บทที่ 2

ทฤษฎีและสถิติที่ใช้ในการวิจัย

2.1 การประมาณค่าด้วยวิธีกำลังสองน้อยสุด (Ordinary Least Squares Estimation)[1]

จากตัวแบบเชิงเส้นของการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ

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

เมื่อ $\underset{\sim}{y}$ คือเวกเตอร์ของตัวแปรตามขนาด $n \times 1$

X คือเมทริกซ์ของตัวแปรอิสระขนาด $n \times (p+1)$

$\underset{\sim}{\beta}$ คือเวกเตอร์ของค่าพารามิเตอร์หรือสัมประสิทธิ์การถดถอยขนาด $(p+1) \times 1$

$\underset{\sim}{\varepsilon}$ คือเวกเตอร์ของความคลาดเคลื่อนขนาด $n \times 1$

n คือจำนวนค่าสังเกต

และ p คือจำนวนตัวแปรอิสระ

วิธีกำลังสองน้อยสุดมีหลักเกณฑ์ว่าจะทำการหาตัวประมาณที่ทำให้ค่าผลรวมของค่าความคลาดเคลื่อนกำลังสองมีค่าน้อยที่สุด นั่นคือจะหาค่า $\underset{\sim}{\hat{\beta}}$ ที่ทำให้

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

มีค่าต่ำที่สุด โดยทำการหาอนุพันธ์ของ SSE เทียบกับ $\underset{\sim}{\hat{\beta}}$ แล้วกำหนดให้เท่ากับศูนย์

เพราะว่า $SSE = \underset{\sim}{\hat{\varepsilon}}' \underset{\sim}{\hat{\varepsilon}}$

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

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

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

---

[1] G.A.F.Seber , Linear Regression Analysis , (John Wiley and Sons 1977) , p. 42 – 47 และ S.R. Searle , Linear Models , (John Wiley and Sons 1971) , p. 75 – 89

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

จะได้ว่า $\quad \dfrac{\partial}{\partial \underset{\sim}{\hat{\beta}}} SSE = -2X' \underset{\sim}{y} + 2X'X \underset{\sim}{\hat{\beta}}$

และเมื่อให้เท่ากับศูนย์จะได้สมการปกติดังนี้

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

ดังนั้นจะได้ตัวประมาณกำลังสองน้อยสุด คือ

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

ซึ่งตัวประมาณที่ได้จากวิธีกำลังสองน้อยสุดจะมีคุณสมบัติที่สำคัญ คือ

1. เป็นตัวประมาณไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator (B.L.U.E.)) นั่นคือ $E(\underset{\sim}{\hat{\beta}}) = \underset{\sim}{\beta}$
2. มีค่าความแปรปรวนของตัวประมาณเท่ากับ $\sigma^2 (X'X)^{-1}$ เมื่อ $\sigma^2 = \dfrac{1}{n-p} \sum_{i=1}^{n} \varepsilon_i^2$

2.2 การประมาณค่าด้วยวิธีความควรจะเป็นสูงสุด (Maximum Likelihood Estimation)[2]

วิธีนี้เป็นวิธีการหาตัวประมาณพารามิเตอร์ที่เป็นไปได้ที่ทำให้ฟังก์ชันความควรจะเป็น $\ell(\underset{\sim}{\theta};\underset{\sim}{x})$ มีค่าสูงสุด โดยมีขั้นตอนดังนี้

สำหรับกรณีของเวกเตอร์ $\underset{\sim}{x}$ ที่เป็นไปได้ทั้งหมด (domain) ในฟังก์ชันความหนาแน่น $f(\underset{\sim}{x};\underset{\sim}{\theta})$ ไม่เกี่ยวข้องกับค่า $\underset{\sim}{\theta}$ จะทำการหาอนุพันธ์เพื่อหาค่า $\underset{\sim}{\theta}$ ที่ทำให้ $\ell(\underset{\sim}{\hat{\theta}};\underset{\sim}{x}) \geq \ell(\underset{\sim}{\theta};\underset{\sim}{x})$ สำหรับทุกๆค่าของ θ_i ที่เป็นไปได้ ค่าวิกฤติของฟังก์ชันความควรจะเป็น $(\underset{\sim}{\hat{\theta}})$ คือคำตอบของสมการ $\dfrac{\partial}{\partial \theta_i} \ell(\underset{\sim}{\theta};\underset{\sim}{x}) = 0$ ซึ่งต้องสอดคล้องกับเงื่อนไขที่พอเพียง (sufficient condition) ที่ทำให้ $\underset{\sim}{\hat{\theta}}$ เป็นค่าสูงสุดของ $\ell(\underset{\sim}{\theta};\underset{\sim}{x})$ กล่าวคือ $\dfrac{\partial^2}{\partial \theta_i^2} \ell(\underset{\sim}{\theta},\underset{\sim}{x}) < 0$ เมื่อ $\underset{\sim}{\theta} = \underset{\sim}{\hat{\theta}}$ เราเรียก $\underset{\sim}{\hat{\theta}}$ เรียกว่าเป็นตัวประมาณแบบ M.L.E. สำหรับ $\underset{\sim}{\theta}$ และสมการ $\dfrac{\partial}{\partial \theta_i} \ell(\underset{\sim}{\theta};\underset{\sim}{x}) = 0$ นี้จะเรียกว่าสมการความควรจะเป็น (likelihood equation)

---

[2] Robert V.Hogg , Elliot A.Tania , Probability and Statistical Inference , (Fith Edition 1997 by Prentice Hall,Inc.) , p. 438 – 442 และ S.R. Searle , Linear Models , (John Wiley and Sons 1971) , p. 75 – 89

ถ้าในฟังก์ชันความหนาแน่นมี p พารามิเตอร์ ; $\underset{\sim}{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ ในการหา M.L.E. ของ $\underset{\sim}{\theta}$ จะได้สมการความควรจะเป็น p สมการดังนี้

$$\frac{\partial}{\partial \theta_1} \ell(\underset{\sim}{\theta}; \underset{\sim}{x}) = 0$$

$$\frac{\partial}{\partial \theta_2} \ell(\underset{\sim}{\theta}; \underset{\sim}{x}) = 0$$

....

$$\frac{\partial}{\partial \theta_p} \ell(\underset{\sim}{\theta}; \underset{\sim}{x}) = 0$$

