บทที่ 2

ตัวสถิติและผลงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ผู้วิจัยมุ่งที่จะศึกษาถึงวิธีการประมาณพารามิเตอร์ในการวิเคราะห์ความแปรปรวนร่วมระหว่างวิธีตัวประมาณเอ็ม และวิธีบูตสแตรป ว่าวิธีใดให้ประสิทธิภาพดีกว่าโดยการเปรียบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบ การศึกษาจะศึกษาจากแบบจำลองค่าสังเกตในตัวแบบ ซึ่งการวิเคราะห์ความแปรปรวนร่วมเป็นวิธีที่ผสมผสานระหว่างการวิเคราะห์ความแปรปรวนกับการวิเคราะห์ความถดถอยเข้าด้วยกัน

สถิติที่ใช้ในการวิจัย

ตัวแบบของการวิเคราะห์ความแปรปรวน

(1) $$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad , \quad i = 1, 2, 3, \ldots, n_j \quad j = 1, 2, 3, \ldots, p$$

ตัวแบบของการวิเคราะห์ความถดถอย เมื่อตัวแปรร่วม X มีความสัมพันธ์เชิงเส้นกับตัวแปรตาม Y คือ

(2) $$y_{ij} = \mu + \sum_{k=1}^{q} \beta_k (x_{ijk} - \bar{x}_{..k}) + \varepsilon_{ij}$$

ผสมผสานตัวแบบการวิเคราะห์ความแปรปรวนจาก (1) กับตัวแบบของการวิเคราะห์ความถดถอย (2) จะได้ตัวแบบการวิเคราะห์ความแปรปรวนร่วม

$$y_{ij} = \mu + \tau_i + \sum_{k=1}^{q} \beta_k (x_{ijk} - \bar{x}_{..k}) + \varepsilon_{ij} \quad , \quad i = 1, 2, 3, \ldots, n_j \quad j = 1, 2, 3, \ldots, p$$

จากตัวแบบกำหนดให้

p เป็นจำนวนวิธีปฏิบัติ

μ หมายถึง ค่าเฉลี่ยของ y

τ_i หมายถึง อิทธิพลของวิธีปฏิบัติ

q เป็นจำนวนตัวแปรร่วม

β_k หมายถึง ค่าสัมประสิทธิ์ของการถดถอยของตัวแปรร่วมที่ k

$\bar{x}_{..k}$ หมายถึง ค่าเฉลี่ยของตัวแปรร่วมที่ k

n_j เป็นขนาดตัวอย่างในวิธีปฏิบัติที่ k

ε_{ij} หมายถึง ค่าความคลาดเคลื่อนของหน่วยทดลองที่ j วิธีปฏิบัติที่ i

การวิเคราะห์ความแปรปรวนร่วมต้องการทดสอบว่าวิธีปฏิบัติมีอิทธิพลแตกต่างกันหรือไม่ จะใช้ตัวแบบดังนี้

ตัวแบบเต็มรูป (Full Model)

$$y_{ij} = \mu + \tau_i + \sum_{k=1}^{q} \beta_k (x_{ijk} - \bar{x}_{..k}) + \varepsilon_{ij}$$

ตัวแบบลดรูป (Reduced Model)

$$y_{ij} = \mu + \sum_{k=1}^{q} \beta_k (x_{ijk} - \bar{x}_{..k}) + \varepsilon_{ij}$$

สมมติฐานการทดสอบ

$H_0 : \tau_1 = \tau_2 = \tau_3 = \ldots = \tau_p = 0$ หรืออิทธิพลของวิธีปฏิบัติไม่แตกต่างกัน

H_a : อิทธิพลของวิธีปฏิบัติอย่างน้อย 1 วิธีที่แตกต่างกัน

ตัวสถิติทดสอบ

$$F = \frac{(SSF - SSR)/(p-1)}{(SST - SSF)/(n-p-q)}$$

เมื่อ SSF หมายถึง ผลบวกกำลังสองของการประมาณของตัวแบบเต็มรูป

SSR หมายถึง ผลบวกกำลังสองของการประมาณของตัวแบบลดรูป

SST หมายถึง ผลบวกกำลังสองรวม

n หมายถึง ขนาดตัวอย่างทั้งหมดซึ่งเท่ากับ $\sum_{j=1}^{p} n_j$

จะปฏิเสธสมมติฐานว่าง (H_0) เมื่อค่าสถิติที่คำนวณได้ $F_e > F_\alpha(p-1, n-p-q)$
ในการทดสอบอิทธิพลของวิธีปฏิบัติจะต้องสร้างตัวแปรหุ่น (Dummy) เพื่อแสดงอิทธิพลของวิธีปฏิบัติรวมไว้ในสมการก่อน ดังนี้

