บทที่ 2

การประมาณค่าสัมประสิทธิ์ความถดถอย

การวิจัยในครั้งนี้เป็นการศึกษาเปรียบเทียบการประมาณค่าสัมประสิทธิ์ในสมการถดถอย ·ชิงเส้นแบบพหุคูณ (Multipie Reqression) คือ ในกรณีที่มีจำนวนตัวแปรอิสระมากกว่า 1 ตัว ระหว่างวิธีกำลังสองต่ำสุด และวิธีบูตสแตรป สำหรับในบทนี้จะกล่าวถึงรายละเอียดของวิธีการประมาณค่าพารามิเตอร์และคุณสมบัติของตัวประมาณที่ได้ ซึ่งจะเสนอในรูปของเวคเตอร์และเมตริกซ์ ส่วนในตอนท้ายของบนจะนำเสนอผลงานวิจัยที่เกี่ยวข้อง โดยมีรายละเอียดต่าง ๆ ดังนี้

2.1 การประมาณค่าพารามิเตอร์โดยวิธีกำลังสองต่ำสุด

ตัวสถิติที่ใช้ในการประมาณค่าพารามิเตอร์ในกรณีที่ความคลาดเคลื่อนมีการแจกแจงแบบปกติ คือตัวประมาณที่ได้จากวิธีกำลังต่ำสุด สำหรับในกรณีที่ไม่ทราบลักษณะการแจกแจงของความคลาดเคลื่อนจะใช้วิธีทางนอนพารามิเตอร์ในการประมาณค่าพารามิเตอร์ ซึ่งมีอยู่ด้วยกันหลายวิธี ในที่นี้จะศึกษาวิธีบูตสแตรป ซึ่งรายละเอียดจะกล่าวต่อไป

2.1.1 วิธีกำลังสองต่ำสุด (Least Square Method)

วิธีการหาตัวประมาณของพารามิเตอร์วิธีนี้ เป็นวิธีที่มีรากฐานมาจากทฤษฎีการประมาณเชิงเส้น (Theory of Linear Estimation) เป็นวิธีการที่คิดขึ้นโดย **คาร์ล เฟรดริก เกาส์** (Karl Friedrich Gauss 1777-1855) และ **อังเดร แอนดรีวิช มาร์คอฟ** (Andrei Andreevich Markov 1856-1922) ซึ่งมีหลักเกณฑ์ดังนี้คือ ให้หาตัว

ประมาณของพารามิเตอร์ที่ทำให้ผลบวกของกำลังที่สองของความคลาดเคลื่อน (Sum Square Errors (SSE)) มีค่าต่ำสุด ซึ่งรายละเอียดในการหาแสดงในรูปของเวคเตอร์และเมตริกซ์ได้ดังนี้

2.1.1.1 <u>การหาตัวประมาณกำลังสองต่ำสุด</u>

ในที่นี้จะพิจารณาตัวแบบเชิงเส้น (Linear model) ในรูป

$$Y = X\beta + \varepsilon$$

เมื่อ Y เป็นเวคเตอร์ของค่าสังเกตหรือตัวแปรตามขนาด $n \times 1$

X เป็นเมตริกซ์ของตัวแปรอิสระหรือค่าคงที่ที่ทราบค่า ขนาด $n \times p$

ซึ่ง $n > p$ และ rank $x = p$

β เป็นเวคเตอร์ของพารามิเตอร์ที่ไม่ทราบค่า ซึ่งเป็นค่าสัมประสิทธิ์ความถดถอยขนาด $p \times 1$

ε เป็นเวคเตอร์ของความคลาดเคลื่อนขนาด $n \times 1$

ซึ่ง $E(\varepsilon) = 0$

$E(\varepsilon\varepsilon') = \sigma^2 I$ โดยที่ I เป็นเมตริกซ์เอกลักษณ์ (single matrix)

จากหลักเกณฑ์ดังกล่าวข้างต้น จะทำการหา $\hat{\beta}$ ซึ่งเป็นตัวประมาณกำลังสองต่ำสุดของ β ที่ทำให้

$$SSE = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \hat{\varepsilon}'\hat{\varepsilon}$$

มีค่าต่ำสุด โดยการหาอนุพันธ์ของ SSE เทียบกับ $\hat{\beta}$ แล้วกำหนดให้เท่ากับ 0 นั่นคือ

$$\begin{aligned}
SSE = \hat{\varepsilon}'\hat{\varepsilon} &= (Y'-X\hat{\beta})'(Y-X\hat{\beta}) \\
&= (Y'-\hat{\beta}'X')(Y-X\hat{\beta}) \\
&= Y'Y-Y'X\hat{\beta} - \hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \\
&= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta}
\end{aligned}$$

$$\frac{\partial}{\partial\hat{\beta}} SSE = -2X'Y + 2X'X\hat{\beta} = 0$$

$$X'X\hat{\beta} = X'Y$$ ซึ่งเป็นสมการปกติ (Normal Equation)

$$\hat{\beta} = (X'X)^{-1} X'Y$$

เมื่อ $X'X$ ไม่เป็นเมตริกซ์เอกเทศ (Singular matrix)

2.2 คุณสมบัติที่สำคัญของตัวประมาณกำลังสองต่ำสุด

2.2.1 ความไม่เอนเอียง (Unbiasedness)

เมื่อ $Y = X\beta + \varepsilon$ และ $E(\varepsilon) = 0$ ตัวประมาณกำลังสองต่ำสุดของ เป็นตัวประมาณที่ไม่เอนเอียง นั่นคือ $E(\hat{\beta}) = \beta$

