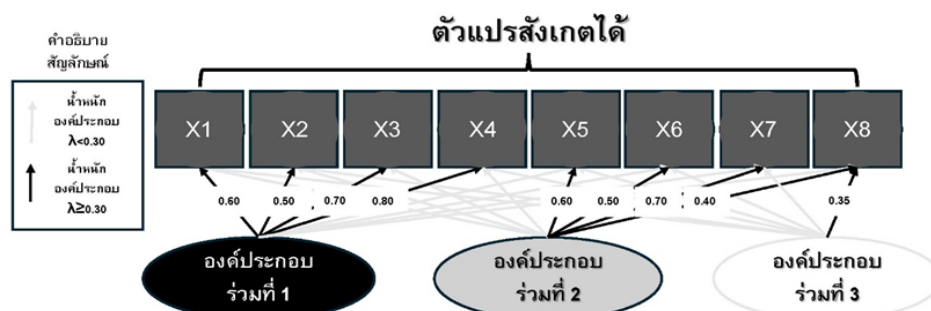
1) องค์ประกอบร่วม (common factor) หมายถึง องค์ประกอบที่มีกลุ่มตัวแปรตั้งแต่ 2 ตัวขึ้นไป รวมอยู่ในองค์ประกอบเดียวกัน โดยอิงจากค่าสัมประสิทธิ์สหสัมพันธ์ที่มีความใกล้เคียงกัน

2) น้ำหนักองค์ประกอบ (factor loading) หมายถึง ค่าความสัมพันธ์ระหว่างตัวแปรกับองค์ประกอบร่วม ซึ่งควรมีค่ามากกว่า 0.30

3) ความร่วมกัน (communalities) แสดงสัดส่วนของความแปรปรวนของแต่ละตัวแปรที่สามารถอธิบายได้โดยองค์ประกอบร่วม ซึ่งคำนวณจากการยกกำลังสองของน้ำหนักองค์ประกอบ

4) ค่าไอเกน (Eigen value) คือ ค่าที่แสดงความผันแปรของตัวแปรทั้งหมดในแต่ละองค์ประกอบในการวิเคราะห์องค์ประกอบ โดยค่าไอเกนของแต่ละ

ภาพที่ 1 แสดงตัวอย่างการสกัดองค์ประกอบในการวิเคราะห์องค์ประกอบเชิงสำรวจอย่างง่าย



ตารางที่ 1 รายละเอียดผลการวิเคราะห์องค์ประกอบเชิงสำรวจจากภาพที่ 1

ตัวแปรสังเกตได้ (Observed Variable)	ความร่วมกัน (Communality)		
	องค์ประกอบร่วมที่ 1	องค์ประกอบร่วมที่ 2	องค์ประกอบร่วมที่ 3
X_1	$(0.60)^2$	$(0.10)^2$	$(0.10)^2$
X_2	$(0.50)^2$	$(0.10)^2$	$(0.10)^2$
X_3	$(0.70)^2$	$(0.10)^2$	$(0.10)^2$
X_4	$(0.80)^2$	$(0.10)^2$	$(0.10)^2$
X_5	$(0.10)^2$	$(0.60)^2$	$(0.10)^2$
X_6	$(0.10)^2$	$(0.50)^2$	$(0.10)^2$
X_7	$(0.10)^2$	$(0.70)^2$	$(0.10)^2$
X_8	$(0.10)^2$	$(0.40)^2$	$(0.35)^2$
ค่าไอเกน (Eigen Value)	1.78	1.26	0.19

ตัวอักษรสีฟ้าแสดงถึงน้ำหนักองค์ประกอบที่ $\lambda < 0.30$ โดยสมมติให้ค่าเป็น 0.10

องค์ประกอบสามารถคำนวณได้จากผลรวมของความร่วมกันขององค์ประกอบนั้นๆ องค์ประกอบที่ 1 จะเป็นองค์ประกอบที่สามารถแยกความผันแปรของตัวแปรออกจากองค์ประกอบอื่นได้มากที่สุด และองค์ประกอบร่วมที่มีค่าไอเกนน้อยกว่า 1 จะถือว่าอธิบายความผันแปรของข้อมูลได้น้อย ดังนั้นในภาพที่ 1 และตารางที่ 1 จึงควรมีองค์ประกอบร่วมเพียง 2 องค์ประกอบเท่านั้น

5) ตัวแปรสังเกตได้ (observed variable) หมายถึง ตัวแปรที่สามารถวัดค่าได้โดยตรงในทางปฏิบัติ

6) ตัวแปรแฝง (latent variable) หมายถึง ตัวแปรที่ไม่สามารถสังเกตได้โดยตรง แต่ถูกคำนวณผ่านแบบจำลองทางคณิตศาสตร์จากตัวแปรสังเกตได้ ซึ่งในกรณีนี้ มักหมายถึงองค์ประกอบร่วม

สรุป

การวิเคราะห์องค์ประกอบเป็นเทคนิคทางสถิติที่ใช้ในการลดจำนวนตัวแปรและจัดกลุ่มตัวแปรที่มีความสัมพันธ์กัน เพื่อตรวจสอบความตรงตามโครงสร้างของข้อมูล การวิเคราะห์องค์ประกอบมีเงื่อนไขสำคัญที่ต้องพิจารณา เช่น การเลือกตัวแปร จำนวนตัวแปร ขนาดตัวอย่าง และการแจกแจงของข้อมูล นอกจากนี้ การวิเคราะห์องค์ประกอบยังช่วยพัฒนาและยืนยันทฤษฎีใหม่ๆ รวมถึงแก้ปัญหาความสัมพันธ์ซ้ำซ้อนระหว่าง

ตัวแปร ช่วยลดความซับซ้อนของการวิเคราะห์ข้อมูล และทำให้การนำเสนอข้อมูลชัดเจนและเข้าใจง่ายขึ้น สุดท้าย องค์ประกอบร่วมและค่าน้ำหนักองค์ประกอบที่คำนวณได้ยังสามารถนำไปใช้ในการวิเคราะห์ทางสถิติต่อไปได้

เอกสารอ้างอิง

1. Angsuchote S, Vijitwanan S, Phinyopanuwat R. Statistical analysis for social science and behavioral science research: techniques for using LISREL. 3rd ed. Bangkok: Charoen Dee Munkong Printing; 2011.
2. Knapp TR, Brown JK. Ten measurement commandments that often should be broken. Res Nurs Health. 1995;18(5):465–469. doi:10.1002/nur.4770180511
3. Kaiyavarn Y. Multivariate statistical analysis for research. 2nd ed. Bangkok: Chulalongkorn University Press; 2014.
4. Robitzsch A. Why ordinal variables can (almost) always be treated as continuous variables: clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. Front Educ. 2020;5. <https://doi.org/10.3389/feduc.2020.589965>.