จากตัวแบบเต็มรูป

$$y_{ij} = \mu + \tau_i + \sum_{k=1}^{q} \beta_k (x_{ijk} - \bar{x}_{..k}) + \varepsilon_{ij}$$

$$D_{i1} = \begin{cases} 1 & \text{เป็นวิธีปฏิบัติที่ 1} \\ 0 & \text{เป็นวิธีปฏิบัติอื่นๆ} \end{cases}$$

$$D_{i2} = \begin{cases} 1 & \text{เป็นวิธีปฏิบัติที่ 2} \\ 0 & \text{เป็นวิธีปฏิบัติอื่นๆ} \end{cases}$$

.

.

.

$$D_{i\,p-1} = \begin{cases} 1 & \text{เป็นวิธีปฏิบัติที่ p-1} \\ 0 & \text{เป็นวิธีปฏิบัติอื่นๆ} \end{cases}$$

ใส่ตัวแปรหุ่นที่สร้างลงในสมการจะได้ตัวแบบเต็มรูปใหม่ดังนี้

$$y_{ij} = \alpha_0 + \alpha_1 D_{i1} + \alpha_1 D_{i2} + \ldots + \alpha_{p-1_1} D_{ip-11} + \sum_{k=1}^{q} \beta_k (x_{ijk} - \bar{x}_{..k}) + \varepsilon_{ij}$$

ให้ $(\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_p)$ เป็นสัมประสิทธิ์ความถดถอยของตัวแปรหุ่น (Dummy) ซึ่งจะแสดงอิทธิ-พลของวิธีปฏิบัติและสามารถเขียนในรูปของเมตริกซ์ได้ดังนี้

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

เมื่อ $\underset{\sim}{y}$ เป็นเวกเตอร์ของค่าสังเกตหรือตัวแปรตามซึ่งมีขนาด n x 1

X เป็นเมตริกซ์ของตัวแปรอิสระซึ่งมีขนาด n x (p+q)

$\underset{\sim}{\beta}$ เป็นเวกเตอร์ของสัมประสิทธิ์ความถดถอยซึ่งมีขนาด (p+q) x 1

$\underset{\sim}{\varepsilon}$ เป็นเวกเตอร์ของค่าความคลาดเคลื่อนที่มีขนาด n x 1

$$\underset{\sim}{y} = \begin{bmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{y}_p \end{bmatrix}$$

$$X = \begin{bmatrix} \underset{\sim}{1} & \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{0} & \ldots & (x_{11} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{12} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{1q} - \bar{x}_{..q}\underset{\sim}{1}) \\ \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{1} & \underset{\sim}{0} & \ldots & (x_{21} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{22} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{2q} - \bar{x}_{..q}\underset{\sim}{1}) \\ \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{1} & \ldots & (x_{31} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{32} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{3q} - \bar{x}_{..q}\underset{\sim}{1}) \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \ldots & (x_{p1} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{p2} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{pq} - \bar{x}_{..q}\underset{\sim}{1}) \end{bmatrix}$$

$$\underset{\sim}{\beta} = (\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{p-1}, \beta_1, \beta_2, \ldots, \beta_q)'$$

$$\underset{\sim}{\varepsilon} = \begin{bmatrix} \underset{\sim}{\varepsilon_1} \\ \underset{\sim}{\varepsilon_2} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{\varepsilon_{p1}} \end{bmatrix}$$

ตัวแบบลดรูป

$$y_{ij} = \alpha_0 + \sum_{k=1}^{q} \beta_k (x_{ijk} - \bar{x}_{..k}) + \varepsilon_{ij}$$

