

ความนำ

การวิเคราะห์ตัวประกอบ (Factor Analysis) และการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) เป็นวิธีการทางสถิติของตัวแปรพหุที่อธิบายความสัมพันธ์ของความแปรปรวนร่วมระหว่างตัวแปรหลายๆ ตัว ด้วยตัวประกอบหรือปัจจัย (Factor) หรือองค์ประกอบ (Component) เพียง 2-3 องค์ประกอบซึ่งเป็นปริมาณสุ่ม (Random quantities) ที่ไม่สามารถสังเกตได้โดยตรง โดยองค์ประกอบที่ได้จะสามารถอธิบายลักษณะความสัมพันธ์ของตัวแปรดังกล่าว รวมทั้งอาจรวมลักษณะเฉพาะกลุ่มของตัวแปรได้ด้วย อาจกล่าวอีกนัยหนึ่งคือ เป็นวิธีการศึกษาความสัมพันธ์ระหว่างตัวแปรหลายๆ ตัวที่มีความสัมพันธ์ซึ่งกันและกัน เพื่อนำไปสู่การจัดกลุ่มตัวแปรให้เป็นองค์ประกอบเพียงไม่กี่องค์ประกอบ โดยที่ตัวแปรในแต่ละองค์ประกอบจะมีความสัมพันธ์กับตัวแปรในองค์ประกอบนั้นๆ มากกว่าตัวแปรในองค์ประกอบอื่น รวมทั้งเพื่อช่วยตีความหมายของกลุ่มตัวแปรที่รวมอยู่ในองค์ประกอบเดียวกันตามความหมายของตัวแปรที่มีความสัมพันธ์สูงกับองค์ประกอบนั้น โดยทั้ง 2 วิธีการมีจุดมุ่งหมายที่คล้ายกัน แตกต่างกันที่การวิเคราะห์ตัวประกอบมีพื้นฐานของตัวแบบเชิงสถิติเพื่อใช้ในการประมาณค่าน้ำหนักขององค์ประกอบ (Factor Loading) และค่าความแปรปรวนเฉพาะ (Specific Variance) ที่ต้องอาศัยวิธีการประมาณด้วยวิธีความควรจะเป็นสูงสุด (Maximum Likelihood Method) หรือ วิธีองค์ประกอบหลัก (Principal Component Method) ในขณะที่การวิเคราะห์องค์ประกอบหลักไม่มีตัวแบบเชิงสถิติเข้ามาเกี่ยวข้อง เพียงแต่เป็นการสร้างเซตของตัวแปรใหม่ (องค์ประกอบ) ให้อยู่ในรูปผลรวมเชิงเส้น (Linear Combination) ของตัวแปรเดิมเพื่อตีความลักษณะของกลุ่มตัวแปรได้ง่ายขึ้น

ปัญหาที่สำคัญประการหนึ่งในการวิเคราะห์ตัวประกอบ หรือการวิเคราะห์องค์ประกอบหลัก คือ การพิจารณาถึงจำนวนองค์ประกอบว่าควรประกอบไปด้วยกี่องค์ประกอบ และจะใช้เกณฑ์ใดในการพิจารณาถึงจำนวนองค์ประกอบที่เหมาะสม หากมีการตัดองค์ประกอบบางองค์ประกอบไป หรือองค์ประกอบที่ใช้ในการสรุปมากเกินไป จะทำให้ผลสรุปที่ได้ไม่สามารถอธิบายลักษณะตัวแปรจากงานวิจัยที่ศึกษาได้อย่างชัดเจนและถูกต้อง โดยเกณฑ์ที่ใช้ในการพิจารณาถึงจำนวนองค์ประกอบที่เหมาะสมมีหลายเกณฑ์ เช่น

1. กฎของ Kaiser (Kaiser's Eigenvalue Greater than 1.0 Rule)

เป็นกฎที่เสนอโดย Kaiser (1960) ซึ่งพิจารณาจำนวนองค์ประกอบจากค่าไอเกนที่มีค่ามากกว่า 1.0 กล่าวคือ หากองค์ประกอบใดมีค่าไอเกนน้อยกว่า 1.0 แล้ว แสดงว่าองค์ประกอบนั้นไม่ควรนำไปอธิบายตัวแปรทั้งหมดที่ศึกษา