หลังจากนั้นจึงแก้สมการทั้ง p สมการนี้เพื่อหาค่า $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_p$ โดยต้องสอดคล้องกับเงื่อนไข $\frac{\partial^2}{\partial \theta_i^2} \ell(\underset{\sim}{\theta}, \underset{\sim}{x}) < 0$ เมื่อ $\theta_i = \hat{\theta}_i$; i = 1, 2, ..., p

ในบางสถานการณ์เราอาจพบปัญหาว่าการแก้สมการความควรจะเป็นเพื่อหาคำตอบของ $\underset{\sim}{\theta}$ กระทำได้ยาก เช่น ในกรณีที่ฟังก์ชันความควรจะเป็นมีค่าที่ติดฟังก์ชันชี้กำลัง (exponential function) อยู่ เราสามารถใช้ $\ln \ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ แทน $\ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ ได้เนื่องจาก $\ln \ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ เป็นฟังก์ชันหนึ่งต่อหนึ่ง (one-to-one function) และเป็นฟังก์ชันเพิ่ม (increasing function) ด้วย ดังนั้นค่า $\ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ หนึ่งค่าจะให้ค่า $\ln \ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ เพียงค่าเดียวเท่านั้น หรือเมื่อหาอนุพันธ์ $\ln \ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ เทียบกับ $\underset{\sim}{\theta}$ จะได้ว่า

$$\frac{\partial}{\partial \underset{\sim}{\theta}} \ln \ell(\underset{\sim}{\theta}; \underset{\sim}{x}) = \frac{1}{\ell(\underset{\sim}{\theta}; \underset{\sim}{x})} \cdot \frac{\partial}{\partial \underset{\sim}{\theta}} \ell(\underset{\sim}{\theta}; \underset{\sim}{x})$$

และเนื่องจาก $\ell(\underset{\sim}{\theta}; \underset{\sim}{x}) > 0$ เสมอ

ดังนั้นเมื่อให้ $\frac{\partial}{\partial \underset{\sim}{\theta}} \ell(\underset{\sim}{\theta}; \underset{\sim}{x}) = \underset{\sim}{0}$ จะได้ว่า $\frac{\partial}{\partial \underset{\sim}{\theta}} \ln \ell(\underset{\sim}{\theta}; \underset{\sim}{x}) = \underset{\sim}{0}$ ด้วย เราจึงสามารถใช้ $\ln \ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ แทน $\ell(\underset{\sim}{\theta}; \underset{\sim}{x})$ ได้ ซึ่งจะสะดวกในการคำนวณมากกว่า

บางกรณีเราจำเป็นต้องใช้วิธีการวิเคราะห์ทำซ้ำทางตัวเลข (iterative numerical analysis) เข้าช่วยในการพิจารณาหาค่า $\underset{\sim}{\theta}$ ที่จะลู่เข้าสู่ค่าคงที่ค่าหนึ่ง ที่จะทำให้ $\frac{\partial}{\partial \underset{\sim}{\theta}} \ell(\underset{\sim}{\theta}; \underset{\sim}{x}) = 0$ กล่าวคือจากสมการ

$$\widetilde{\theta}_{i+1} = \widetilde{\theta}_i - \frac{\ell'(\widetilde{\theta}_i, \underset{\sim}{x})}{\ell''(\widetilde{\theta}_i, \underset{\sim}{x})}$$

เมื่อ $\ell'(\widetilde{\theta}_i; \underset{\sim}{x}) = \frac{\partial}{\partial \widetilde{\vartheta}_i} \ell(\underset{\sim}{\theta}, \underset{\sim}{x})$

และ $\ell''(\widetilde{\theta}_i; \underset{\sim}{x}) = \frac{\partial^2}{\partial \widetilde{\vartheta}_i^2} \ell(\underset{\sim}{\theta}, \underset{\sim}{x})$

กำหนดค่าเริ่มต้นสำหรับ $\tilde{\theta}_i$ เพื่อหาค่า $\tilde{\theta}_{i+1}$ หลังจากนั้นทำซ้ำๆโดยการแทนค่า $\tilde{\theta}_i$ ด้วยค่า $\tilde{\theta}_{i+1}$ (อันเดิม) ทางขวามือของสมการเพื่อจะคำนวณหาค่าใหม่ จนกระทั่งค่าของ $\tilde{\theta}_{i+1}$ (อันใหม่) ไม่ค่อยเปลี่ยนแปลง ซึ่งจะเป็น M.L.E. ของ $\underset{\sim}{\theta}$ นั่นเอง วิธีการวิเคราะห์ทำซ้ำทางตัวเลขมีอยู่หลายวิธี เช่น วิธีของนิวตัน-ราฟสัน (Newton-Raphson) วิธีของไฟโบแนคซี (Fibonacci) และวิธีของโกลเด้น (Golden) เป็นต้น

<u>ตัวประมาณ MLE สำหรับสัมประสิทธิ์ความถดถอย</u>[3]

จากความสัมพันธ์เชิงเส้นระหว่างตัวแปรตาม(Y) กับตัวแปรอิสระ(X) p ตัวโดยมีค่าความคลาดเคลื่อน (Ɛ) และมีขนาดตัวอย่างเท่ากับ n เราสามารถเขียนในรูปของเมตริกซ์ดังนี้

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

จากข้อสมมติเบื้องต้นของการวิเคราะห์ความถดถอย จะได้ $\underset{\sim}{y} \sim N_n(X\underset{\sim}{\beta}, \sigma_\varepsilon^2 I_n)$ และจะได้ฟังก์ชันความหนาแน่นของ y สำหรับค่าสังเกตที่ i คือ

$$f_{Y_i}(y_i) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp\left[-\frac{\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)'\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)}{2\sigma_\varepsilon^2}\right] \quad (1)$$

เนื่องจากค่าสังเกตทั้ง n ค่าเป็นอิสระซึ่งกันและกัน จะได้ว่าฟังก์ชันความหนาแน่นร่วมอยู่ในรูปของ

$$p(\underset{\sim}{y}) = \prod_{i=1}^{n} f_{Y_i}(y_i) = \frac{1}{\left(\sigma_\varepsilon\sqrt{2\pi}\right)^n}\exp\left[-\frac{\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)'\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)}{2\sigma_\varepsilon^2}\right] \quad (2)$$

ฟังก์ชันความควรจะเป็นคือ

$$\ell(\underset{\sim}{\beta};\sigma_\varepsilon^2) = \frac{1}{\left(2\pi\sigma_\varepsilon^2\right)^{n/2}}\exp\left[-\frac{\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)'\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)}{2\sigma_\varepsilon^2}\right] \quad (3)$$