เขียนอยู่ในรูปเมตริกซ์ได้ดังนี้

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

เมื่อ $\underset{\sim}{y}$ เป็นเวกเตอร์ของค่าสังเกตหรือตัวแปรตามซึ่งมีขนาด n x 1

X เป็นเมตริกซ์ของตัวแปรอิสระซึ่งมีขนาด n x (q+1)

$\underset{\sim}{\beta}$ เป็นเวกเตอร์ของสัมประสิทธิ์ความถดถอยซึ่งมีขนาด (q+1) x 1

$\underset{\sim}{\varepsilon}$ เป็นเวกเตอร์ของค่าความคลาดเคลื่อนที่มีขนาด n x 1

$$X = \begin{bmatrix} \underset{\sim}{1} & (x_{11} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{12} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{1q} - \bar{x}_{..q}\underset{\sim}{1}) \\ \underset{\sim}{1} & (x_{21} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{22} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{2q} - \bar{x}_{..q}\underset{\sim}{1}) \\ \underset{\sim}{1} & (x_{31} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{32} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{3q} - \bar{x}_{..q}\underset{\sim}{1}) \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{1} & (x_{p1} - \bar{x}_{..1}\underset{\sim}{1}) & (x_{p2} - \bar{x}_{..2}\underset{\sim}{1}) & \ldots & (x_{pq} - \bar{x}_{..q}\underset{\sim}{1}) \end{bmatrix}$$

$$\underset{\sim}{\beta} = (\alpha_0, \beta_1, \beta_2, \ldots, \beta_q)'$$

$$\underset{\sim}{\varepsilon} = \begin{bmatrix} \underset{\sim}{\varepsilon_1} \\ \underset{\sim}{\varepsilon_2} \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{\varepsilon_{p1}} \end{bmatrix}$$

วิธีการประมาณพารามิเตอร์

1. <u>วิธีบูตสแตรป</u> **(Bootstrap Method)**

การหาตัวประมาณพารามิเตอร์โดยวิธีบูตสแตรป ซึ่งสถานการณ์ต่างๆ ที่ใช้ในการวิจัยนั้นถูกสร้างมาโดยอาศัยการจำลองโดยใช้เทคนิคการจำลองมอนติคาร์โล **(Monte - Carlo Simulation Technique)** โดยอาศัยเครื่องคอมพิวเตอร์ AMDAHL 5860 ใช้ภาษาฟอร์แทรน **(Fortran)** มีรายละเอียดดังนี้

1.1 สร้างตัวเลขสุ่มตามประเภทการแจกแจง

1.2 สุ่มตัวอย่างจากกลุ่มตัวเลขสุ่มที่สร้างขึ้นมาแบบใส่คืน

1.3 คำนวณหาค่า $\underset{\sim}{\hat{\beta}}$ โดยวิธีกำลังสองน้อยที่สุด มีขั้นตอนดังนี้

จากตัวแบบ $$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \qquad (1)$$

$\underset{\sim}{\varepsilon}$ มีการแจกแจงแบบปกติ $E(\underset{\cdot}{\varepsilon}) = \underset{\cdot}{0}$ และ $V(\underset{\cdot}{\varepsilon}) = E(\underset{\cdot}{\varepsilon}'\underset{\cdot}{\varepsilon}) = \sigma^2 I_n$

วิธีกำลังสองน้อยที่สุดคือ ทำให้ผลบวกกำลังสองของความคลาดเคลื่อนมีค่าน้อยที่สุด

$$\begin{aligned} SSE &= \underset{\sim}{\hat{\varepsilon}}' \underset{\sim}{\hat{\varepsilon}} \\ &= (\underset{\sim}{y} - X\underset{\sim}{\hat{\beta}})'(\underset{\sim}{y} - X\underset{\sim}{\hat{\beta}}) \\ &= (\underset{\sim}{y}' - \underset{\sim}{\hat{\beta}}' X')(\underset{\sim}{y} - X\underset{\sim}{\hat{\beta}}) \end{aligned}$$

$$= \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{y}'X\hat{\underset{\sim}{\beta}} - \hat{\underset{\sim}{\beta}}'X'\underset{\sim}{y} + \hat{\underset{\sim}{\beta}}'X'X\hat{\underset{\sim}{\beta}}$$

$$= \underset{\sim}{y}'\underset{\sim}{y} - 2\hat{\underset{\sim}{\beta}}'X'\underset{\sim}{y} + \hat{\underset{\sim}{\beta}}'X'X\hat{\underset{\sim}{\beta}}$$