2. การทดสอบโดย Cattell's Scree Test

เป็นแผนภาพที่เสนอโดย Cattell (1966) ซึ่งพิจารณาจำนวนองค์ประกอบที่เหมาะสมจากแผนภาพ Scree Plot กล่าวคือ แผนภาพจะพลอตกราฟเส้นระหว่างค่าไอเกนแต่ละค่า (แกน y) กับลำดับขององค์ประกอบที่สัมพันธ์กับค่าไอเกนดังกล่าว (แกน x) โดยเรียงจากค่าไอเกนที่มากที่สุดถึงค่าไอเกนที่น้อยที่สุด แนวคิดของ Scree Plot คือ ค่าไอเกนในตำแหน่งท้าย ๆ จะมีค่าน้อยมาก สามารถพิจารณาจากจุดที่มีลักษณะที่เรียกว่า ข้อศอก (Elbow) ซึ่งจากจุดนั้นค่าไอเกนจะค่อยๆ ลดลงอย่างช้าๆ (ความชันเริ่มน้อยลง) ซึ่งองค์ประกอบที่สัมพันธ์กับค่าไอเกนเหล่านั้น จะให้ค่าความแปรปรวนที่เพิ่มขึ้นน้อยมาก จึงสามารถละทิ้งได้

3. การทดสอบโดย Bartlett's Chi-square Test

เป็นวิธีการที่เสนอโดย Bartlett (1950) ซึ่งมีการทดสอบสมมติฐานโดยใช้ตัวสถิติทดสอบแบบไค-สแควร์ เพื่อทดสอบการเท่ากันของกลุ่มค่าไอเกน โดยมีการแบ่งกลุ่มค่าไอเกนอย่างคร่าวๆ ระหว่างค่าไอเกนที่มีขนาดใหญ่ กับค่าไอเกนที่มีขนาดเล็ก และทดสอบสมมติฐานการเท่ากันของกลุ่มค่าไอเกนที่มีขนาดเล็ก หากไม่สามารถปฏิเสธสมมติฐานดังกล่าวได้ แสดงว่ากลุ่มองค์ประกอบที่สัมพันธ์กับกลุ่มค่าไอเกนที่มีขนาดใหญ่ ควรนำไปใช้ในการสรุป และแปลผลต่อไป

ในขณะที่กลุ่มองค์ประกอบที่สัมพันธ์กับกลุ่มไอเกนที่มีขนาดเล็ก ก็จะสามารถละทิ้งองค์ประกอบนั้นไปได้ ซึ่งสามารถคำนวณได้จาก

$$\chi^2 = - \left[n - \frac{1}{6}(2p+1) \right] \ln|R|$$

เมื่อ $|R|$ แทน ดีเทอร์มิแนนต์ของเมทริกซ์สหสัมพันธ์ n แทนจำนวนขนาดตัวอย่าง p แทน จำนวนตัวแปร และ χ^2 ประมาณเป็นการแจกแจงแบบไค-สแควร์ด้วยองศาอิสระ $p(p-1)/2$

4. วิธี Velicer's Minimum Average Partial (MAP)

เป็นวิธีการที่เสนอโดย Velicer (1976) ซึ่งมีการประยุกต์หลักการของการวิเคราะห์องค์ประกอบหลักกับการวิเคราะห์ลำดับที่ของเมทริกซ์สหสัมพันธ์บางส่วน นั่นคือ เป็นการหาค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่าง 2 ตัวแปร ในทุกคู่ของตัวแปรที่เป็นไปได้ในองค์ประกอบที่พิจารณา ซึ่งจะคงองค์ประกอบดังกล่าวไว้ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์อย่างน้อย 2 ค่าในองค์ประกอบมีสหสัมพันธ์กันสูง วิธีการนี้เป็นการพยายามหาองค์ประกอบหลักที่ดีที่สุดมากกว่าการหาจุดตัดเพื่อตัดองค์ประกอบที่ให้ผลของความแปรปรวนต่ำ ซึ่งสามารถคำนวณได้จาก

$$\bar{f}_k = \sum_{\substack{i=1 \\ i \neq j}}^p \sum_{\substack{j=1 \\ i \neq j}}^p (r_{ij,k})^2 / p(p-1)$$

เมื่อ p แทน จำนวนตัวแปร หรือองค์ประกอบ $r_{ij,k}$ แทน สัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปร i และ j โดยการขจัดองค์ประกอบแรกจำนวน k องค์ประกอบ

5. การทดสอบโดย Lawley's Test

เป็นวิธีการที่เสนอโดย Lawley (1956) ซึ่งมีการทดสอบสมมติฐานโดยใช้ตัวสถิติทดสอบแบบไค-สแควร์ เพื่อทดสอบการมีนัยสำคัญของค่าไอเกนจากองค์ประกอบหลักที่ 2 มีการตั้งสมมติฐานว่างเกี่ยวกับความสัมพันธ์ระหว่างองค์ประกอบ 2 องค์ประกอบ ซึ่งสามารถคำนวณได้จาก

$$\chi^2 = \frac{n-1}{1-\bar{r}} \sum_{\substack{i=1 \\ i \neq j}}^p \sum_{\substack{j=1 \\ i \neq j}}^p (r_{ij} - \bar{r})^2 - \mu \sum_{k=1}^p (\bar{r}_k - \bar{r})^2$$