พิสูจน์ จาก
$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1} X'Y \\
&= (X'X)^{-1} X'(X\beta + \varepsilon) \\
&= (X'X)^{-1} X'X\beta + (X'X)^{-1} X'\varepsilon \\
&= \beta + (X'X)^{-1} X'\varepsilon \\
E(\hat{\beta}) &= E(\beta) + E(X'X)^{-1} X'\varepsilon \\
&= \beta + (X'X)^{-1} X' E(\varepsilon) \\
&= \beta \qquad\qquad E(\varepsilon) = 0
\end{aligned}$$

$\therefore\ \hat{\beta}$ เป็นตัวประมาณที่ไม่เอนเอียงสำหรับ β

2.2.2 ความแปรปรวนของ $\hat{\beta} = 6^2 (X'X)^{-1}$

เมื่อ $E(\varepsilon) = 0$, $V(\varepsilon) = 6^2 I$

พิสูจน์ จาก $\hat{\beta} - \beta = (X'X)^{-1} X'\varepsilon$

$$\begin{aligned}
V(\hat{\beta}) &= E(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))' \\
&= E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'
\end{aligned}$$

016075

$$= E\,[(X'X)^{-1} X' \varepsilon . \varepsilon' X(X'X)^{-1}]$$

$$= (X'X)^{-1} X' E(\varepsilon . \varepsilon)' X (X'X)^{-1}$$

$$\therefore V(\hat{\beta}) = 6^2 (X'X)^{-1}$$

2.2.3 ค่าประมาณความแปรปรวนของ $\hat{\beta} = 6^2 (X'X)^{-1}$

เมื่อ $6^2 = \frac{1}{n-p} \sum_{i=1}^{n} \varepsilon_i^2$

จาก $V(\hat{\beta}) = 6^2 (X'X)^{-1}$

ในกรณีที่ไม่ทราบค่า 6^2 จึงต้องทำการประมาณ

จาก $\hat{\beta} = \beta + (X'X)^{-1} X' \varepsilon$

พิจารณา

$$\begin{aligned}
\hat{\varepsilon} &= Y - \hat{Y} \\
&= Y - X\hat{\beta} \\
&= X\beta + \varepsilon - X\,[\beta + (X'X)^{-1} X' \varepsilon] \\
&= X\beta + \varepsilon - X\beta - X(X'X)^{-1} X' \varepsilon \\
&= \varepsilon - X(X'X)^{-1} X' \varepsilon \\
&= [I_n - X(X'X)^{-1} X'] \varepsilon \\
&= M \varepsilon
\end{aligned}$$

เมื่อ $M = I_n - X(X'X)^{-1} X'$ เป็น symmetric matrix

และเป็น Idempotent matrix นั่นคือ $M^2 = M$ และ $M' = M$

พิสูจน์

$$\begin{aligned}
MM &= [I_n - X(X'X)^{-1} X'][I_n - X(X'X)^{-1} X'] \\
&= I_n - X(X'X)^{-1} X' - X(X'X)^{-1} X' + X(X'X)^{-1} X'X(X'X)^{-1} X' \\
&= I_n - X(X'X)^{-1} X' - X(X'X)^{-1} X' + X(X'X)^{-1} X' \\
&= I_n - X(X'X)^{-1} X' \\
&= M
\end{aligned}$$

<u>พิสูจน์</u> $M' = [I_n - X(X'X)^{-1} X]'$

$= [I_n - X(X'X)^{-1} X]'$

$= M$

M เป็น Idempotent matrix จริง

จาก $E(\hat{\varepsilon}'\hat{\varepsilon}) = E(\sum_{i=1}^{n} \hat{\varepsilon}_i^2)$

$= E(\hat{\varepsilon} - E(\hat{\varepsilon}))'(\hat{\varepsilon} - E(\hat{\varepsilon}))$

$= E(\hat{\varepsilon}'\hat{\varepsilon})$

$= E(\varepsilon' M'M\varepsilon)$

$= E(\varepsilon' M\varepsilon)$

$= E(tr\varepsilon' M\varepsilon)$

$= E(tr\, M\varepsilon\varepsilon')$

$= tr\, E(M\varepsilon\varepsilon')$

$= tr\, M\, E(\varepsilon\varepsilon')$

$= \sigma^2\, tr(M)$

$= \sigma^2 [tr\,(I_n - X(X'X)^{-1} X)]$

$= \sigma^2 [tr\, I_n - tr\,(X(X'X)^{-1} X']$

$= \sigma^2 [tr\, I_n - tr\, X'X(X'X)^{-1}]$

$= \sigma^2 (tr\, I_n - tr\, I_p)$

$= \sigma^2 (n-p)$

$E\left(\frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-p}\right) = \sigma^2$

$\sum_{i=1}^{n} \frac{\hat{\varepsilon}_i^2}{n-p}$ เป็นตัวประมาณที่ไม่เอนเอียงสำหรับ σ^2

ค่าประมาณความแปรปรวนของ $\hat{\beta} = \hat{\sigma}^2 (X'X)^{-1}$ เมื่อ $\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{\hat{\varepsilon}_i^2}{n-p}$

2.2.4 ความคลาดเคลื่อนมีการแจกแจงแบบปกติ

$\hat{\beta}$ และ $\hat{\sigma}^2$ จะเป็นตัวประมาณที่มีประสิทธิภาพ(Efficiency Estimator) คือมีค่าความแปรปรวนต่ำสุดในบรรดาตัวประมาณที่ไม่เอนเอียง (Uniformly Minimum Variance Unbiased Estimator (UMVUE)) โดยมีการพิสูจน์ไว้แล้วในทฤษฎีบทดังต่อไปนี้