การหาค่าน้อยที่สุดของผลบวกกำลังสองของความคลาดเคลื่อนทำได้โดยหาอนุพันธ์ (Differentiate) เทียบกับ $\hat{\underset{\sim}{\beta}}_i$ แล้วกำหนดให้เท่ากับ 0

$$\frac{\partial\, SSE}{\partial \hat{\underset{\sim}{\beta}}} = \begin{bmatrix} \frac{\partial\, SSE}{\partial \hat{\underset{\sim}{\beta}}_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial\, SSE}{\partial \hat{\underset{\sim}{\beta}}_p} \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = \underset{\sim}{0}$$

$$-2X'\underset{\sim}{y} + 2X'X\hat{\underset{\sim}{\beta}} = \underset{\sim}{0}$$

$$(X'X)\hat{\underset{\sim}{\beta}} = X'\underset{\sim}{y}$$

$$\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

แทนค่า $\hat{\underset{\sim}{\beta}}$ ลงในสมการจะได้

$$\hat{\underset{\sim}{y}} = X\hat{\underset{\sim}{\beta}} \qquad (2)$$

จาก (1) และ (2) คำนวณหา

$$\hat{\underset{\sim}{\varepsilon}} = \underset{\sim}{y} - \hat{\underset{\sim}{y}}$$

นั่นคือ $\hat{\underset{\sim}{\varepsilon}}_i = \underset{\sim}{y}_i - \hat{\underset{\sim}{y}}_i$, $i = 1, 2, \ldots, n$

1.4 สุ่ม $\hat{\varepsilon}_i$ แบบใส่คืน (With Replacement) ขนาด n จะได้ผลลัพธ์ดังนี้

$$\varepsilon^*_1, \varepsilon^*_2, \varepsilon^*_3, \ldots, \varepsilon^*_n$$

1.5 นำค่า ε^*_i มาพิจารณาโดยรวมไว้ในสมการ

$$\underset{\sim}{y}^{*} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}^{*}$$
$$= \underset{\sim}{\hat{y}} + \underset{\sim}{\varepsilon}^{*}$$

1.6 คำนวณหาค่า $\underset{\sim}{\hat{\beta}}^{*}$ โดยวิธีกำลังสองน้อยที่สุด โดยยึดหลักเกณฑ์เดียวกันกับการหา $\underset{\sim}{\hat{\beta}}$ คือทำการหา $\underset{\sim}{\hat{\beta}}^{*}$ ซึ่งเป็นตัวประมาณกำลังสองน้อยที่สุดของ $\underset{\sim}{\hat{\beta}}$ ที่ทำให้

$$SSE^{*} = \sum_{i=1}^{n} \hat{\varepsilon}_i^{*2} = \underset{\sim}{\hat{\varepsilon}}^{*\prime} \underset{\sim}{\hat{\varepsilon}}^{*} \quad \text{มีค่าต่ำสุด}$$

โดยการหาอนุพันธ์ของ SSE^{*} เทียบกับ $\underset{\sim}{\hat{\beta}}^{*}$ แล้วกำหนดให้เท่ากับ 0 นั่นคือ

$$\begin{aligned} SSE^{*} &= \underset{\sim}{\hat{\varepsilon}}^{*\prime} \underset{\sim}{\hat{\varepsilon}}^{*} \\ &= (\underset{\sim}{y}^{*} - X\underset{\sim}{\hat{\beta}}^{*})'(\underset{\sim}{y}^{*} - X\underset{\sim}{\hat{\beta}}^{*}) \\ &= (\underset{\sim}{y}^{*\prime} - \underset{\sim}{\hat{\beta}}^{*\prime} X')(\underset{\sim}{y}^{*} - X\underset{\sim}{\hat{\beta}}^{*}) \\ &= \underset{\sim}{y}^{*\prime}\underset{\sim}{y}^{*} - \underset{\sim}{y}^{*\prime} X\underset{\sim}{\hat{\beta}}^{*} - \underset{\sim}{\hat{\beta}}^{*\prime} X'\underset{\sim}{y}^{*} + \underset{\sim}{\hat{\beta}}^{*\prime} X'X\underset{\sim}{\hat{\beta}}^{*} \\ &= \underset{\sim}{y}^{*\prime}\underset{\sim}{y}^{*} - 2\underset{\sim}{\hat{\beta}}^{*\prime} X'\underset{\sim}{y}^{*} + \underset{\sim}{\hat{\beta}}^{*\prime} X'X\underset{\sim}{\hat{\beta}}^{*} \end{aligned}$$