เมื่อ p แทน จำนวนตัวแปร หรือองค์ประกอบ r_{ij} แทน สัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปร i และ j โดย

$$\bar{r} = \frac{2}{p(p-1)} \sum_{i=k+1}^p \sum_{j=1}^p r_{ij} \quad \mu = \frac{(p-1)^2(1-(1-\bar{r})^2)}{p-(p-2)(1-\bar{r})^2} \quad \bar{r} = \frac{1}{(p-1)} \sum_{\substack{i=1 \\ i \neq k}}^p r_{ik}$$

6. วิธี Broken-stick

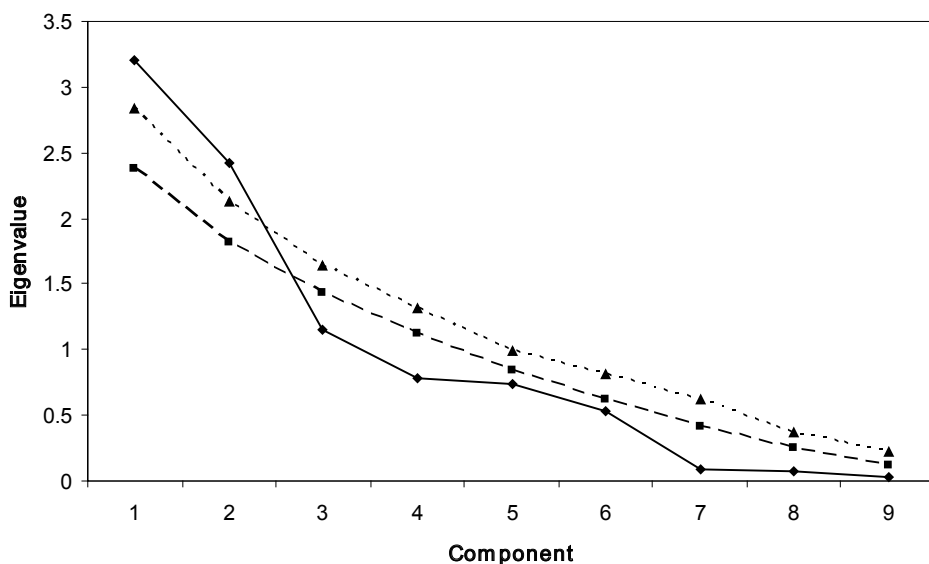
เป็นวิธีการที่เสนอโดย Frontier (1976 อ้างถึงใน Peres-Neto, et al. 2003) ซึ่งมีแนวทางการพิจารณา ค่าไอเกนจากข้อมูลจำลองแบบ โดยมีข้อตกลงว่าหากความแปรปรวนรวม (ผลรวมของค่าไอเกน) ที่ถูกแบ่งอย่างสุ่ม แก่องค์ประกอบต่างๆ แล้ว ค่าคาดหวังของการแจกแจงของค่าไอเกนจะมีการแจกแจงแบบ Broken-stick นั่นคือ หากค่าไอเกนจากข้อมูลจริงมีค่ามากกว่าค่าไอเกนที่ได้จากการคำนวณด้วยตัวแบบ Broken-stick แสดงว่าองค์ประกอบนั้น เหมาะสมที่จะใช้ในการอธิบายตัวแปรที่ศึกษา ซึ่งสามารถคำนวณได้จาก

$$b_k = \sum_{i=k}^p \frac{1}{i}$$

เมื่อ p แทน จำนวนตัวแปร b_k แทน ขนาดของค่าไอเกนจากองค์ประกอบที่ k ภายใต้ตัวแบบ Broken-stick

7. วิธี Horn's Parallel Analysis (PA)

เป็นวิธีการที่เสนอโดย Horn (1965) ซึ่งมีการจำลองข้อมูลจากการแจกแจงแบบปกติที่มีลักษณะเดียวกับ ข้อมูลจริง เพื่อให้ได้ค่าไอเกนในแต่ละครั้งของการจำลองแบบ จากนั้นหาค่าเฉลี่ยของค่าไอเกน และหาค่าไอเกนตำแหน่ง เปอร์เซ็นไทล์ที่ 95 เพื่อเปรียบเทียบกับค่าไอเกนจากข้อมูลจริง นั่นคือ หากค่าไอเกนจากข้อมูลจริงมีค่ามากกว่าทั้ง ค่าเฉลี่ยของค่าไอเกน และค่าไอเกนตำแหน่งเปอร์เซ็นไทล์ที่ 95 แสดงว่าองค์ประกอบนั้นเหมาะสมที่จะใช้ในการอธิบาย ตัวแปรที่ศึกษา ดังตัวอย่างจากภาพที่ 1 ซึ่งเป็นการพลอตกราฟเส้นระหว่างค่าไอเกน และลำดับขององค์ประกอบที่สัมพันธ์ กับค่าไอเกน จะเห็นได้ว่ามีเพียง 2 ค่าแรกของค่าไอเกนจากข้อมูลจริง (Real Data) ที่อยู่เหนือเส้นประยาว และเส้นประ ซึ่งเป็นการพลอตของค่าเฉลี่ยของค่าไอเกน (Mean) และกราฟเส้นของค่าไอเกนตำแหน่งเปอร์เซ็นไทล์ที่ 95 (95th Percentile) ตามลำดับ จึงสามารถสรุปได้ว่าข้อมูลชุดนี้ควรประกอบไปด้วย 2 องค์ประกอบที่มีนัยสำคัญทางสถิติ ที่ระดับ .05 ตามที่ Glorfeld (1995) ได้เสนอว่า เพียงแค่ใช้ค่าไอเกนตำแหน่งเปอร์เซ็นไทล์ที่ 95 ก็เพียงพอต่อการ พิจารณามีนัยสำคัญทางสถิติที่ระดับ .05