---

[3] Robert V.Hogg , Elliot A.Tania , <u>Probability and Statistical Inference</u> , (Fith Edition 1997 by Prentice Hall,Inc.) p.438 – 442 และ S.R. Searle , <u>Linear Models</u> , (John Wiley and Sons) , p.75 – 89

ดังนั้น $\ln \ell(\underset{\sim}{\beta};\sigma_\varepsilon^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)'\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)$ (4)

ขั้นต่อไปเราจะหาค่าพารามิเตอร์ $\underset{\sim}{\beta}$ และ σ_ε^2 ที่ทำให้ฟังก์ชันความควรจะเป็นมีค่ามากที่สุด โดยการหาอนุพันธ์บางส่วน (patial derivative) ของสมการ $\ln\ell(\underset{\sim}{\beta};\sigma_\varepsilon^2)$ เทียบกับค่าพารามิเตอร์ $\underset{\sim}{\beta}$ และ σ_ε^2 จะได้

$$\frac{\partial}{\partial \underset{\sim}{\beta}}\ln\ell\left(\underset{\sim}{\beta};\sigma_\varepsilon^2\right) = -\frac{1}{2\sigma_\varepsilon^2}\left(2X'X\underset{\sim}{\beta} - 2X'\underset{\sim}{y}\right)$$

$$\frac{\partial}{\partial \underset{\sim}{\beta}}\ln\ell\left(\underset{\sim}{\beta};\sigma_\varepsilon^2\right) = -\frac{n}{2}\left(\frac{1}{\sigma_\varepsilon^2}\right) + \frac{1}{2\sigma_\varepsilon^4}\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)'\left(\underset{\sim}{y} - X\underset{\sim}{\beta}\right)$$

กำหนดให้สมการเท่ากับศูนย์และแก้สมการจะได้ตัวประมาณความควรจะเป็นสูงสุด $\hat{\underset{\sim}{\beta}}$ และ S_E^2 คือ

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

และ $$S_E^2 = \frac{\left(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}}\right)'\left(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}}\right)}{n}$$

2.3 การประมาณค่าด้วยวิธีตัวประมาณ M (M-Estimation)[4]

เมื่อข้อกำหนดของรูปแบบการถดถอยเกี่ยวกับการแจกแจงค่าความคลาดเคลื่อนไม่เป็นจริง จะมีผลทำให้ค่าประมาณสัมประสิทธิ์การถดถอยที่ได้จากตัวประมาณกำลังสองน้อยสุดไม่มีคุณสมบัติของความแกร่ง เนื่องจากตัวประมาณดังกล่าวมีความไวต่อข้อมูลที่ผิดปกติ ดังนั้นช่วงความเชื่อมั่นของค่าคาดหวังของตัวแปรตาม $\underset{\sim}{y}$ และ $\underset{\sim}{\beta}$ จะไม่มีประสิทธิภาพและไม่สามารถทดสอบสมมติฐานเกี่ยวกับค่าความแปรปรวนของประชากรได้ ฮิวเบอร์ (2507) จึงได้ศึกษาฟังก์ชันของความคลาดเคลื่อนและได้พัฒนาตัวประมาณขึ้นมาใหม่ โดยใช้หลักการพื้นฐานของตัวประมาณความควรจะเป็นสูงสุด แล้วเรียกตัวประมาณที่ได้นี้ว่าตัวประมาณ M (M-estimator) มีขั้นตอนในการหาตัวประมาณดังนี้

---

[4] Peter J.Rousseeuw , Annick M. Le Roy , Robust Regression and Outlier Detection , (Wiley series in Probability and Mathematical Statistic) , p.145 - 154

จากที่กล่าวในหัวข้อ 2.1 การประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีความควรจะเป็นสูงสุด จะต้องทราบรูปแบบฟังก์ชันการแจกแจงของค่าความคลาดเคลื่อน $f(\varepsilon_i)$ ซึ่ง

$$\varepsilon_i = y_i - \underset{\sim}{x}_i^{'}\underset{\sim}{\beta}$$

เมื่อ ε_i คือค่าความคลาดเคลื่อนของค่าสังเกตที่ i

y_i คือตัวแปรตามของค่าสังเกตที่ i

$\underset{\sim}{x}_i^{'}$ คือเวกเตอร์ของตัวแปรอิสระ p ตัว ของค่าสังเกตที่ i

กล่าวคือ $\underset{\sim}{x}_i^{'} = \begin{bmatrix} 1 & x_{i1} & x_{i2} & \ldots & x_{ip} \end{bmatrix}_{1\times(p+1)}$

$\underset{\sim}{\beta}$ คือเวกเตอร์ของสัมประสิทธิ์การถดถอย

$$\underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1)x1}$$

p คือ จำนวนตัวแปรอิสระ

และ n คือ ขนาดตัวอย่าง

จากนั้นจึงทำการหา $\hat{\underset{\sim}{\beta}}$ ที่ทำให้ฟังก์ชันความควรจะเป็นของค่าความคลาดเคลื่อนมีค่าสูงสุด กล่าวคือ

$$\max_{\underset{\sim}{\beta}} \sum_{i=1}^{n} \ln f(y_i - \underset{\sim}{x}_i^{'}\underset{\sim}{\beta}) = \min_{\underset{\sim}{\beta}} \left\{ - \sum_{i=1}^{n} \ln f(y_i - \underset{\sim}{x}_i^{'}\underset{\sim}{\beta}) \right\}$$

เมื่อ $\sum_{i=1}^{n} \ln f(y_i - \underset{\sim}{x}_i^{'}\underset{\sim}{\beta})$ เป็นฟังก์ชันภาวะน่าจะเป็น

ในการหา $\hat{\underset{\sim}{\beta}}$ ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นมีค่าสูงสุด ทำได้โดยการหาอนุพันธ์ของฟังก์ชันภาวะน่าจะเป็นเทียบกับ $\hat{\underset{\sim}{\beta}}$ แล้วกำหนดให้ฟังก์ชันมีค่าเท่ากับศูนย์ ดังนี้

$$\sum_{i=1}^{n} \underset{\sim}{x}_i \frac{\frac{d}{d\hat{\underset{\sim}{\beta}}} \ln f(y_i - \underset{\sim}{x}_i^{'}\hat{\underset{\sim}{\beta}})}{f(y_i - \underset{\sim}{x}_i^{'}\hat{\underset{\sim}{\beta}})} = \underset{\sim}{0}$$