$$\frac{\partial\, SSE^{*}}{\partial \underset{\sim}{\hat{\beta}}^{*}} = \begin{bmatrix} \dfrac{\partial\, SSE}{\partial \underset{\sim}{\hat{\beta}}^{*}{}_{1}} \\ . \\ . \\ . \\ \dfrac{\partial\, SSE}{\partial \underset{\sim}{\hat{\beta}}^{*}{}_{p}} \end{bmatrix} = \begin{bmatrix} 0 \\ . \\ . \\ . \\ . \\ . \\ . \\ 0 \end{bmatrix} = \underset{\sim}{0}$$

$$\begin{aligned} -2X'\underset{\sim}{y}^{*} + 2X'X\underset{\sim}{\hat{\beta}}^{*} &= \underset{\sim}{0} \\ (X'X)\underset{\sim}{\hat{\beta}}^{*} &= X'\underset{\sim}{y}^{*} \\ \underset{\sim}{\hat{\beta}}^{*} &= (X'X)^{-1}X'\underset{\sim}{y}^{*} \end{aligned}$$

1.7 กระทำตามขั้นตอนในข้อ 1.4-1.6 ซ้ำ k ครั้ง จะได้

$\hat{\underset{\sim}{\beta}}^{*1}, \hat{\underset{\sim}{\beta}}^{*2}, \ldots, \hat{\underset{\sim}{\beta}}^{*k}$

1.8 คำนวณหาค่า $\overline{\hat{\underset{\sim}{\beta}}} = \sum_{i=1}^{k} \hat{\underset{\sim}{\beta}}^{*i} \Big/ k$

จากการศึกษาพบว่าการคำนวณหาค่า $\overline{\hat{\underset{\sim}{\beta}}} = \sum_{i=1}^{k} \hat{\underset{\sim}{\beta}}^{*i} \Big/ k$ ซ้ำๆ กัน k ตัวนั้นพบว่าจำนวนที่เหมาะสมจะอยู่ในช่วง 50-200 ตัว ในที่นี้จะใช้ $k = 50$

2. <u>วิธีตัวประมาณเอ็ม</u>

ตัวประมาณเอ็ม ผู้เสนอคือ P.J. Huber ในปี ค.ศ. 1964 โดยศึกษามาจากฟังก์ชันความคลาดเคลื่อน ตัวประมาณที่จะประมาณพารามิเตอร์ $\sum_{i=1}^{n} \rho\,(\varepsilon_i/S)$ เมื่อ ε_i เป็นค่าคลาดเคลื่อนของค่าสังเกตที่ i และ S เป็นตัวประมาณของการกระจายของตัวอย่าง

จากสมการ

(1) $$\underset{\underset{\sim}{\beta}}{\mathrm{MIN}} \sum_{i=1}^{n} \rho\,(\varepsilon_i/S) = \underset{\underset{\sim}{\beta}}{\mathrm{MIN}} \sum_{i=1}^{n} \rho \left| (y_i - X_i{}' \underset{\sim}{\beta}) \Big/ S \right| \quad , \; X_i = \underset{\sim}{X_i}{}'$$

ในขั้นแรกจะต้องประมาณ $\hat{\underset{\sim}{\beta}}$ โดยวิธีกำลังสองน้อยที่สุดเสียก่อน แล้วจึงนำมาหาค่า ε_i จากนั้นนำค่า ε_i ที่ได้มาหาค่า S ดังนี้

$$S = \frac{\mathrm{Median}\left|\varepsilon_i - \mathrm{Median}(\varepsilon_i)\right|}{0.6745}$$

โดยปกติ S เป็นตัวประมาณที่เอนเอียง ถ้าจะไม่ให้เอนเอียงต้องใช้ค่ามัธยฐานของส่วนเบี่ยงเบนมาตรฐานและทำการปรับด้วยค่าคงที่ 0.6745

