บทที่ 2

สถิติที่ใช้ในการวิจัย

2.1 การวิเคราะห์ปัจจัย

ในขั้นตอนการสกัดปัจจัย (Factor Extraction) จะใช้วิธีองค์ประกอบหลัก (Principal Components) ซึ่งอาศัยหลักความสัมพันธ์เชิงเส้นระหว่างตัวแปรที่ใช้เป็นข้อมูล มีวิธีการดังต่อไปนี้

1. กำหนดปัจจัยที่คาดว่าจะมีผลหรืออิทธิพลต่อตัวแปรหรือลักษณะที่เราสนใจศึกษา (x_i)
2. ตรวจสอบปัจจัยแต่ละตัวเกี่ยวกับลักษณะต่าง ๆ
   - 2.1 เป็นตัวแปรเชิงปริมาณชนิดต่อเนื่องหรือไม่
   - 2.2 หน่วยของตัวแปรเป็นหน่วยเดียวกันหรือไม่
   - 2.3 ค่าของตัวแปรที่นำมาใช้เปรียบเทียบกันได้หรือไม่
3. หา Dispersion Matrix ของปัจจัยข้างต้น (S)
4. หา Characteristic root ของ S คือ $\lambda_1,\ldots,\lambda_p$
5. นำ Characteristic root ของ S ที่มีค่าสูงสุด (λ_1) มาหา Characteristic vector เรียกว่า $\underline{a}_1$
6. Normalized Characteristic vector $\underline{a}_1$ เรียกว่า $\underline{a}_1^*$ จะได้ Component ที่ 1 อยู่ในรูป

$$Y_1 = \underline{a}_1^* \underline{X} = \sum_{j=1}^{p} a_{ij} X_j$$

ซึ่งใช้สำหรับวัด "ขนาด" หรือ " SIZE" ของตัวแปรหรือลักษณะที่สนใจ

7. หา Component อื่น ๆ ที่เหลือจาก Characteristic root อื่น ๆ ที่มีค่าสูงรองลงมาจาก λ_1 ตามลำดับ คือ $\lambda_2, \lambda_3, \ldots, \lambda_p$ โดยวิธีเดียวกัน เพื่อนำมาใช้ในการทดสอบสมมติฐานเกี่ยวกับปัจจัยต่าง ๆ (Orthogonal Comparison)

ในกรณีที่นำตัวแปรมาวิเคราะห์ p ตัว และใช้ตัวอย่างขนาด n

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots\cdots & X_{1p} \\ X_{21} & X_{22} & \cdots\cdots & X_{2p} \\ \vdots & & & \\ \vdots & & & \\ X_{n1} & X_{n2} & \cdots\cdots & X_{np} \end{bmatrix}_{n\times p}$$

ให้ Y เป็นตัวแปรตัวใหม่ที่จะใช้แทนปัจจัยหรือตัวแปรที่นำมาวิเคราะห์ทั้ง p ตัว $(X_1, X_2, \ldots\ldots, X_p)$

$$Y = \underset{1\times p}{\underline{a}'}\ \underset{1\times p}{\underline{X}} \qquad \underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ : \\ a_p \end{bmatrix}$$

เพื่อที่จะสามารถหาค่า $\underline{a}$ ที่ประกอบด้วยค่า $a_1, a_2, \ldots, a_p$ ได้ จึงต้องมีสมการแสดงความสัมพันธ์ระหว่าง Y กับ X ทั้งหมด p สมการ

ให้ $\underline{Y_i} = \underline{a_i'}\underline{X}$ เมื่อ $\underline{a_i} = \begin{bmatrix} a_{i1} \\ a_{i2} \\ : \\ a_{ip} \end{bmatrix}_{p\times 1}$

การหาค่า vector $\underline{a_i}$

ให้ S แทน Dispersion matrix ของ $\underline{X}$

$$S = \begin{bmatrix} s_{11} & s_{12} \ldots\ldots\ldots s_{1p} \\ s_{21} & s_{22} \ldots\ldots\ldots s_{2p} \\ : & \\ : & \\ s_{p1} & s_{p2} \ldots\ldots\ldots s_{pp} \end{bmatrix}_{p\times p}$$

Characteristic root ของ S คือ $\lambda_1 > \lambda_2 > \lambda_3 > \ldots\ldots > \lambda_p$ หาจาก $|s - \lambda I| = 0$

Characteristic vector $\underline{a_i}$ หาจาก $(s - \lambda_i I)\underline{a_i} = \underline{0}$

Normalized Characteristic vector $\underline{a_i}$ หาจาก

$$\begin{bmatrix} a_{i1} / \sqrt{a_{i1}^2 + a_{i2}^2 + \ldots + a_{ip}^2} \\ a_{i2} / \sqrt{a_{i1}^2 + a_{i2}^2 + \ldots + a_{ip}^2} \\ : \\ : \\ a_{ip} / \sqrt{a_{i1}^2 + a_{i2}^2 + \ldots + a_{ip}^2} \end{bmatrix}$$

2.2 การจัดกลุ่มของค่าถ่วงน้ำหนัก โดยวิธี Cumulative$\sqrt{f(y)}$

การหาขอบเขตของกลุ่ม (Construction of Strata) เพื่อแบ่งค่าถ่วงน้ำหนักออกเป็นกลุ่มๆ โดยพิจารณาจากการสะสมของ $\sqrt{f(y)}$ (Cumulative$\sqrt{f(y)}$) เมื่อ f(y) เป็นจำนวนรายที่มีคะแนนความสำคัญอยู่ในช่วงหนึ่งๆ โดยที่ y คือ ค่าที่ได้จากการแทนตัวแปรลงแบบจำลอง ที่แสดงถึงค่าถ่วงน้ำหนักของปัจจัยที่ใช้ในการจำแนกกลุ่ม ขั้นตอนในการจำแนกกลุ่มมีดังนี้

1) แบ่งคะแนนความสำคัญออกเป็นช่วง โดยให้แต่ละช่วงมีขนาดพอสมควร

2) หาความถี่ (frequency) ของคะแนนความสำคัญในแต่ละช่วง คือ f(y) และ $\sqrt{f(y)}$

3) หาค่าสะสมของ $\sqrt{f(y)}$ (Cumulative$\sqrt{f(y)}$)

4) แบ่งยอดรวมของค่าสะสม $\sqrt{f(y)}$ โดยหารค่าจำนวนกลุ่มที่ต้องการผลหารที่ได้เป็นตัวแบ่งค่าสะสมของ $\sqrt{f(y)}$ ออกเป็นช่วงซึ่งแสดงถึงขอบเขตของค่าถ่วงน้ำหนัก