ทฤษฎีบท 2.1[1] จาก $Y = X\beta + \varepsilon$ เมื่อ ε $N(0, \sigma^2 I)$

จะได้ว่า

1. $\hat{\beta} = (X'X)^{-1} X'Y$ เป็นตัวประมาณแบบภาวะน่าจะเป็นสูงสุดของ β

2. $\hat{\sigma}^2 = \dfrac{1}{n-p} \sum_{i=1}^{n} \varepsilon_i^2$ เป็นตัวประมาณแบบภาวะน่าจะเป็นสูงสุดสำหรับ σ^2

ซึ่งได้ปรับ (adjusted) แล้ว เพื่อให้เป็นตัวประมาณที่ไม่เอนเอียง

3. $\hat{\beta}$ $N(\beta : \beta, \sigma^2 c)$ เมื่อ $c = (X'X)^{-1}$

4. $\dfrac{(n-p)\hat{\sigma}^2}{\sigma^2} = U \sim \chi^2 (u : n-p)$

5. $\hat{\beta}$ และ $\hat{\sigma}^2$ เป็นอิสระ

6. $\hat{\beta}$ และ $\hat{\sigma}^2$ เป็นสถิติที่พอเพียงสำหรับ β และ σ^2

7. $\hat{\beta}$ และ $\hat{\sigma}^2$ เป็นสถิติที่สมบูรณ์

ทฤษฎีบท 2.2[2] จาก $Y = X\beta + \varepsilon$ เมื่อ ε $N(0, \sigma^2 I)$ ให้ $t(\beta, \sigma^2)$ เป็นฟังก์ชั่นใด ๆ ของพารามิเตอร์ β และ σ^2 ซึ่งสามารถหาตัวประมาณที่ไม่เอนเอียงได้ แล้วมี $q(\hat{\beta}, \hat{\sigma}^2)$ เป็นฟังก์ชั่นของสถิติที่พอเพียงเป็นตัวประมาณที่ไม่เอนเอียงของฟังก์ชั่น $t(\beta, \sigma^2)$ จะได้ว่า $q(\hat{\beta}, \hat{\sigma}^2)$ เป็นตัวประมาณที่มีค่าความแปรปรวนต่ำสุดในบรรดา

----

[1]Franklin A.Graybill, Theory and Application of the Linear Model (North Scituate, Massachusetts : DUXBURY PRESS),p.176.

[2]Ibid.

ตัวประมาณที่ไม่เอนเอียง (Uniformly Minimum Variance Unbiased Estimator (UMVUE)) สำหรับ $t(\beta, \sigma^2)$

2.2.5 ความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติ

ค่า $\hat{\beta}$ และ $\hat{\sigma}^2$ ที่ได้จากวิธีกำลังสองต่ำสุดจะไม่เป็นตัวประมาณที่มีคุณสมบัติเป็น UMVUE แต่ $\hat{\beta} = (X'X)^{-1} X'Y$ จะเป็นตัวประมาณที่มีค่าความแปรปรวนต่ำสุดในบรรดาตัวประมาณที่ไม่เอนเอียงและเป็นฟังก์ชั่นเชิงเส้นของ Y เท่านั้น ซึ่งไม่เป็นตัวประมาณที่ดีในการนำไปใช้ประมาณค่าฟั ่ชั่นอื่น ๆ ที่ไม่ใช่ฟังก์ชั่นเชิงเส้นของ Y โดยมีนิยามและการพิสูจ ไว้แล้วในทฤษฎีบทดังต่อไปนี้

นิยามที่ 2.1[3] ให้ $y_1, y_2, \ldots, y_n$ เป็นค่าสังเกตของตัวแปรสุ่ม ซึ่งมี joint c.d.f. $F_y(.: r), r < \Omega$ และ $s(r)$ เป็นฟังก์ชั่นของ r, $\theta = s(r)$ ถ้า $\hat{\theta}$ เป็นฟังก์ชั่นเชิงเส้นของ Y ที่มีค่าความแปรปรวนต่ำสุดในบรรดาตัวประมาณที่ไม่เอนเอียงและเป็นฟังก์ชั่นเชิงเส้นของ Y ด้วยกัน แล้วจะเรียก นี้ว่า Uniformly Best (minimum variance) Linear Unbiased Estimator สำหรับ θ ซึ่งนิยมเรียกตัวประมาณนี้อีกแบบหนึ่งว่า Best Linear Unbiased Estimator (BLUE))

ทฤษฎีบท 2.3[4] เมื่อ $Y = X\beta + \varepsilon$, $E(\varepsilon) = 0$ $cov(\varepsilon) = \sigma^2 I$ และ $\Omega = \{(\beta, \sigma^2) : \beta \in E_p, \sigma^2 > 0\}$ ตัวประมาณกำลังต่ำสุดของ l (เมื่อ l เป็นเวคเตอร์ของค่าคงที่ขนาด p x 1) คือ $l'\hat{\beta}$ และเรียกตัวประมาณนี้ว่า Uniformly Minimum Variance Linear Unbiased Estimator (BLUE) ของ $l'\beta$

---

[3]Ibid., p.218.

[4]Ibid., p.219.