เนื่องจากไม่ทราบการแจกแจงของค่าความคลาดเคลื่อน $f(\varepsilon_i)$ ทำให้ไม่สามารถหาอนุพันธ์ของฟังก์ชันความควรจะเป็นและ $\hat{\underset{\sim}{\beta}}$ ได้ ดังนั้นฮิวเบอร์จึงได้เสนอวิธีตัวประมาณ M โดยมี

การกำหนดฟังก์ชันที่เหมาะสมแทนอนุพันธ์ของฟังก์ชันความควรจะเป็น เรียกว่าฟังก์ชันความแกร่ง ψ (Robust function)

ดังนั้นหลักการของวิธีตัวประมาณ M ก็คือ การหาค่า $\hat{\underset{\sim}{\beta}}$ ที่ทำให้ผลรวมของฟังก์ชันความคลาดเคลื่อน ρ ที่ถูกเลือกให้เหมาะสมกับลักษณะการแจกแจงของ ε_i มีค่าน้อยที่สุด กล่าวคือการหาค่า $\hat{\underset{\sim}{\beta}}$ ที่ทำให้ $\sum_{i=1}^{n}\rho\left(\frac{y_i-\underset{\sim}{x}_i'\underset{\sim}{\beta}}{\sigma}\right)$ มีค่าน้อยที่สุด

ถ้าเลือก $\rho(z)=\frac{1}{2}z^2$ แล้ว $\hat{\underset{\sim}{\beta}}$ จะเป็นตัวประมาณที่ได้จากวิธีกำลังสองน้อยสุด แต่ถ้าเลือก $\rho(z)=-\ln f(z)$ แล้ว $\hat{\underset{\sim}{\beta}}$ จะเป็นตัวประมาณที่ได้จากวิธีความควรจะเป็นสูงสุด

ในที่นี้ σ คือค่าพารามิเตอร์ของเสกล (Scale Parameter) จากการแจกแจงของความคลาดเคลื่อน

การหาค่า $\hat{\underset{\sim}{\beta}}$ ที่ทำให้ให้ผลรวมของฟังก์ชันความคลาดเคลื่อนมีค่าน้อยที่สุด ทำได้โดยการหาอนุพันธ์ของ $\sum_{i=1}^{n}\rho\left(\frac{y_i-\underset{\sim}{x}_i'\underset{\sim}{\beta}}{\sigma}\right)$ เทียบกับ $\hat{\underset{\sim}{\beta}}$ แล้วกำหนดให้เท่ากับศูนย์ จะได้สมการคือ

$$\sum_{i=1}^{n}\psi\left(\frac{\varepsilon_i}{s}\right)\underset{\sim}{x}_i=\underset{\sim}{0} \qquad (5)$$

เมื่อ $\psi(z)=\frac{\partial}{\partial u}\rho(z)$ และ

s คือตัวประมาณที่แกร่งของพารามิเตอร์ของสเกล ซึ่งเป็นตัวประมาณไม่เอนเอียงของส่วนเบี่ยงเบนมาตรฐาน โดยทั่วไปวิธีการคำนวณหา s ที่ดีจะต้องไม่ขึ้นกับค่าความคลาดเคลื่อนที่มีขนาดใหญ่ และสามารถนำค่าไปใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยที่เหมาะสมได้ ซึ่งจะทำให้ฟังก์ชันความคลาดเคลื่อนมีค่าน้อยที่สุดเมื่อกำหนดค่า s คงที่ที่เหมาะสม การเลือกตัวประมาณที่แกร่งของพารามิเตอร์สเกล (s) ที่เหมาะสมควรเลือกตัวประมาณที่มีความเอนเอียงและความแปรปรวนน้อย ซึ่งนิยมใช้มัธยฐานของค่าสัมบูรณ์ความเบี่ยงเบน (Median Absolute Deviation : MAD) และทำการปรับด้วยค่าคงที่ 0.6745 ซึ่งเป็นค่าที่ทำให้ s เป็นตัวประมาณไม่เอนเอียงของ σ เมื่อ n มีขนาดใหญ่และความคลาดเคลื่อนมีการแจกแจงเป็นปกติ นั่นคือ

$$s=\frac{median\ |\varepsilon_i-median\ (\varepsilon_i)|}{0.6745}$$

โดยที่ ε_i เป็นค่าความคลาดเคลื่อนของค่าสังเกตตัวที่ i ที่ได้จากตัวประมาณกำลังสองน้อยที่สุด

การเลือกตัวประมาณนี้จะประมาณพารามิเตอร์ของค่าน้อยที่สุดของ $\sum_{i=1}^{n} \rho\left(\frac{\varepsilon_i}{s}\right)$ ในตัวแบบการถดถอยเชิงเส้น เราสามารถหาตัวประมาณ M จาก

$$\min_{\underset{\sim}{\beta}} \sum_{i=1}^{n} \rho\left(\frac{\varepsilon_i}{s}\right) = \min_{\underset{\sim}{\beta}} \sum_{i=1}^{n} \rho\left(\frac{y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}}{s}\right) \qquad (6)$$

สำหรับฟังก์ชัน $\rho(\varepsilon_i)$ ของวิธีกำลังสองน้อยสุดนั้นจะอยู่ในรูปของ ε_i^2 โดยค่า s ไม่จำเป็นต้องใช้เนื่องจาก $\rho(\varepsilon_i)$ เป็นฟังก์ชันที่มีความแปรปรวนเพียงค่าเดียวเท่านั้น และผลลัพธ์ของการประมาณ $\underset{\sim}{\beta}$ ไม่แปรเปลี่ยนตาม s

ในการหาค่าน้อยที่สุดจากสมการ (5) จะหาอนุพันธ์ของ ρ เทียบกับ $\underset{\sim}{\beta}$ แล้วกำหนดให้เท่ากับศูนย์ โดย $\psi = \rho'$ ดังนี้

$$\sum_{i=1}^{n} x_{ij} \psi\left(\frac{y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}}{s}\right) = \underset{\sim}{0} \qquad ; \quad j = 1, 2, \ldots, m \qquad (7)$$

จะได้ว่า ฟังก์ชัน ψ ไม่เป็นฟังก์ชันเชิงเส้น วิธีการที่ใช้ในการประมาณค่าสัมประสิทธิ์ความถดถอยส่วนใหญ่จะใช้วิธีกำลังสองน้อยสุดถ่วงน้ำหนักซ้ำหลายรอบ (Iteratively Reweighted Least Square) ของบีทตันและตูกี (Beaton and Tukey) เราสามารถเขียนสมการ m สมการของสมการที่ (6) ได้ดังนี้