ภาพที่ 1 กราฟเส้นค่าไอเกนจากข้อมูลจริง ค่าเฉลี่ยของค่าไอเกน และค่าไอเกนตำแหน่งเปอร์เซ็นไทล์ที่ 95

จากงานวิจัยของ Humphreys and Montanelli (1975) ซึ่งได้ประเมินความถูกต้องของเกณฑ์ในการพิจารณาจำนวนองค์ประกอบโดยเปรียบเทียบระหว่างวิธี MAP และวิธี PA โดยจำลองข้อมูลจากเมทริกซ์สหสัมพันธ์ที่มีจำนวนตัวแปรเท่ากับ 20 และ 40 ตัว ค่าสังเกตขนาด 100 และ 500 ค่า เพื่อพิจารณาจำนวนองค์ประกอบ 3 และ 7 องค์ประกอบ ได้ทำซ้ำ 40 และ 50 ซ้ำในแต่ละเงื่อนไข ผลการศึกษาชี้ให้เห็นว่า วิธี MAP ให้ผลของจำนวนองค์ประกอบที่มากเกินไปและยังให้ผลที่แย่งอีก เมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้น ในขณะที่วิธี PA ให้ผลของจำนวนองค์ประกอบได้ถูกต้องเกือบทุกกรณีการศึกษา

Zwick and Velicer (1986) ได้ประเมินความถูกต้องของเกณฑ์ในการพิจารณาจำนวนองค์ประกอบโดยเปรียบเทียบเกณฑ์ทั้งหมด 5 เกณฑ์ คือ กฎของ Kaiser การทดสอบโดย Cattell's scree test การทดสอบโดย Bartlett's chi-square test วิธี MAP และวิธี PA โดยจำลองข้อมูลจากเมทริกซ์สหสัมพันธ์ที่มีจำนวนตัวแปรเท่ากับ 36 และค่าสังเกตขนาด 72 และ 180 จำนวนตัวแปรเท่ากับ 72 ตัวและค่าสังเกตขนาด 144 และ 360 ค่าน้ำหนักองค์ประกอบเท่ากับ 0.5 และ 0.8 รวมทั้งสัดส่วนจำนวนตัวแปรและจำนวนองค์ประกอบหลักเท่ากับ 6 และ 12 ได้ประเมินความถูกต้องจากค่าเฉลี่ยของความแตกต่างระหว่างจำนวนองค์ประกอบหลักที่ได้จากการใช้เกณฑ์นั้น ๆ กับจำนวนองค์ประกอบหลักที่ได้จากผลที่ศึกษาทั้งหมดโดยอ้างอิงจากค่าไอเกนที่มีค่ามากกว่า 1.0 และมีค่าน้ำหนักองค์ประกอบมากกว่า 0.3 ผลการศึกษาเปรียบเทียบพบว่า วิธี PA ให้ผลของความถูกต้องได้ดีที่สุดคิดเป็นร้อยละ 92 รองลงมาคือวิธี MAP ให้ผลของความถูกต้องร้อยละ 84 การทดสอบโดย Cattell's scree test ให้ผลของความถูกต้องร้อยละ 57 ในขณะที่ยังพบว่าการทดสอบโดย Bartlett's chi-square test ให้ผลของความถูกต้องเพียงร้อยละ 30 ส่วนกฎของ Kaiser ให้ผลของความถูกต้องเพียงร้อยละ 22 และยังพบว่าให้จำนวนองค์ประกอบที่มากเกินไปความจำเป็น Zwick and Velicer (1986) ยังพบว่ากฎของ Kaiser จะให้ผลของความถูกต้อง เมื่อจำนวนองค์ประกอบหลักต่อจำนวนตัวแปรที่ศึกษาอยู่ในอัตราส่วน 1 : 3 หรือ 1 : 5 หรือ 1 : 6 และ Fabrigar, et al. (1999) พบว่าหากค่าไอเกนที่ได้มีค่าใกล้เคียง 1.0 มากๆ เช่น 1.01 หรือ 0.99 ย่อมเกิดปัญหาในการตัดสินใจต่อผู้วิจัยได้ สำหรับการทดสอบโดย Cattell's scree test นั้น หากค่าความชันระหว่าง 2 จุด มีอยู่หลายช่วงในแผนภาพ (ดังเช่นภาพที่ 1 ซึ่งอาจจะต้องเลือกระหว่าง 4 หรือ 7 องค์ประกอบ) จะเป็นการยากในการตัดสินใจว่า ควรจะประกอบไปด้วยกี่องค์ประกอบ แต่หากเปรียบเทียบระหว่างวิธีการนี้กับกฎของ Kaiser จะพบว่า วิธีการนี้ให้ผลของความถูกต้องมากกว่า และไม่ซับซ้อนกำหนดในเรื่องจำนวนตัวแปรที่ศึกษา