2.3 การประมาณค่าพารามิเตอร์โดยวิธีบูตสแตรป (Bootstrap Method)

วิธีการหาตัวประมาณของพารามิเตอร์วิธีนี้ เป็นวิธีที่เสนอขึ้นโดย Bradley Efron ในปี ค.ศ. 1979 โดยมีหลักเกณฑ์ดังนี้คือ ทำการสุ่มตัวอย่างจากข้อมูลที่เก็บรวบรวมมาแบบใส่คืน (with replacement) ขนาดเท่ากับจำนวนตัวอย่างหรือข้อมูลที่มีอยู่แล้วนั้น เพื่อสร้างข้อมูลชุดใหม่แล้วนำมาใช้ในการประมาณค่าพารามิเตอร์ที่สนใจ

สำหรับการหาตัวประมาณโดยวิธีนี้ สามารถทำได้ 3 แบบ คือ โดยการคำนวณหา.าก สูตร โดยใช้เทคนิคมอนติคาร์โล และโดยหาจากการกระจายของอนุกรมเทย์เลอร์ (Taylor series) ซึ่งส่วนใหญ่จะใช้เทคนิคมอนติคาร์โล โดยนำคอมพิวเตอร์มาใช้เป็นเครื่องมือช่วยสำหรับจำนวนครั้งที่ทำการสุ่มตัวอย่าง (Bootstrap sampling) นั้นควรจะอยู่ในช่วง 50-200 ก็เพียงพอที่จะทำให้ได้ตัวประมาณที่ดี โดยมีขั้นตอนการทำงานดังนี้

1. สร้างตัวเลขสุ่ม เพื่อนำไปใช้ในการสุ่มตัวอย่างแบบใส่คืน

2. จากตัวอย่างที่ได้แต่ละชุด นำมาหาค่าสถิติหรือค่าประมาณของพารามิเตอร์ที่สนใจ

ตัวอย่าง การประมาณค่า standard error ของสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient ($\hat{\rho}$)) หรือ $\widehat{SD}(\hat{\rho})$ โดยวิธีบูตสแตรป

ให้ $X_1, X_2, \ldots, X_n$ iid F ซึ่ง F เป็นการแจกแจงที่ไม่ทราบ

$x_i = (y_i, z_i)$; $i = 1, 2, \ldots, n$

จาก
$$\hat{\rho} = \frac{\sum_{i=1}^{n} Y_i Z_i - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} Z_i / n}{\left[\{\sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)^2/n\}\{\sum_{i=1}^{n} Z_i^2 - (\sum_{i=1}^{n} Z_i)^2/n\}\right]^{1/2}}$$

จะเห็นว่าไม่สามารถหาค่าประมาณของ SD ($\hat{\rho}$) ได้จากข้อมูลชุดนี้โดยใช้สูตรทั่วไป ซึ่ง $\widehat{SD}(x) = \sum_{i=1}^{n} (x_i - \bar{x})^2 / n-1$ เมื่อ n คือ ขนาดของตัวอย่าง จึงนำเอาวิธีบูตสแตรปมาใช้ในการประมาณค่า โดยใช้เทคนิคมอนติคาร์โล

<u>ขั้นตอน</u> การประมาณค่า SD ($\hat{\rho}$) โดยวิธีบูตสแตรป

1) ให้ $\hat{F}$ เป็น empirical probability distribution ของ x_i ; $i = 1,2, \ldots, n$

2) สุ่ม x_i แบบใส่คืน (with replacement) ขนาด n ดังนั้น x_i แต่ละตัวจะมีโอกาสถูกเลือกเป็นตัวอย่าง = $\frac{1}{n}$

ได้ $x_1^*, x_2^*, \ldots, x_n^* \sim$ iid $\hat{F}$

เรียก x_i^* นี้ว่า bootstrap sample คำนวณหาค่า $\hat{\rho}$

3) กระทำในข้อ 2) B ครั้ง จะได้ $\hat{\rho}^{*1}, \hat{\rho}^{*2}, \ldots, \hat{\rho}^{*B}$

4) คำนวณหาค่าประมาณของ SD ($\hat{\rho}$) โดยวิธีบูตสแตรป ($\delta_B(\hat{\rho})$)

ได้ $\delta_B(\hat{\rho}) = \text{var}(\hat{\rho}^*)^{1/2}$

ซึ่ง var ($\hat{\rho}^*$) คือ ค่าประมาณความแปรปรวน (variance) ของ $\hat{\rho}$ ภายใต้ $\hat{F}$

$$\delta_B(\hat{\rho}) = \left[\frac{\sum_{i=1}^{B} (\hat{\rho}^{*i} - \bar{\hat{\rho}}^*)}{B-1} \right]^{1/2}$$

เมื่อ

$$\bar{\hat{\rho}}^* = \frac{\sum_{i=1}^{B} \hat{\rho}^{*i}}{B}$$

และได้ว่า $\bar{\hat{\rho}}^*$ เป็นค่าประมาณของค่าคาดหมายของ $\hat{\rho}$ ($\hat{E}(\hat{\rho})$) ที่ได้จากวิธีบูตสแตรป

ในการศึกษาครั้งนี้ได้นำเอาวิธีบูตสแตรปมาใช้ในการประมาณค่าพารามิเตอร์ในการวิเคราะห์ความถดถอยเชิงเส้น และหาค่าความแปรปรวน ในกรณีที่ไม่ทราบลักษณะการแจกแจง

ของความคลาดเคลื่อน ซึ่งมีรายละเอียดในการหาแสดงในรูปของเวคเตอร์และเมตริกซ์ได้ดังนี้

2.3.1 การหาตัวประมาณโดยวิธีบูตสแตรป

จาก $Y = X\beta + \varepsilon$

เมื่อ $E(\varepsilon) = 0$

$E(\varepsilon\varepsilon') = \sigma^2 I$

$\varepsilon_i \sim$ iid F ; ไม่ทราบการแจกแจง F

การหาตัวประมาณกระทำตามขั้นตอนต่อไปนี้

1. คำนวณหาค่า $\hat{\beta}$ โดยวิธีกำลังสองต่ำสุด

ได้ $\hat{\beta} = (X'X)^{-1} X'Y$

และ $\hat{Y} = X\hat{\beta}$

$\hat{\varepsilon} = Y - \hat{Y}$ ซึ่ง $\sum_{i=1}^{n} \hat{\varepsilon}_i = 0$

ให้ F เป็นการแจกแจงของตัวอย่าง (empirical distribution) $\hat{\varepsilon}_i$

2. สุ่ม $\hat{\varepsilon}_i$ แบบใส่คืน (with replacement) ขนาด n

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

$\varepsilon^*_i \sim$ iid F

3. นำค่า ε^*_i มาพิจารณาโดยรวมไว้ในสมการ

จะได้ $Y^* = X\hat{\beta} + \varepsilon^* = \hat{Y} + \varepsilon^*$

คำนวณหาค่า $\hat{\beta}^*$ โดยวิธีกำลังสองต่ำสุด โดยยึดหลักเกณฑ์เดียวกัน คือทำการหา $\hat{\beta}^*$ ซึ่งเป็นตัวประมาณกำลังสองต่ำสุดของ $\hat{\beta}$ ที่ทำให้