เนื่องจาก
$$\sum_{i=1}^{n} x_{ij} \psi\left(\frac{y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}}{s}\right) = \frac{\sum_{i=1}^{n} x_{ij} \psi\left[\left(y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}\right)/s\right]}{\left(y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}\right)/s} \cdot \left(y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}\right)/s$$

จาก (6) จะได้ว่า
$$\frac{\sum_{i=1}^{n} x_{ij} \psi\left[\left(y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}\right)/s\right]}{\left(y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}\right)/s} \cdot \left(y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}\right)/s = \underset{\sim}{0}$$

$$\sum_{i=1}^{n} x_{ij} w_{i0} \left(y_i - \underset{\sim}{x}_i{}' \underset{\sim}{\beta}\right) = \underset{\sim}{0} \qquad ; \quad j = 1, 2, \ldots, m \qquad (8)$$

เมื่อ

$$w_{i0} = \begin{cases} \psi \dfrac{\left|\left(y_i - \underset{\sim}{x}_i{}' \beta_0\right)/s\right|}{\left(y_i - \underset{\sim}{x}_i{}' \beta_0\right)/s} & ; \quad y_i \neq \underset{\sim}{x}_i{}' \hat{\beta}_0 \\ 1 & ; \quad y_i = \underset{\sim}{x}_i{}' \hat{\beta}_0 \end{cases}$$

สมการเหล่านี้อาจเขียนในรูปเมทริกซ์ได้ ดังนี้

$$X'W_0 X \underset{\sim}{\beta} = X'W_0 \underset{\sim}{y} \qquad (9)$$

เมื่อ W เป็นเมทริกซ์ทแยงมุมขนาด n X n ของน้ำหนัก (Weight) และมีสมาชิกตามเส้นทแยงมุมหลักเป็น w_{10} , w_{20} ,…, w_{n0}

ดังนั้นตัวประมาณ $\underset{\sim}{\hat{\beta}}$ ที่หาได้ในครั้งแรกจะมีค่าเป็น

$$\underset{\sim}{\hat{\beta}} = \left(X'W_0 X\right)^{-1} X'W_0 \underset{\sim}{y} \qquad (10)$$

เราจะทำครั้งถัดไปด้วยการคำนวณน้ำหนัก ซึ่งใช้ $\underset{\sim}{\beta}$ แทน $\underset{\sim}{\beta}$ เดิมและทำซ้ำๆกัน จนกระทั่งได้ค่า $\underset{\sim}{\hat{\beta}}$ ที่ค่อนข้างคงที่

ถ้า $\underset{\sim}{\hat{\beta}}$ มีค่าคงที่ในรอบที่ k จะได้ตัวประมาณ คือ

$$\underset{\sim}{\hat{\beta}} = \left(X'W_0 X\right)^{-1} X'W_0 \underset{\sim}{y}$$

จากสมการที่ (5) จะต้องเลือกฟังก์ชันความแกร่ง ψ หรืออาจเรียกว่าเกณฑ์ความแกร่ง (Robust Criterion) ที่เหมาะสม เพื่อทำให้สัมประสิทธิ์การถดถอยที่ได้มีความแกร่งและมีประสิทธิภาพ

ในงานวิจัยนี้ได้ศึกษาถึงเกณฑ์ความแกร่ง 2 เกณฑ์ คือ

1) **เกณฑ์ความแกร่งของแรมเซย์** (Ramsay's robust criterion function)[5] มีรูปแบบฟังก์ชัน ρ , ฟังก์ชัน ψ และฟังก์ชัน w ดังนี้

$$\rho(z) = a^{-2}[1 - \exp(-a|z|) \cdot (1 + a|z|)] \qquad ; |z| < \infty$$

ถ้า $\psi(z) = \frac{\partial}{\partial z}\rho(z)$ แล้ว

$$\psi(z) = z\exp(-a|z|) \qquad ; |z| < \infty$$

และ $w(z) = \frac{\psi(z)}{z}$ แล้ว

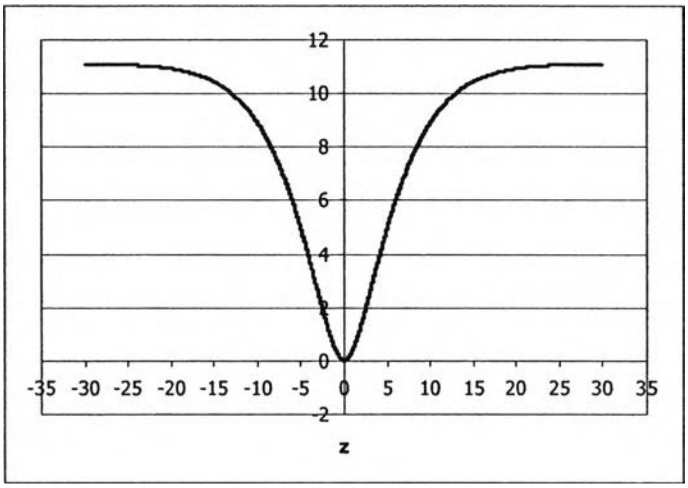
$$w(z) = \exp(-a|z|) \qquad ; |z| < \infty$$

[5] ปราณี รัตนัง , "การประมาณค่าสัมประสิทธิ์การถดถอยพหุเมื่อความผิดพลาดมีการแจกแจงแบบเบ้และมีการแจกแจงแบบหางยาวกว่าปกติ" , (วิทยานิพนธ์ปริญญามหาบัณฑิตภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย 2531) , หน้า 19-21

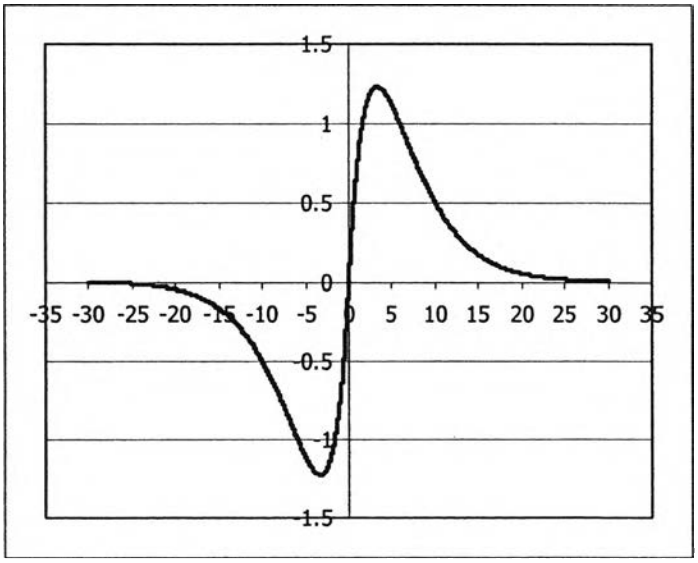
เมื่อ z คือค่าความคลาดเคลื่อนมาตรฐาน (ε_i/s)
และ a คือค่าคงที่ (turning constant) ซึ่งกำหนดมาเพื่อให้ตัวประมาณที่ได้มีประสิทธิภาพตามต้องการการเปรียบเทียบกับตัวประมาณ OLS ซึ่งในที่นี้กำหนดให้เท่ากับ 0.3 และคำนวณโดยให้มีประสิทธิภาพ 95%

