บทความปริทัศน์ (Review)

Factor Analysis in Medical Education Research

ไพโรจน์ บุญลักษณ์ศิริ

รับบทความ: 11 สิงหาคม 2568
ปรับแก้ไขบทความ: 19 สิงหาคม 2568
ตอบรับตีพิมพ์: 10 ตุลาคม 2568

บทนำ

งานวิจัยรูปแบบสำรวจความเห็นที่ใช้ในทางแพทยศาสตรศึกษานั้นมีจำนวนมาก เก็บข้อมูลโดยใช้แบบสอบถาม (questionnaire) ที่สร้างขึ้นใหม่ตามเนื้อหาที่ต้องการ และวัดผลเป็นคะแนน Likert scale ตัวอย่างเช่น The Dundee Ready Educational Environment Measure (DREEM) ซึ่งมี 50 คำถาม การวิเคราะห์ผลโดยแสดงคะแนนดิบ ตามรายข้อทุกข้อคงจะมากเกินไป ดังนั้น การจัดกลุ่มเนื้อหาที่มีจำนวนข้อคำถามมากตั้งแต่ตอนการสร้างแบบสอบถามโดยมี panel of expert พิจารณานอกจาก content validity และยังรวมถึง construct validity ด้วย ซึ่งแปลผลด้วยคะแนนตามกลุ่มจะทำให้เข้าใจง่ายขึ้น นอกจากนี้ยังจัดกลุ่มภายหลังการสำรวจโดยใช้สถิติวิเคราะห์มาช่วยในเรื่อง construct validity ได้ด้วยเทคนิคที่นิยมใช้ได้แก่ Factor analysis

Factor analysis คืออะไร

เป็นวิธีทางสถิติเพื่อลดจำนวนตัวแปรหรือข้อคำถามในแบบสอบถาม โดยรวมกลุ่มข้อคำถามที่มีความสัมพันธ์กันและอาจจะตัดบางข้อออก ถ้าจัดเข้ากลุ่มใดไม่ได้ ทำให้วิเคราะห์สรุปผลและประยุกต์ใช้งานได้ง่ายขึ้น การสร้างตัวแปรแฝง (latent variable) ใหม่นี้เรียกว่า factor วิเคราะห์จาก correlation matrix หรือ covariance matrix ของตัวแปร ซึ่งได้มีการคำนวณค่าเฉพาะ โดยสื่อถึงปริมาณของ variance ที่อธิบายแต่ละ factor และทิศทางได้ เรียกค่าเหล่านี้ว่า eigenvalues และ eigenvectors ตามลำดับ สูตรที่ใช้สำหรับคำนวณมีความซับซ้อน แต่โปรแกรมสถิติทั่วไปสามารถคำนวณค่าเหล่านี้ให้ได้ สมการอธิบายอย่างง่ายคือ eigenvalue ของแต่ละ factor = จำนวนตัวแปรมาตรฐานที่ factor นั้นสามารถอธิบายได้ การแปลผลที่ใช้ Kaiser's criterion โดยมีการเลือก factors ที่มี eigenvalue มากกว่า 1 ขึ้นไป นำมาใช้ค่า eigenvalue และค่า eigenvector จะได้นำเอามาคำนวณถึงสัมประสิทธิความสัมพันธ์ของแต่ละข้อคำถามในแต่ละ factors เรียกว่า factor loading ที่ได้คำนวณจากสมการดังต่อไปนี้ Factor loading = eigenvector × sq.root (eigenvalue) มีค่าพิสัย -1 to +1 ข้อกำหนดเบื้องต้น (assumption) ของข้อมูลสำหรับ Factor analysis มีดังนี้ 1. No outlier 2. Sufficient sample size (จำนวนผู้ที่ตอบแบบสอบถามควรมีมากกว่าข้อคำถามจำนวน 3 - 5 เท่าขึ้นไป และข้อคำถามที่มีมากกว่าจำนวน factors) 3. Interval scale 4. No perfect multi- collinearity สามารถตรวจสอบความเหมาะสมของข้อมูลที่ใช้วิเคราะห์จาก correlation matrix ก่อน หรือใช้สถิติเฉพาะโดยใช้ 2 วิธีต่อไปนี้คู่กัน ได้แก่ 1. Test of the identity of correlation matrix โดย Bartlett's test of sphericity ถ้าพบ p-value < 0.05 แสดงว่าเหมาะสมในการสกัด factors และ 2. Kaiser-Meyer-Olkin measure of sampling adequacy, KMO test บอกขนาด partial correlation ตั้งแต่ 0.6 ขึ้นไปถือว่าเหมาะสม ใช้คำสั่ง STATA: factor test (สามารถที่จะเพิ่มคำสั่งนี้ได้จาก Stata software center, ssc install)