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

โดยการหาอนุพันธ์ของ SSE* เทียบกับ β แล้วกำหนดให้เท่ากับ 0 นั่นคือ

$$SSE^* = \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* = (Y^* - X\hat{\beta}^*)'(Y^* - X\hat{\beta}^*)$$

$$= (Y^{*\prime} - \hat{\beta}^{*\prime}X')(Y^* - X\hat{\beta}^*)$$

$$= Y^{*\prime} Y^{*} - Y^{*\prime} X \hat{\beta}^{*} - \hat{\beta}^{*\prime} X' Y^{*} + \hat{\beta}^{*\prime} X' X \hat{\beta}^{*}$$

$$= Y^{*\prime} Y^{*} - 2\hat{\beta}^{*\prime} X' Y^{*} + \hat{\beta}^{*\prime} X' X \hat{\beta}^{*}$$

$$\frac{\partial SSE}{\partial \beta} = -2X'Y^{*} + 2X'X\hat{\beta} = 0$$

$$\therefore X'X\hat{\beta}^{*} = X'Y^{*}$$

$$\therefore \hat{\beta}^{*} = (X'X)^{-1} X' Y^{*}$$

4. กระทำตามขั้นตอนในข้อ 2-3 ซ้ำ B ครั้ง จะได้ $\hat{\beta}^{*1}, \hat{\beta}^{*2}, \ldots, \hat{\beta}^{*B}$

5. คำนวณหาค่า $\bar{\hat{\beta}}^{*} = \sum_{i=1}^{B} \hat{\beta}^{*i} / B$

จากหลักการของวิธีบูตสแตรป จะได้ว่า $\bar{\hat{\beta}}^{*}$ เป็นค่าประมาณของ $E(\hat{\beta})$ ซึ่ง $E(\hat{\beta}) = \beta$ การศึกษาครั้งนี้จึงสนใจและใช้ $\bar{\hat{\beta}}^{*}$ เป็นตัวประมาณของ β ที่ได้จากวิธีบูตสแตรป

ในที่นี้จะเห็นว่า ε_i^{*} เป็นตัวอย่างที่สุ่มได้โดยวิธีบูตสแตรป (bootstrap sample) ซึ่งมีคุณสมบัติดังนี้คือ

1. $E(\varepsilon_i^{*}) = 0 \quad ; \quad i = 1,2,\ldots, n$
2. $cov(\varepsilon_i^{*}\varepsilon_j^{*}) = 0$
3. $V(\varepsilon_i^{*}) = \frac{1}{n} \sum_{i=1} \hat{\varepsilon}_i^2$

พิสูจน์ $\qquad E(\varepsilon_i^{*}) = 0$

เมื่อ ε_i^{*} เป็นการสุ่มตัวอย่างแบบใส่คืนจาก $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n$

ให้ b_i = จำนวนครั้งที่หน่วย i ตกอยู่ในตัวอย่าง

$b_i \quad \text{Binomial} \quad (n, p = \frac{1}{n})$

$E(b_i) = \frac{n}{n} = 1$

$V(b_i) = n.\frac{1}{n}(1 - \frac{1}{n}) = \frac{n-1}{n}$

$cov(b_i, b_j) = -n.p_i.p_j = -n.\frac{1}{n}.\frac{1}{n} = -\frac{1}{n}$

$E(\varepsilon_i^{*}) = E\{E(\varepsilon_i^{*} / \hat{\varepsilon}_i, \ldots, \hat{\varepsilon}_n)\}$

$$= E\{\sum_{i=1}^{n} p_i \, \varepsilon_i^*\}$$

$$= E\{\sum_{i=1}^{n} \frac{1}{n} \, \varepsilon_i^*\}$$

$$= E\{\frac{1}{n} \sum_{i=1}^{n} b_i \, \hat{\varepsilon}_i\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i \, E(b_i)$$

$$E(\varepsilon_i^*) = 0 \qquad\qquad \sum_{i=1}^{n} \hat{\varepsilon}_i = 0$$

พิสูจน์ $\quad \text{cov}(\varepsilon_i^*, \varepsilon_j^*) = 0$

$$\text{cov}(\varepsilon_i^*, \varepsilon_j^*) = 0$$

เพราะว่า ε_i^* และ ε_j^* เป็นอิสระแก่กัน เนื่องจากการสุ่มแบบใส่คืนจาก $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n$

พิสูจน์ $\quad V(\varepsilon_i^*) = \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i^2$