ฟังก์ชันของ ρ และ ψ ของแรมเซย์ เมื่อพบค่าผิดพลาดที่มีค่ามาก (Extreme Value) จะมีอิทธิพลลดลงเรื่อยๆ และอิทธิพลเหล่านี้จะอยู่ในสัดส่วนที่ถูกตัดออกไป นั่นคือค่าสังเกตที่ผิดปกติมากๆจะอยู่ในสัดส่วนที่ถูกตัดออกจากตัวอย่าง ลักษณะกราฟของฟังก์ชัน เป็นดังนี้

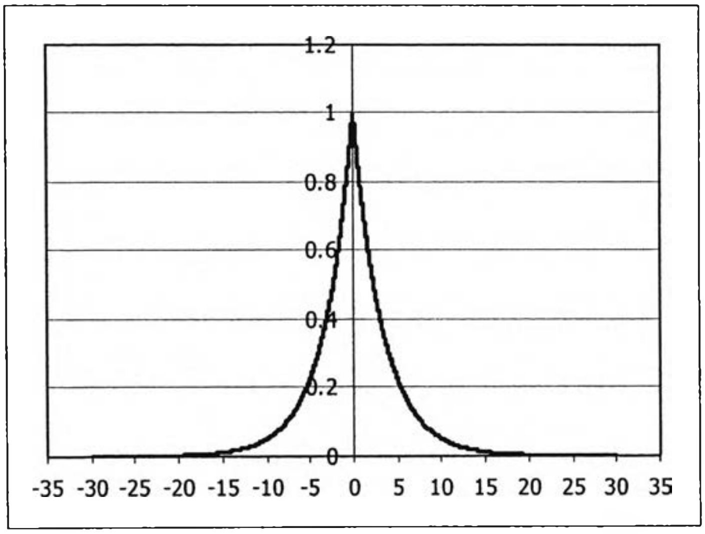
รูปที่ 2.1 ฟังก์ชันความคลาดเคลื่อน $\rho(z)$ ของแรมเซย์



รูปที่ 2.2 ฟังก์ชันความแกร่ง $\psi(z)$ ของแรมเซย์



รูปที่ 2.3 ฟังก์ชันถ่วงน้ำหนัก $w(z)$ ของแรมเซย์



1) **เกณฑ์ความแกร่งของตูกี** (Tukey's robust criterion function)[6] มีรูปแบบฟังก์ชัน ρ , ฟังก์ชัน ψ และฟังก์ชัน w ดังนี้

$$\rho(z) = \begin{cases} \dfrac{z^2}{2} - \dfrac{z^4}{2a^2} + \dfrac{z^6}{6a^4} & ;|z| \le a \\ \dfrac{a^2}{6} & ;|z| > a \end{cases}$$

ถ้า $\psi(z) = \dfrac{\partial}{\partial z}\rho(z)$ จะได้ว่า

$$\psi(z) = \begin{cases} 0 & ;z < -a \\ z\left[1-\left(\dfrac{z}{a}\right)^2\right]^2 & ;-a \le z \le a \\ 0 & ;z > a \end{cases}$$

และถ้า $w(z) = \dfrac{\psi(z)}{z}$ จะได้ว่า

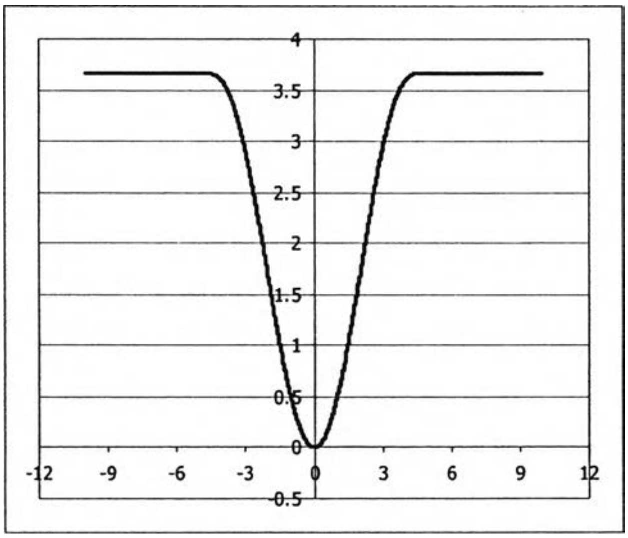
[6] จิตรวี วีระประดิษฐ์ , "การประมาณค่าพารามิเตอร์ในสมการถดถอยเชิงเส้นพหุเมื่อข้อมูลมีค่าผิดปกติ" , (วิทยานิพนธ์ปริญญามหาบัณฑิตภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย 2539) , หน้า 27-29

$$w(z)=\begin{cases}0 & ;z<-a\\ \left[1-\left(\frac{z}{a}\right)^2\right]^2 & ;-a\le z\le a\\ 0 & ;z>a\end{cases}$$

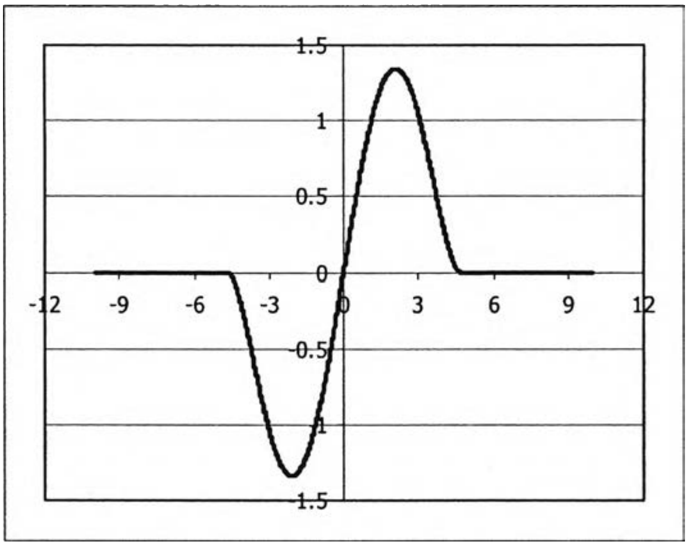
เมื่อ z คือค่าความคลาดเคลื่อนมาตรฐาน $(\varepsilon_i \,/\, s)$
และ a คือค่าคงที่ (Turning Constant) ซึ่งกำหนดมาเพื่อให้ตัวประมาณที่ได้มีประสิทธิภาพตามต้องการการเปรียบเทียบกับตัวประมาณ OLS ซึ่งในที่นี้กำหนดให้เท่ากับ 4.685 และคำนวณโดยให้มีประสิทธิภาพ 95%