Silverstein (1987) ได้ประเมินความถูกต้องของเกณฑ์ในการพิจารณาจำนวนองค์ประกอบโดยเปรียบเทียบกฎของ Kaiser กับวิธี PA โดยจำลองข้อมูล 24 ชุดข้อมูลที่มีการแจกแจง จำนวนค่าสังเกต และจำนวนตัวแปรต่างกัน โดยประเมินความถูกต้องจากค่าเฉลี่ยของความแตกต่างระหว่างจำนวนองค์ประกอบหลักที่ได้จากการใช้กฎ Kaiser หรือวิธี PA กับจำนวนองค์ประกอบหลักที่ได้จากผลที่ศึกษาทั้งหมด โดยอ้างอิงจากค่าไอเกนที่มีค่ามากกว่า 1.0 และร้อยละสะสมของความแปรปรวนที่อธิบายได้ตั้งแต่ 70% ขึ้นไป ผลการศึกษาแสดงว่า วิธี PA ให้ผลของความถูกต้องได้ดีกว่ากฎของ Kaiser

Glorfeld (1995) ได้เสนอให้ใช้ค่าไอเกนตำแหน่งเปอร์เซ็นต์ไทล์ที่ 95 ในวิธี PA แทนค่าเฉลี่ยของค่าไอเกน ซึ่งค่าไอเกนดังกล่าวจะสอดคล้องกับนัยสำคัญทางสถิติที่ระดับ .05 โดยจำลองข้อมูลจากเมทริกซ์สหสัมพันธ์ที่มีจำนวนค่าสังเกต 90 ค่า และจำนวนตัวแปร 9 ตัว จำนวนค่าสังเกต 180 ค่า และจำนวนตัวแปร 36 ตัว ซึ่งมีการแจกแจงใน 4 รูปแบบคือ 1) การแจกแจงแบบปกติมาตรฐาน 2) การแจกแจงแบบเอกรูป 3) การแจกแจงแบบ G ที่มีการเอนเอียงทางขวาในระดับสูง และ 4) การแจกแจงแบบผสมที่มาจากการแจกแจงแบบปกติ แบบเอกรูป แบบเอนเอียงทางขวาและซ้าย การทำซ้ำทั้งหมด 5,000 ซ้ำ ผลการคำนวณค่าไอเกนลำดับที่ 1-3 จากเมทริกซ์สหสัมพันธ์ที่มีจำนวนค่าสังเกต 90 ค่า และจำนวนตัวแปร 9 ตัว และค่าไอเกนลำดับที่ 1-12 จากเมทริกซ์สหสัมพันธ์ที่มีจำนวนค่าสังเกต 180 ค่า และจำนวน

ตัวแปร 36 ตัว พบว่าวิธี PA ให้ค่าเฉลี่ยของค่าไอเกน ค่ามัธยฐานของค่าไอเกน และค่าไอเกนตำแหน่งเปอร์เซ็นต์ไทล์ที่ 95 จากทุกการแจกแจงมีค่าที่ใกล้เคียงกัน อย่างไรก็ตาม Glorfeld (1995) ยังได้จำลองเมทริกซ์สหสัมพันธ์ที่มีจำนวนค่าสังเกต 120 ค่าและจำนวนตัวแปร 12 ตัว การทำซ้ำทั้งหมด 5,000 ซ้ำ เพื่อศึกษาค่าเฉลี่ยของค่าไอเกน และค่าไอเกนตำแหน่งเปอร์เซ็นต์ไทล์ที่ 95 ผลที่ได้พบว่า ค่าเฉลี่ยของค่าไอเกนจะให้จำนวนองค์ประกอบที่มากเกินไปจนความจำเป็น จึงได้เสนอให้ใช้ค่าไอเกนตำแหน่งเปอร์เซ็นต์ไทล์ที่ 95 แทนค่าเฉลี่ยของค่าไอเกน ซึ่งค่าไอเกนดังกล่าวจะสอดคล้องกับนัยสำคัญทางสถิติที่ระดับ .05