จาก $\quad V(\varepsilon_i^*) = E(\varepsilon_i^{*2}) - \{E(\varepsilon_i^*)\}^2$

$$= E(\varepsilon_i^{*2})$$

$$= E\{\sum_{i=1}^{n} p_i \, \varepsilon_i^*\}$$

$$= E\{\sum_{i=1}^{n} \frac{1}{n} \, \varepsilon_i^{*2}\}$$

$$= E\{\sum_{i=1}^{n} \frac{1}{n} b_i \, \hat{\varepsilon}_i^2\}$$

$$= \frac{1}{n}\{\sum_{i=1}^{n} \hat{\varepsilon}_i^2 \, E(b_i)\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i^2 \cdot 1$$

$$= \frac{1}{n} \sum_{i=1}^{n} \hat{\varepsilon}_i^2$$

2.4 คุณสมบัติที่สำคัญของตัวประมาณที่ได้จากวิธีบูตสแตรป

2.4.1 ความไม่เอนเอียง (Unbiasedness)

นั่นคือ ตัวประมาณที่ได้จากบูตสแตรป $\bar{\hat{\beta}}^*$ เป็นตัวประมาณที่ไม่เอนเอียงสำหรับ β หรือ $E(\bar{\hat{\beta}}^*) = \beta$

พิสูจน์

$$E(\hat{\beta}^*) = E[E(\hat{\beta}^*/\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n)] \quad \ldots\ldots(1)$$

จาก $Y^* = X\hat{\beta} + \varepsilon^*$

โดยวิธีกำลังสองต่ำสุดได้ว่า $\hat{\beta}^* = [(X'X)^{-1} X'Y^*]$

$$E(\hat{\beta}^*) = E[(X'X)^{-1} X'(X\hat{\beta} + \varepsilon^*)]$$

$$= E[(X'X)^{-1} X'X\hat{\beta} + (X'X)^{-1} X'\varepsilon^*]$$

$$= (X'X)^{-1}X'XE(\hat{\beta}) + (X'X)^{-1}X' E(\varepsilon^*)$$

$$= \hat{\beta} \quad \ldots\ldots(2)$$

$\because$ จากสมการ $Y^* = X\hat{\beta} + \varepsilon^*$ มี $\hat{\beta}$ เป็นพารามิเตอร์ หรือค่าคงที่

นำ (2) ไปแทนใน (1) จะได้ว่า

$$E(\hat{\beta}^*) = E(\hat{\beta}) = \beta$$

$\because$ จากวิธีกำลังสองต่ำสุด $E(\hat{\beta}) = \beta$

$\therefore$ $\hat{\beta}^*$ เป็นตัวประมาณที่ไม่เอนเอียงสำหรับ β จากการทำบูตสแตรป B ครั้ง จะได้ $\beta^{*1}, \beta^{*2}, \ldots, \beta^{*B}$

ให้ $$\bar{\hat{\beta}}^* = \frac{\sum_{i=1}^{B} \hat{\beta}^{*i}}{B}$$

$$E(\bar{\hat{\beta}}^*) = E\left(\frac{\sum_{i=1}^{B} \hat{\beta}^{*i}}{B}\right)$$

$$= \frac{1}{B} \sum_{i=1}^{B} E(\hat{\beta}^{*i})$$

$$= \frac{1}{B} \cdot B\beta$$

$$= \beta$$

$\therefore$ $\bar{\hat{\beta}}^*$ เป็นตัวประมาณที่ไม่เอนเอียงสำหรับ β

จะเห็นว่า $\hat{\beta}^*$ และ $\bar{\hat{\beta}}^*$ ต่างก็เป็นตัวประมาณที่ไม่เอนเอียงสำหรับ β แต่เนื่องจาก $\hat{\beta}^*$ มีค่าความแปรปรวนมากกว่า $\hat{\beta}$ ซึ่งเป็นตัวประมาณที่ได้จากวิธีกำลังสองต่ำสุด และมากกว่า $\bar{\hat{\beta}}^*$ ด้วย จึงไม่พิจารณา $\hat{\beta}^*$ ในการประมาณค่า β

<u>พิสูจน์</u> ความแปรปรวนของ $\hat{\beta}^*$ มากกว่า $\hat{\beta}$ **หรือ** $V(\hat{\beta}^*) > V(\hat{\beta})$

จาก $Y = X\beta + \varepsilon$

ประมาณค่า β โดยวิธีกำลังสองต่ำสุด ได้ $\hat{\beta} = (X'X)^{-1}X'Y$

$$\hat{\varepsilon} = Y - X\hat{\beta} = Y - \hat{Y}$$

แล้วสุ่ม $\hat{\varepsilon}$ แบบใส่คืนได้ ε^*

ให้ $Y^* = X\hat{\beta} + \varepsilon^*$

ประมาณค่า $\hat{\beta}$ โดยวิธีกำลังสองต่ำสุด ได้ $\hat{\beta}^* = (X'X)^{-1}X'Y^*$

$$V(\hat{\beta}^*) = V\{E(\hat{\beta}^*/\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n)\} + E\{V(\hat{\beta}^*/\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n)\}$$

$$= V[\hat{\beta}] + E[V(\hat{\beta}^*/\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n) \quad \ldots\ldots(3)$$

จาก $\hat{\beta}^* = (X'X)^{-1}X'Y^*$

$$= (X'X)^{-1}X'(X\hat{\beta} + \varepsilon)$$

$$= (X'X)^{-1}X'X\hat{\beta} + (X'X)^{-1}X'\varepsilon^*$$

$$= \hat{\beta} + (X'X)^{-1}X'\varepsilon^*$$

$$\hat{\beta}^* - \hat{\beta} = (X'X)^{-1}X'\varepsilon^*$$

และจาก $V(\hat{\beta}^*/\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n) = E\{\hat{\beta}^* - E(\hat{\beta}^*)\}\{\hat{\beta}^* - E(\hat{\beta}^*)\}'$

$$= E\{\hat{\beta}^* - \hat{\beta}\}\{\hat{\beta}^* - \hat{\beta}\}'$$

$$= E[(X'X)^{-1} X' \varepsilon^{*}][(X'X)' X' \varepsilon^{*}]'$$

$$= E[(X'X)^{-1} X' \varepsilon^{*} \varepsilon^{*'} X(X'X)^{-1}]$$

$$= (X'X)^{-1} X' E(\varepsilon^{*} \varepsilon^{*'}) X(X'X)^{-1}$$

$$= (X'X)^{-1} X'X (X'X)^{-1} V(\varepsilon^{*})$$

$$= \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n} (X'X)^{-1} \qquad \ldots\ldots (4)$$