หลักการวิเคราะห์มี 2 แบบ ได้แก่

1. Exploratory factor analysis (EFA) ซึ่งนำมาใช้บ่อยในแบบสอบถามที่สร้างใหม่ทั่วไปที่ยังไม่มีโครงสร้างใดๆ

2. Confirmatory factor analysis (CFA) เป็นวิธีที่ซับซ้อนขึ้น สำหรับแบบสอบถามที่มีโครงสร้างอยู่แล้ว

ในที่นี้จะใช้ EFA โดยใช้คำสั่งในโปรแกรม STATA แสดงตัวอย่างประกอบ ด้วย basic extraction method ที่เป็น default ของโปรแกรมคือ principal factor method (pf) โดยพิจารณาจาก common variance (communality) ที่อธิบายความสัมพันธ์ของตัวแปรร่วม ซึ่งมีค่าเริ่มต้นจาก squared multiple correlations (SMC) วิธีอื่นที่มีชื่อที่คล้ายกัน ได้แก่ principal component factor (pcf) ใช้ total variance และค่าเริ่มต้น variance = 1 ในทุกตัวแปร ซึ่งอ้างว่าเสถียรกว่า pf และใช้กับข้อมูลจำนวนน้อยได้ใน STATA command ต้องใส่ pcf ในส่วนของ option เข้าไปด้วย ตัวอย่าง ประเมินบรรยากาศการเรียนด้วย 20-item questionnaire with 4-point Likert scale ทำ 2 ขั้นตอนหลัก ใช้คำสั่ง STATA ดังนี้

1. factor q1-q20 คำสั่งที่ใช้หลัก STATA command: factor จะแสดง output เป็นตารางจัดกลุ่ม factors พร้อมค่า eigenvalues และ factor loadings ของข้อคำถามที่ยังไม่ได้ปรับหมุนแกนความสัมพันธ์ของการแปลผล ให้เลือก factors ที่มีค่าของ eigen values>1 มาใช้ การพิจารณาจำนวน factors กับค่า eigenvalues ตรวจสอบได้ง่ายด้วยการสร้างกราฟในระหว่างจำนวน factor (x) และ eigenvalue (y) ที่เรียกว่า scree plot จุดที่กราฟมีการเลี้ยวหักศอกคือจุดที่ใช้นำมากำหนดจำนวน factors, STATA command: scree plot

2. rotate, varimax ซึ่งการใช้คำสั่งต่อไปคือ STATA command: rotate เป็นการปรับค่า factor loadings ชัดเจนขึ้น โดยการหมุนแกน vector ของ factors วิธีที่นิยมใช้กันคือทำการหมุนแกนตั้งฉากกัน (orthogonal varimax rotation) ให้แต่ละ factor เป็นอิสระจากกัน (uncorrelated) อธิบายอย่างง่าย คือ ให้มองเห็นรูปภาพตั้งตรงตามกรอบที่ตั้งฉาก (STATA ใช้ varimax เป็น default) แต่ถ้าปรับแบบ partial correlated โดยการหมุนแกนเป็นมุมแหลม (oblique Promax rotation) ให้เห็นภาพเบ้ไปตามกรอบที่ไม่ตั้งฉากใน STATA ให้ใส่ Promax เป็น option เพิ่มเข้าไป