เมื่อพิจารณาจากงานวิจัยข้างต้น หรืออีกหลายๆ งานวิจัยดังเช่น งานวิจัยของ Jackson (1993); Press-Neto, Jackson, and Somers (2003; 2005) Dinno (2009) พบว่า วิธี PA ให้ผลของความถูกต้องเกี่ยวกับจำนวนองค์ประกอบที่เหมาะสมกว่าวิธีการอื่นๆ เช่น การใช้กฎของ Kaiser วิธี MAP การทดสอบโดย Bartlett's chi-square test หรือการทดสอบโดย Cattell's scree test แต่หากจะใช้วิธี PA เป็นทางเลือกในการพิจารณาจำนวนองค์ประกอบนั้น ก็ยังไม่มีคำสั่งให้เลือกใช้ในโปรแกรมสำเร็จรูปทางสถิติ โดยมีเพียงโปรแกรม ViSta-PARAN ที่พัฒนาโดย Young (2003) ซึ่งเพิ่มคำสั่งขั้นตอนวิธี PA โดย Ledesma and Valero-Mora (2007) ดังเช่น อาฟีฟิ ลาเต๊ะ (2551) ได้เสนอไว้ อย่างไรก็ตาม บทความนี้ได้นำเสนอการใช้โปรแกรมสำเร็จรูปทางสถิติอีก 2 โปรแกรมเพื่อเป็นอีกหนึ่งทางเลือกในการ นำวิธี PA มาใช้ในการพิจารณาองค์ประกอบที่เหมาะสมต่อไป

โปรแกรมสำเร็จรูปทางสถิติสำหรับวิธีการ Parallel Analysis

โปรแกรมสำเร็จรูป SPSS

โปรแกรมสำเร็จรูป SPSS (Statistics Package for the Social Sciences) เป็นโปรแกรมคอมพิวเตอร์ที่นิยมใช้แพร่หลายมากที่สุดสำหรับการวิเคราะห์ทางสถิติในสังคมศาสตร์ ไม่ว่าจะเป็นการวิเคราะห์สถิติเชิงบรรยาย หรือสถิติเชิงอ้างอิง ทั้งกรณีการวิเคราะห์แบบตัวแปรเดียว หรือหลายตัวแปร อย่างไรก็ตาม โปรแกรม SPSS ยังไม่มีคำสั่งโดยตรงเพื่อใช้ขั้นตอนวิธีจากวิธี PA ในการพิจารณาองค์ประกอบที่เหมาะสม ในปี ค.ศ. 2004 Hayton et al. ได้เสนอ Syntax Command ในการจำลองข้อมูล และคำนวณค่าไอเกน เพียงแค่เปิดหน้าต่างใหม่ในส่วน Syntax และป้อนคำสั่งดังภาพที่ 2

บรรทัดที่ 1 เป็นการนำเข้าข้อมูลโดยการกำหนดจำนวนค่าสังเกต (n) และจำนวนตัวแปร (v) ที่สอดคล้องกับข้อมูลจริงเพื่อทำการจำลองข้อมูล โดยบรรทัดที่ 2-3 จาก Syntax Command เป็นการกำหนดจำนวนค่าสังเกตทั้งหมด 100 ค่า (1 ถึง 100) และจำนวนตัวแปร 9 ตัว (v1 ถึง v9) บรรทัดที่ 4 เป็นการจำลองข้อมูลให้สอดคล้องกับข้อมูลจริงที่มีการแจกแจงแบบปกติ ซึ่งหากงานวิจัยที่นำมาศึกษาเป็นมาตราส่วนประมาณค่าจำนวน 7 ระดับ สามารถแทนค่าสูงสุดและค่ากลางของข้อมูล จาก 5 และ 3 เป็น 7 และ 4 ได้ ตามลำดับ ส่วนบรรทัดที่ 5 และ 6 เป็นค่าต่ำสุด และค่าสูงสุดของข้อมูลที่ใช้ในการจำลอง บรรทัดที่ 12-20 จาก Syntax Command เป็นขั้นตอนการคำนวณค่าไอเกนจากเมทริกซ์สหสัมพันธ์ของข้อมูลจำลองแบบ ซึ่งในขั้นตอนนี้ได้กำหนดเกณฑ์ค่าเริ่มต้นของค่าไอเกนเท่ากับ 0 และทำซ้ำ 25 ซ้ำ เพื่อให้ได้ค่าไอเกนในแต่ละตัวแปรที่ได้จำลอง โดยมีการสกัดองค์ประกอบด้วยวิธีการวิเคราะห์องค์ประกอบหลัก และไม่มี การหมุนแกนขององค์ประกอบ