จากสมการ (1) หาอนุพันธ์ของ ρ เทียบกับ β_i แล้วกำหนดให้เท่ากับ 0 โดย $\Psi = \rho'$

(2) $$\sum_{i=1}^{n} X_{ij} \Psi \left[(y_i - \underset{\sim}{X}' \underset{\sim}{\beta}) \Big/ S \right] = 0 \quad , \; j = 1, 2, \ldots, M$$

โดยที่ $M = p+q+1$ ในตัวแปรเต็มรูป

$M = q$ ในตัวแปรลดรูป

การประมาณค่าสัมประสิทธิ์การถดถอยส่วนใหญ่จะใช้วิธีกำลังสองน้อยที่สุดซึ่งถ่วงน้ำหนักซ้ำหลายรอบ (Iteratively Reweighted Least Squares) จากสมการ (2) สามารถเขียนได้ดังนี้

$$\sum_{i=1}^{n} X_{ij}\Psi\left[(y_i - \underset{\sim}{X}'\underset{\sim}{\beta})\Big/S\right] = \frac{\sum_{i=1}^{n} X_{ij}\Psi\left[(y_i - \underset{\sim}{X}'\underset{\sim}{\beta})\Big/S\right]}{(y_i - \underset{\sim}{X}'\underset{\sim}{\beta})\Big/S} \times (y_i - \underset{\sim}{X}'\underset{\sim}{\beta})\Big/S = 0$$

กำหนดให้

$$W_{i0} = \begin{cases} \dfrac{\Psi_i\left|(y_i - \underset{\sim}{X}_i'\underset{\sim 0}{\beta})\Big/S\right|}{(y_i - \underset{\sim}{X}_i'\underset{\sim 0}{\beta})\Big/S} & , y_i \neq \underset{\sim}{X}_i'\underset{\sim 0}{\hat{\beta}} \\ 1 & , y_i = \underset{\sim}{X}_i'\underset{\sim 0}{\hat{\beta}} \end{cases}$$

ดังนั้นสามารถเขียนเป็น m สมการได้ดังนี้

$$\sum_{i=1}^{n} X_{ij} W_{i0}(y_i - \underset{\sim}{X}'\underset{\sim}{\beta}) = 0 \qquad , \quad j = 1, 2, \ldots, m$$

และเขียนให้อยู่ในรูปเมตริกซ์ได้เป็น

$$X' W_0' X \underset{\sim}{\hat{\beta}} = X' W_0 \underset{\sim}{y}$$

เมื่อ W เป็นเมตริกซ์ขนาด n x n และมีสมาชิกตามเส้นทแยงมุมเป็น $W_{10}, W_{20}, \ldots, W_{n0}$

$$W = \begin{bmatrix} W_{10} & 0 & . & . & . & . & . & 0 \\ 0 & W_{20} & . & . & . & . & . & 0 \\ . & . & & & & & & . \\ . & . & & & & & & . \\ . & . & & & & & & . \\ 0 & . & . & . & . & . & . & W_{n0} \end{bmatrix}$$

การประมาณที่แกร่งจำเป็นอย่างยิ่งต้องเลือกค่าเริ่มต้นของ β_j อย่างระมัดระวังซึ่งการใช้ $\hat{\beta}_j$ จากวิธีกำลังสองน้อยที่สุดจะให้ผลดี ในที่นี้จึงใช้ $\hat{\underset{\sim}{\beta}}$ กำลังสองน้อยที่สุดมาคำนวณค่า W_i แล้วก็นำไปหาค่า $\hat{\beta}_j$ ใหม่โดยกระทำซ้ำ ๆ จนกระทั่งได้ค่า $\hat{\underset{\sim}{\beta}}$ ที่ค่อนข้างคงที่ ดังนั้นตัวประมาณ $\hat{\underset{\sim}{\beta}}_M$ จึงหาได้จาก