การแปลผลจากตัวอย่างสมมุตินี้ จัดกลุ่มหรือลดตัวแปรเหลือ 4 factors (Figure1) โดยเลือกจาก eigenvalues ที่มีมากกว่า 1 เท่านั้น output ซึ่งแสดงตาราง factor loading ของข้อคำถาม (q) ตาม factor ที่หมุนแกนความสัมพันธ์แล้ว ให้เลือกค่า 0.4 ขึ้นไป เป็นค่าที่พอยอมรับได้ (ค่าที่ดีคือ 0.7 ขึ้นไป ทิศทางบวกหรือทิศทางลบก็ได้) มีค่าน้อยกว่า 0.4 ในทุก factors ซึ่งอาจจะสามารถที่จะตัดออกได้ ในที่นี้ factor1 โดยมี q 2,6,7,11,12,13,18,20 การเขียนผลนำเสนอให้ตั้งชื่อ factor โดยสื่อแทนเนื้อหาของกลุ่ม เช่น ด้านการสอนและผู้สอนใน factors ที่เหลือก็ทำแบบเดียวกัน เมื่อได้ factors แล้วสามารถให้โปรแกรมสร้างหรือทำนายข้อมูลคะแนนของแต่ละ factor ขึ้นมาใหม่ได้ด้วย เรียกว่า factor scores โดยใช้หลักการ regression method ที่นำไปใช้วิเคราะห์เชิงลึกอื่นต่อไปได้ ในที่นี้ใช้คำสั่ง STATA command: predict f1 f2 f3 f4 โปรแกรมจะสร้าง column ข้อมูลเพิ่มให้ใน dataset

จากตารางที่ 1 แสดงค่า uniqueness ในส่วน column สุดท้ายแปลผล คือ ค่าที่บ่งบอกความเป็นเอกลักษณ์ของข้อคำถามนั้น ถ้าค่าสูงแสดงว่าสัมพันธ์กับข้ออื่นน้อย หรือแปลผกผันกับค่า communality นั้นเอง ซึ่งจะสังเกตได้ว่าถ้าค่ายิ่งน้อยจะให้ค่า factor loading สูง

Variable	Factor1	Factor2	Factor3	Factor4	Uniqueness
q1	0.1792	0.5547	0.0726	0.2456	0.5946
q2	0.5904	0.2237	0.0779	0.2887	0.5120
q3	0.2457	0.5466	0.1229	0.1874	0.5907
q4	0.0017	-0.0185	0.3798	-0.7675	0.2663
q5	0.0096	0.2462	0.3908	0.5206	0.5155
q6	0.7752	0.0417	-0.1197	-0.0633	0.3791
q7	0.8129	0.2593	0.0359	-0.0131	0.2705
q8	-0.0705	0.0684	0.8849	-0.1154	0.1940
q9	0.0217	-0.0143	0.9177	-0.0746	0.1516
q10	0.2559	0.2346	0.3637	-0.1182	0.7333
q11	0.7921	0.0536	-0.0714	-0.0086	0.3645
q12	0.4887	0.3550	0.2852	-0.0970	0.5444
q13	0.5348	0.0033	0.1774	0.0379	0.6810
q14	0.2545	0.4491	0.1955	0.3767	0.5534
q15	0.0436	0.7654	-0.1120	0.0338	0.3986
q16	0.1318	0.5247	0.1565	0.0511	0.6802
q17	0.0373	-0.1822	0.8051	0.0692	0.3125
q18	0.7208	-0.1325	-0.0164	0.2406	0.4047
q19	0.0499	0.8357	-0.1437	0.0236	0.2779
q20	0.5890	0.5580	0.1263	-0.1254	0.3100

Figure1 factor loadings after orthogonal varimax rotation

เอกสารอ้างอิง (References)

1. Tavakol M, Wetzel A. Factor analysis: a means for theory and instrument development in support of construct validity. Int J Med Educ. 2020 :11: 245-247.
2. Shrestha N. Factor analysis as a Tool for Survey Analysis. Am J Appl Math Stat. 2021. :9(1); 4-11.