เมื่อใช้คำสั่ง Run และเลือก All แล้วผลลัพธ์แสดงค่าสังเกต 100 ค่าจากตัวแปรทั้ง 9 ตัว ค่า Communalities และความแปรปรวนรวมที่สามารถอธิบายได้ (Total Variance Explained) ทั้งส่วนเริ่มต้น (Initial) และส่วนสกัด (Extraction) รวมทั้งเมทริกซ์องค์ประกอบในแต่ละองค์ประกอบที่ได้ (Component Matrix) จากขั้นตอนข้างต้น นำค่า

ไอเกนที่ได้เก็บไว้และทำซ้ำขั้นตอนดังกล่าวตามจำนวนซ้ำที่ต้องการ จากนั้นนำค่าไอเกนในแต่ละซ้ำมาคำนวณค่าเฉลี่ยของค่าไอเกน และค่าไอเกนตำแหน่งเปอร์เซ็นต์ไทล์ต่างๆ ในแต่ละลำดับขององค์ประกอบที่สัมพันธ์กันเพื่อนำไปเปรียบเทียบกับค่าไอเกนของข้อมูลจริงตามแนวทางของวิธี PA ต่อไป

```

1)      INPUT PROGRAM.
        LOOP LOOP#1=1 TO 100.
        DO REPEAT V=V1 TO V9.
        COMPUTE V=RND(NORMAL(5/6)+3).
5)      IF (V LT 1)V=1.
        IF (V GT 5)V=5.
        END REPEAT.
        END CASE.
        END LOOP#1.
10)     END FILE.
        END INPUT PROGRAM.
        FACTOR
        /VARIABLES v1 v2 v3 v4 v5 v6 v7 v8 v9
        /MISSING LISTWISE
15)     /ANALYSIS v1 v2 v3 v4 v5 v6 v7 v8 v9
        /PRINT INITIAL EXTRACTION
        /CRITERIA MINEIGEN(0) ITERATE(25)
        /EXTRACTION PC
    
```

ภาพที่ 2 Syntax Command เพื่อจำลองข้อมูลและคำนวณค่าไอเกนด้วยโปรแกรม SPSS

หน่วยประมวล WPA

Patil et al. (2007) ได้นำขั้นตอนวิธีจากโปรแกรมสำเร็จรูป SAS (Statistical Analysis System) ซึ่งเป็นโปรแกรมที่ใช้ในการวิเคราะห์ข้อมูลที่มีประสิทธิภาพสูงโปรแกรมหนึ่ง และสามารถใช้ได้กับข้อมูลในทุกประเภท ทั้งข้อมูลทางด้านวิทยาศาสตร์ หรือทางสังคมศาสตร์ โดย Patil et al. (2007) ได้นำขั้นตอนวิธีที่เสนอโดย O'Conner (2000) มาพัฒนาเป็นหน่วยประมวลผลเพื่อคำนวณค่าเฉลี่ยของค่าไอเกน และค่าไอเกนตำแหน่งเปอร์เซ็นต์ไทล์ที่ต้องการ ซึ่งมีชื่อเรียกว่า Web-based Parallel Analysis (WPA) เพียงแค่ป้อนจำนวนตัวแปร (Number of Variables) และจำนวนค่าสังเกต (Sample Size) ลงไปในช่องที่กำหนด เพื่อนำผลที่ได้ไปเปรียบเทียบกับค่าไอเกนของข้อมูลจริงในแต่ละลำดับขององค์ประกอบที่สัมพันธ์กัน โดยสามารถเปลี่ยนจำนวนซ้ำจาก 100 ซ้ำ ในช่อง Number of Random Correlation Matrices to Generate และตำแหน่งเปอร์เซ็นต์ไทล์จาก 95 ในช่อง Percentile of Eigenvalues ซึ่งเป็นค่าที่กำหนดจากหน่วยประมวลเป็นค่าอื่นที่ต้องการได้ รวมทั้งสามารถเลือกการกำหนดเมทริกซ์เริ่มต้นในช่อง Type of Analysis

ในการจำลองเมทริกซ์สหสัมพันธ์เป็นแบบ Principal Components Analysis หรือ Principal Axis Factoring และเลือกค่าเริ่มต้น (Seed) ได้ด้วย ผลที่ได้จะแสดงค่าเฉลี่ยของค่าไอเกน และค่าไอเกนตำแหน่งเปอร์เซ็นต์ไทล์ที่กำหนด ในแต่ละลำดับขององค์ประกอบที่สัมพันธ์กัน เพื่อนำไปเปรียบเทียบกับค่าไอเกนของข้อมูลจริงตามแนวทางของวิธี PA ต่อไป ดังภาพที่ 3

Please Enter Your Specifications:

Number of Variables in your Dataset to be Factor Analyzed		
Sample Size of Your dataset		
Type of Analysis	1	(default is '1' [RECOMMENDED] for Principal components Analysis; Use '2' for Principal Axis Factoring)
Number of Random Correlation Matrices to Generate	100	(default is 100)
Percentile of Eigenvalues	95	(default is 95)
Seed	1000	(please modify as desired)
Submit Query		Reset

Parallel Analysis: Eigenvalues Extracted Using Principal Components Analysis		
Specifications for this Run:		
Sample Size	100	
Number of Variables	9	
Number of Correlation Matrices	100	
Percent	95	
Seed	1000	
Random Data Eigenvalues		
Root	Means	Percentile
1.000000	1.478776	1.596786
2.000000	1.314124	1.419372
3.000000	1.185228	1.273411
4.000000	1.084153	1.139899
5.000000	0.985399	1.042112
6.000000	0.882676	0.956883
7.000000	0.791295	0.870297
8.000000	0.690452	0.772990
9.000000	0.587897	0.669331

ภาพที่ 3 การประมวลผล และผลลัพธ์จาก Web-based Parallel Analysis

สรุปและข้อเสนอแนะ

การพิจารณาจำนวนองค์ประกอบที่เหมาะสมจากค่าไอเกนด้วยวิธี PA ด้วยโปรแกรมสำเร็จรูปทั้ง 2 โปรแกรม ทั้งการใช้โปรแกรมสำเร็จรูป SPSS และหน่วยประมวลผล WPA ทำให้สามารถพิจารณาความมีนัยสำคัญของจำนวนองค์ประกอบว่า ควรประกอบไปด้วยกี่องค์ประกอบได้อย่างถูกต้องและชัดเจน ซึ่งให้ผลที่เหมาะสมกว่าวิธีอื่นๆ จึงน่าจะเป็นอีกทางเลือกหนึ่งในการนำวิธี PA ไปใช้ในการพิจารณาองค์ประกอบที่เหมาะสมต่อไป อย่างไรก็ตาม การอธิบายถึงลักษณะขององค์ประกอบจากกลุ่มตัวแปรก็ยังเป็นสิ่งสำคัญในงานวิจัยที่มีการวิเคราะห์ตัวแปรหลาย ๆ ตัว เพื่อค้นหาลักษณะความสัมพันธ์ของตัวแปรเหล่านั้น หรือเพื่อลดจำนวนข้อมูลให้น้อยลง โดยสูญเสียสารสนเทศที่มีอยู่ในข้อมูลน้อยที่สุด

เอกสารอ้างอิง

- อาฟีฟี ลาเต๊ะ. (2551). การพิจารณาจำนวนองค์ประกอบที่เหมาะสมโดยใช้วิธีการ Parallel Analysis และโปรแกรมคอมพิวเตอร์ ViSta-PARAN. *วารสารวิจัยและวัดผลการศึกษา มหาวิทยาลัยบูรพา*, 6, 1-14.
- Bartlett, M.S. (1950). Tests of significance in factor analysis. *British Journal of Psychology, Statistician Section*, 3, 77-85.
- Dinno, A. (2009). Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research*, 44, 362-388.
- Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299.
- Glorfeld, L.W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55, 377-393.
- Hayton, J.C. Allen, D.G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7, 191-205.
- Horn, J.L. (1965). A rational and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Humphreys, L.G., & Montanelli, R.G. (1975). An investigated of the parallel analysis criterion for determining the number of common Factors. *Multivariate Behavioral Research*, 10, 193-205.
- Jackson, D.A. (1993). Stopping rules in principal component analysis: A comparison of heuristical and statistical approaches. *Ecology*, 74, 2204-2214.
- Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Lawley, D.N. (1956). Tests of the significance for the latent roots of covariance and correlation matrices. *Biometrika*, 43, 128-136.
- Ledesma, P.D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research and Evaluation*, 12, 1-11.
- O'conner, B.P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavioral Research Methods, Instruments, and Computers*, 32, 396-402.
- Patil, V., Singh, S., Mishra, S., & Todd, D.D. (2007). Efficient theory development and factor retention criteria: Abandon the 'Eigenvalue Greater than One' criterion. *Journal of Business Research*, 61, 162-170.
- Peres-Neto, P.R., Jackson, D.A., & Somers, K.M. (2003). Giving meaningful interpretation to ordination axes: Assessing loading significance in principal component analysis. *Ecology*, 84, 2347-2363.

- Peres-Neto, P.R., Jackson, D.A., & Somers, K.M. (2005). How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics and Data Analysis*, 49, 974-997.
- Silverstein, A.B. (1987). Note on the parallel analysis criterion for determining the number of common factors or principal components. *Psychological Reports*, 61, 351-354.
- Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321-327.
- Zwick, W.R., & Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432-442.