นำ (4) และ $V(\hat{\beta}) = 6^2(X'X)^{-1}$ ไปแทนใน (3) จะได้

$$V(\hat{\beta}^{*}) = 6^2 (X'X)^{-1} + E \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n} (X'X)^{-1}$$

$$= 6^2(X'X)^{-1} + (X'X)^{-1} E \frac{\sum_{i=1} \hat{\varepsilon}_i^2}{n}$$

จากที่พิสูจน์มาแล้วข้างต้นได้ว่า

$$E\left(\frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-p}\right) = 6^2$$

$$E\left(\frac{\sum_{i=1} \hat{\varepsilon}_i^2}{n}\right) = \left(\frac{n-p}{n}\right) 6^2$$

$$\therefore \quad V(\hat{\beta}^{*}) = 6^2(X'X)^{-1} + \left(\frac{n-p}{n}\right)6^2 (X'X)^{-1}$$

จากวิธีกำลังสองต่ำสุดได้ว่า $V(\hat{\beta}) = 6^2(X'X)^{-1}$

จะเห็นได้ว่า $V(\hat{\beta}^{*}) > V(\hat{\beta})$ จึงไม่นำ $\hat{\beta}^{*}$ มาใช้ในการประมาณค่า β

2.4.2 ความแปรปรวนของ $\bar{\hat{\beta}}^{*} = \frac{1}{B} \left(\frac{2n-p}{n}\right) 6^2(X'X)^{-1}$

พิสูจน์ จากการทำบูตสแตรป B ครั้ง

$$\bar{\hat{\beta}}^{*} = \frac{\sum_{i=1}^{B} \hat{\beta}^{*i}}{B}$$

$$V(\hat{\beta}^{*}) = V\left(\frac{\sum_{i=1} \hat{\beta}^{*i}}{B}\right)$$

$$= \frac{1}{B^2}\{V(\sum_{i=1}^{B} \hat{\beta}^{*i})\}$$

$$= \frac{1}{B^2}\{\sum_{i=1} V(\hat{\beta}^{*i}) + 2\sum_{i<j} cov(\beta^{*i}, \beta^{*j})\} \quad \ldots\ldots (5)$$

จากการทำบูตสแตรปแต่ละครั้ง ได้ว่า

$$V(\hat{\beta}^{*i}) = 6^2(X'X)^{-1} + \left(\frac{n-p}{n}\right)6^2(X'X)^{-1}$$

$$= \left(1 + \frac{n-p}{n}\right) 6^2 (X'X)^{-1}$$

$$V(\hat{\beta}^{*i}) = \left(\frac{2n-p}{n}\right)6^2(X'X)^{-1} \quad \ldots\ldots (6)$$

พิจารณา $cov(\hat{\beta}^{*i}, \hat{\beta}^{*j}) = cov\ [(X'X)^{-1} X' Y^{*i}, (X'X)^{-1} X' Y^{*j}]$

เมื่อ Y^{*i} และ Y^{*j} คือ vector ของ Y^{*} ในการบูตสแตรปครั้งที่ i และ j ตามลำดับ

จาก $cov(\hat{\beta}^{*i}, \beta^{*j}) = (X'X)^{-1}X'cov(y^{*i}, y^{*j})\, X(X'X)^{-1} \quad \ldots\ldots (7)$

จาก $Y^{*} = X\hat{\beta} + \varepsilon^{*}$

$$E(Y^{*}) = E(X\hat{\beta} + \varepsilon^{*})$$

$$= XE(\hat{\beta}) + E(\varepsilon^{*})$$

$$= X\hat{\beta}$$

$$cov(Y^{*i}, Y^{*j}) = E[\{Y^{*i} - E(Y^{*i})\}\{Y^{*j} - E(Y^{*j})\}']$$

$$= E[\{Y^{*i} - X\hat{\beta}\}\{Y^{*j} - X\hat{\beta}\}']$$

$$= E(\varepsilon^{*i}, \varepsilon^{*j\prime})$$

$$= cov(\varepsilon^{*i}, \varepsilon^{*j})$$

$= 0$ ..... นำไปแทนใน (7)

จะได้ว่า $cov(\beta^{*i}, \beta^{*j}) = 0$ ..... (8)

นำ (6) และ (8) ไปแทนใน (5) จะได้ว่า

$$V(\bar{\hat{\beta}}^*) = \frac{1}{B^2} \sum_{i=1}^{B} \{ \frac{2n-p}{n} 6^2 (X'X)^{-1} + 0 \}$$

$$= \frac{1}{B^2} \{ B \cdot \frac{2n-p}{n} 6^2 (X'X)^{-1} \}$$

$$\therefore V(\bar{\hat{\beta}}^*) = \frac{1}{B} \frac{(2n-p)}{n} 6^2 (X'X)^{-1} \quad .... (9)$$

จะเห็นได้ว่า $V(\bar{\hat{\beta}}^*) < V(\hat{\beta})$ แต่ไม่ทราบค่า 6^2 จึงต้องทำการประมาณค่า $V(\bar{\hat{\beta}}^*)$

2.4.3 ค่าประมาณความแปรปรวนของ $\bar{\hat{\beta}}^*$ ที่ได้จากการเทคนิคมอนติคาร์โล