$$\hat{\underset{\sim}{\beta}}_M = (X' W X)^{-1} X' W \underset{\sim}{y}$$

ผลงานวิจัยที่เกี่ยวข้อง

1. วิธีบูตสแตรป

มีการพัฒนามาจากการนำวิธีการของตัวแบบสมการถดถอยที่ไม่ทราบการแจกแจงของความคลาดเคลื่อนไปใช้ในการประมาณหาค่าส่วนเบี่ยงเบนมาตรฐานของสัมประสิทธิ์ของความถดถอย ผู้ศึกษาคือแบรดเลย์ เอฟรอน (BRADLEY EFRON) ในปีค.ศ. 1979 ต่อมาในปี ค.ศ. 1983 เดวิด เอ ฟรีเมนต์ (DAVID A. FREEMAN) และสตีเฟน ซี ปีเตอร์ (STEPHEN C. PETER) ได้นำเอาวิธีบูตสแตรปมาใช้ในการประมาณค่าส่วนเบี่ยงเบนมาตรฐานและการพยากรณ์ของการวิเคราะห์ความถดถอยเชิงเส้น เมื่อขนาดตัวอย่างมีจำนวนจำกัดและไม่เป็นไปตามข้อตกลงเบื้องต้น พบว่าวิธีบูตสแตรปสามารถประมาณค่าส่วนเบี่ยงเบนมาตรฐานของสัมประสิทธิ์ความถดถอยได้ค่าใกล้เคียงค่าจริงมากกว่าค่าประมาณที่หาได้จากสูตรทั่วไป

ปี พ.ศ. 2532 มาลี ตระการศิรินันท์ ได้เปรียบเทียบการประมาณค่าพารามิเตอร์ของรูปแบบสมการความถดถอยเชิงเส้นระหว่างวิธีกำลังสองน้อยที่สุด และวิธีบูตสแตรปพบว่าเมื่อความคลาดเคลื่อนมีการแจกแจงเบี่ยงเบนไปจากแบบปกติแต่ไม่มากนัก ขนาดตัวอย่างและจำนวนตัวแปรอิสระมีค่าน้อย วิธีบูตสแตรปมีประสิทธิภาพดีกว่าวิธีกำลังสองน้อยที่สุด

2. วิธีตัวประมาณเอ็ม

ได้มีผู้ศึกษาลักษณะการประมาณพารามิเตอร์ในรูปแบบของการวิเคราะห์การถดถอยเชิงเส้นได้คือ ปี ค.ศ. 1964 พี. เจ. ฮิวเบอร์ (P.J. Huber) ได้ศึกษากับฟังก์ชั่นของความคลาดเคลื่อน ซึ่งเรียกว่าตัวประมาณเอ็ม (M-Estimator) โดยที่ตัวประมาณนี้จะประมาณพารามิเตอร์ของ $\sum_{i=1}^{n} \rho\,(\varepsilon_i/S)$ เมื่อ ε_i เป็นค่าคลาดเคลื่อนของตัวสังเกตที่ i และ S เป็นตัวประมาณที่เหมาะสมของการกระจายของตัวอย่าง

ปี พ.ศ. 2530 ปราณี รัตนังได้ศึกษาเปรียบเทียบวิธีการประมาณสัมประสิทธิ์การถด-ถอยพหุ เมื่อความผิดพลาดมีการแจกแจงหางยาวกว่าปกติและเบ้ ระหว่างวิธีกำลังสองน้อยที่สุด และวิธีตัวประมาณเอ็ม ที่ใช้เกณฑ์ความแกร่งของ Ramsay พบว่าสเกลแฟคเตอร์และเปอร์เซนต์การปลอมปนจะมีอิทธิพลจากมากไปหาน้อยที่ทำให้ตัวประมาณเอ็มดีกว่าวิธีกำลังสองน้อยที่สุด

ปี พ.ศ. 2535 กรรณิการ์ อุโฆษกุล ได้ศึกษาเปรียบเทียบวิธีการประมาณพารามิเตอร์ในการวิเคราะห์ความแปรปรวนร่วมเมื่อความผิดพลาดมีการแจกแจงแบบหางยาว ระหว่างวิธีกำลังสองน้อยที่สุดและวิธีตัวประมาณเอ็มที่ใช้เกณฑ์ความแกร่งของ Ramsay พบว่าจำนวนตัวแปรร่วมและจำนวนวิธีปฏิบัติมีอิทธิพลที่ทำให้ตัวประมาณเอ็มดีกว่าวิธีกำลังสองน้อยที่สุด