จากการทำบูตสแตรป B ครั้ง สามารถหาค่า $\hat{V}(\bar{\hat{\beta}}^*)$ ได้จากสูตรทั่วไป นั่นคือ

จาก $$\hat{V}(\hat{\beta}^*) = \frac{1}{B-1} \sum_{i=1} (\hat{\beta}^{*i} - \bar{\hat{\beta}}^*)(\hat{\beta}^{*i} - \bar{\hat{\beta}}^*)'$$

$$\hat{V}(\bar{\hat{\beta}}^*) = V\left(\frac{\sum_{k=1}^{B} \hat{\beta}^{*k}}{B}\right)$$

$$= \frac{1}{B^2} \{ V(\sum_{k=1}^{B} \hat{\beta}^{*k}) \}$$

$$= \frac{1}{B^2} \{ \sum_{k=1}^{B} V(\hat{\beta}^{*k}) + 2 \sum_{i<j} cov(\beta^{*i}, \beta^{*j}) \}$$

$$= \frac{1}{B^2} \cdot B\, V(\hat{\beta}^*)$$ จากที่พิสูจน์มาแล้วตาม-ข้างต้นได้ว่า

$$= \frac{1}{B} V(\hat{\beta}^*)$$ $cov(\beta^{*i}, \beta^{*j}) = 0$

$$\hat{V}(\bar{\hat{\beta}}^*) = \frac{1}{B(B-1)} \sum_{i=1} (\hat{\beta}^{*i} - \bar{\hat{\beta}}^*)(\hat{\beta}^{*i} - \bar{\hat{\beta}}^*)' \quad .....(10)$$

ถ้าทำการสุ่มตัวอย่างเป็นจำนวนครั้งมาก ๆ นั่นคือ B มีค่าเข้าใกล้ ค่าประมาณของ $V(\bar{\hat{\beta}}^*)$ ที่ได้จากเทคนิคมอนติคาร์โล จะมีค่าเข้าใกล้ค่าของ $V(\bar{\hat{\beta}}^*)$ ที่หาได้จากสูตร

2.5 ผลงานวิจัยที่เกี่ยวข้อง

เนื่องจาก σ^2 เป็นค่าที่ไม่ทราบ จึงไม่สามารถหาค่า $V(\hat{\beta})$ ซึ่งเท่ากับ $\sigma^2(X'X)^{-1}$ ได้ ในปี ค.ศ.1979 Bradley Efron ได้ศึกษาและนำเอาวิธีบูตสแตรปมาใช้ในการประมาณค่า Standard error ของสัมประสิทธิ์ความถดถอย ซึ่งก็คือ $[V(\hat{\beta})]^{1/2}$ ในสมการถดถอยที่มีรูปแบบ(model) ที่ยุ่งยากซับซ้อนและไม่ทราบลักษณะการแจกแจงของความคลาดเคลื่อน ทั้งนี้ตั้งอยู่บนสมมติฐานที่ว่า รูปแบบของสมการถูกต้องและใช้ตัวประมาณของสัมประสิทธิ์ความถดถอยที่ได้จากวิธีกำลังสองต่ำสุดเป็นตัวประมาณของพารามิเตอร์ โดยนำคอมพิวเตอร์มาใช้ในการสุ่มตัวอย่างโดยใช้เทคนิควิธีมอนติคาร์โล ในกรณีที่สมการมีรูปแบบทั่วไปคือ $Y = X\beta + \varepsilon$ จะมีขั้นตอนการหาค่าประมาณของ Standard error ของสัมประสิทธิ์ความถดถอย เช่นเดียวกับตัวอย่าง ในหัวข้อ 2.3 ได้ว่า $V(\hat{\beta}^*) = \sum_{i=1} (\hat{\beta}^{*i}-\bar{\hat{\beta}}^*)(\hat{\beta}^{*i}-\bar{\hat{\beta}})' / B-1$ และสามารถหา $V(\hat{\beta}^*)$ โดยไม่ใช้เทคนิคมอนติคาร์โลได้ซึ่งมีค่าเท่ากับ $\hat{\sigma}^2(X'X)^{-1}$ ซึ่ง $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \varepsilon_i}{n}$ เป็น biased downward ของ σ^2 ทำให้ค่าประมาณของ $V(\hat{\beta})$ น้อยกว่าความเป็นจริง จึงใช้ $V(\hat{\beta}^*)$ ที่ได้จากวิธีบูตสแตรปเป็นค่าประมาณของ $V(\hat{\beta})$ และสรุปผลได้ว่าเมื่อ $B \to \infty$ ค่า $V(\hat{\beta}^*)$ ที่หาได้จะมีค่าเข้าใกล้ $V(\hat{\beta})$ และค่า B ซึ่งเป็นจำนวนครั้งในการทำบูตสแตรปที่เหมาะสมควรอยู่ในช่วง 50 - 200 ครั้ง ต่อมาในปี ค.ศ.1983 David A.Freedman และ Stephen C.Peters ได้นำวิธีการนี้ไปใช้ในการประมาณค่า Standard error และพยากรณ์ในการวิเคราะห์ความถดถอยเชิงเส้น เมื่อขนาดของตัวอย่างมีจำนวนจำกัด (finite sample size) และไม่เป็นไปตามข้อตกลงเบื้องต้น ซึ่งพบมากในทางเศรษฐศาสตร์ เช่นในกรณีที่เกิด Lag Structure Autoregressive, Structure Heteroscedasticity เป็นต้น สรุปผลได้ว่าวิธีบูตสแตรปสามารถให้ค่าประมาณของ Standard error ของสัมประสิทธิ์ความถดถอยที่ดีและเข้าใกล้ค่าจริงมากกว่าค่าประมาณที่หาได้จากสูตรทั่วไป