ฟังก์ชันของ ρ และ ψ ของตูกี เมื่อพบค่าผิดพลาดที่มีค่ามาก (Extreme Value) จะมีอิทธิพลลดลงเรื่อยๆ และอิทธิพลเหล่านี้จะอยู่ในสัดส่วนที่ถูกตัดออกไป นั่นคือค่าสังเกตที่ผิดปกติมากๆจะอยู่ในสัดส่วนที่ถูกตัดออกจากตัวอย่าง ลักษณะกราฟของฟังก์ชัน เป็นดังนี้

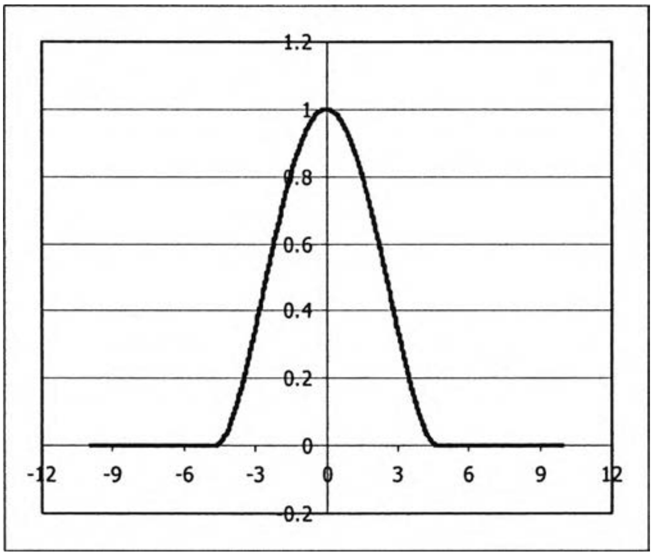
รูปที่ 2.4 ฟังก์ชันความคลาดเคลื่อน $\rho(z)$ ของตูกี



รูปที่ 2.5 ฟังก์ชันความแกร่ง $\psi(z)$ ของตูกี



รูปที่ 2.6 ฟังก์ชันถ่วงน้ำหนัก $w(z)$ ของตูกี



เกณฑ์ความแกร่งของตัวประมาณ M นั้นมีผู้ที่คิดเกณฑ์ขึ้นมาอีกหลายวิธี ดังแสดงไว้ในตารางต่อไปนี้

ตารางที่ 2.1 แสดงฟังก์ชันความแกร่งชนิดต่างๆ

เกณฑ์วิธีการต่างๆ	ρ(z) (ฟังก์ชันความคลาดเคลื่อน)	ψ(z) (ฟังก์ชันความแกร่ง)	w(z) (ฟังก์ชันถ่วงน้ำหนัก)	ขอบเขตของความคลาดเคลื่อน
ตัวประมาณกำลังสองน้อยสุด	$\frac{1}{2}z^2$	z	1.0	$\|z\| < \infty$
ฟังก์ชันความแกร่งของฮิวเบอร์เมื่อ t = 2	$\frac{1}{2}z^2$ $\|z\|t - \frac{1}{2}t^2$	z tsign(z)	1.0 $\frac{t}{\|z\|}$	$\|z\| \le t$ $\|z\| > t$
ฟังก์ชันความแกร่งของแฮมเพล เมื่อ a = 1.7 b = 3.4 c =8.5	$\frac{1}{2}z^2$ $a\|z\| - \frac{1}{2}a^2$ $\frac{a\left(c\|z\| - \frac{1}{2}z^2\right)}{c-b} - \frac{z}{6}a^2$ a(b+c-a)	z asign(z) $\frac{asign(z)(c-\|z\|)}{c-b}$ 0	1.0 $\frac{a}{\|z\|}$ $\frac{a(c-\|z\|)}{\|z\|(c-b)}$ 0	$\|z\| \le a$ $a < \|z\| \le b$ $b < \|z\| \le c$ $\|z\| > c$
ฟังก์ชันความแกร่งของแอนดริว เมื่อ a = 1.399	$a\left[1 - \cos\left(\frac{z}{a}\right)\right]$ 0	$\sin\left(\frac{z}{a}\right)$ 0	$\frac{\sin(z/a)}{z/a}$ 0	$\|z\| \le a\pi$ $\|z\| > a\pi$
ฟังก์ชันความแกร่งของแรมเซย์ เมื่อ a = 0.3	$a^{-2}[1 - \exp(-a\|z\|) \cdot (1 + a\|z\|)]$	$z\exp(-a\|z\|)$	$\exp(-a\|z\|)$	$\|z\| < \infty$
ฟังก์ชันความแกร่งของตูกี เมื่อ a = 4.685	$\frac{z^2}{2} - \frac{z^4}{2a^2} + \frac{z^6}{6a^4}$ $\frac{a^2}{6}$	0 $z\left[1 - \left(\frac{z}{a}\right)^2\right]^2$	0 $\left[1 - \left(\frac{z}{a}\right)^2\right]^2$	$\|z\| \le a$ $\|z\| > a$

2.4 การประมาณค่าด้วยวิธีบูตสแตรป (Bootstrap Estimation)[7]

วิธีบูตสแตรป (Bootstrap method) เป็นวิธีการประมาณค่าไม่อิงพารามิเตอร์ (non-parametric estimation) ที่เสนอโดยเบรดลี่ เอฟรอน (Bradley Efron) ในปี พ.ศ.2520 วิธีบูตสแตรปเป็นวิธีที่ให้ผลดีที่สุดในบรรดาวิธีไม่อิงพารามิเตอร์อื่นๆเพราะการหาค่าประมาณด้วยวิธีบูตสแตรปนี้จะได้ตัวประมาณความควรจะเป็นสูงสุดไม่อิงพารามิเตอร์ (nonparametric maximum likelihood estimater) หลักเกณฑ์ของวิธีบูตสแตรป คือ ทำการสุ่มตัวอย่างจากข้อมูลที่เก็บรวบรวมมาแบบคืนที่ (with replacement) ขนาดเท่ากับจำนวนตัวอย่างหรือข้อมูลที่มีอยู่แล้วนั้น เพื่อสร้างข้อมูลชุดใหม่แล้วนำมาใช้ในการประมาณค่าพารามิเตอร์ที่สนใจ

ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ มีรูปแบบเป็นสมการเมตริกซ์ คือ

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

และเงื่อนไขของการหาตัวประมาณสัมประสิทธิ์ความถดถอย คือค่าเฉลี่ยความคลาดเคลื่อนต้องเป็นศูนย์ ($E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$) และความแปรปรวนมีค่าเป็นค่าคงที่ค่าหนึ่ง ($Var(\varepsilon_i) = \sigma^2$) หรือเขียนในรูปเมตริกซ์ได้เป็น $E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}') = \sigma^2 \mathrm{I}_n$

นอกจากนี้ค่าความคลาดเคลื่อนต้องมีการแจกแจงแบบเดียวกันและเป็นอิสระซึ่งกันและกัน เราสามารถทำการประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีบูตสแตรปได้ดังนี้

<u>ขั้นที่ 1</u> เริ่มจากใช้ตัวประมาณกำลังสองน้อยสุด

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

จะได้ว่า $\hat{\underset{\sim}{y}} = X\hat{\underset{\sim}{\beta}}$ และ $\hat{\underset{\sim}{\varepsilon}} = \underset{\sim}{y} - \hat{\underset{\sim}{y}}$ ซึ่ง $\sum_{i=1}^{n} \hat{\varepsilon}_i = 0$

<u>ขั้นที่ 2</u> ให้ f เป็นฟังก์ชันการแจกแจงของตัวอย่าง (empirical distribution) $\hat{\varepsilon}_i$

สุ่ม $\hat{\varepsilon}_i$ แบบคืนที่ (with replacement) ขนาด n จะได้ $\varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_n^*$

<u>ขั้นที่ 3</u> นำค่า ε_i^* มาพิจารณารวมไว้ในสมการ จะได้

---

[7] Freedman D.A. , "Bootstrapping Regression Models" , <u>The Annals of Statistics 9</u> (April 1981) : 233-234

Bradley Efron , Robert J.Tibshirani , <u>An Introduction to the Bootstrap</u> , (Chapman and Hall Inc.1993) , p. 105 – 112 และ มาลี ตระการศิรินนท์ "การเปรียบเทียบการประมาณค่าพารามิเตอร์ของรูปแบบสมการความถดถอยเชิงเส้นด้วยวิธีกำลังสองต่ำสุดและวิธีบูตสแตรป" , (วิทยานิพนธ์ปริญญามหาบัณฑิตภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย 2532) , หน้า 22-23

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

คำนวณหาค่า $\underset{\sim}{\hat{\beta}}^{*}$ ด้วยวิธีกำลังสองน้อยสุดโดยยึดหลักเกณฑ์เดียวกัน คือทำการหา $\underset{\sim}{\hat{\beta}}^{*}$ ซึ่งเป็นตัวประมาณกำลังสองน้อยสุดของ $\underset{\sim}{\hat{\beta}}$ ที่ทำให้

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

เราสามารถทำได้โดยการหาอนุพันธ์ของ SSE* เทียบกับ $\underset{\sim}{\beta}$ แล้วกำหนดให้เท่ากับศูนย์

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

จะได้ว่า $\frac{\partial}{\partial \underset{\sim}{\hat{\beta}}^{*}} SSE = -2X'\underset{\sim}{y}^{*} + 2X'X\underset{\sim}{\hat{\beta}}^{*}$

และเมื่อให้เท่ากับศูนย์จะได้สมการปกติดังนี้

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

จะได้ตัวประมาณกำลังสองน้อยสุดในขั้นแรกเป็น

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

<u>ขั้นที่ 4</u> เก็บค่า $\underset{\sim}{\hat{\beta}}^{*}$ ไว้ใน $\underset{\sim}{\hat{\beta}}^{*i}$ แล้วกลับไปทำในขั้นตอนที่ 2 และ 3 ซ้ำทั้งหมด B รอบ จะได้ตัวประมาณ $\underset{\sim}{\hat{\beta}}^{*i}$ ทั้งหมด B ค่า กล่าวคือจะได้ตัวประมาณ $\underset{\sim}{\hat{\beta}}^{*1}, \underset{\sim}{\hat{\beta}}^{*2}, \ldots, \underset{\sim}{\hat{\beta}}^{*B}$

<u>ขั้นที่ 5</u> นำค่า $\underset{\sim}{\hat{\beta}}^{*i}$ ทั้งหมด B ค่ามาคำนวณหาค่าเฉลี่ยจากสูตร $\overline{\underset{\sim}{\hat{\beta}}}^{*} = \frac{\sum_{i=1}^{B} \underset{\sim}{\hat{\beta}}^{*i}}{B}$

จากหลักการของวิธีบูตสแตรป จะได้ว่า $\overline{\underset{\sim}{\hat{\beta}}}^{*}$ เป็นค่าประมาณของ $E\left(\underset{\sim}{\hat{\beta}}\right)$ ซึ่ง $E\left(\underset{\sim}{\hat{\beta}}\right) = \underset{\sim}{\beta}$ การศึกษาครั้งนี้จึงสนใจและใช้ $\overline{\underset{\sim}{\hat{\beta}}}^{*}$ เป็นตัวประมาณของ $\underset{\sim}{\beta}$ ที่ได้จากวิธีบูตสแตรป ซึ่งตัวประมาณที่ได้จะมีคุณสมบัติที่สำคัญ คือ

1) $\bar{\hat{\beta}}^{*}_{\sim}$ เป็นตัวไม่เอนเอียงประมาณของ $\underset{\sim}{\beta}$ นั่นคือ $E\left(\bar{\hat{\beta}}^{*}_{\sim}\right) = \underset{\sim}{\beta}$

2) ค่าความแปรปรวนร่วมของ $\bar{\hat{\beta}}^{*}_{\sim}$ คือ

$$Cov\left(\bar{\hat{\beta}}^{*}_{\sim}\right) = \frac{1}{B}\left(\frac{2n-p}{n}\right)\sigma^{2}(X'X)^{-1}$$

และจะได้ว่า $Var(\bar{\hat{\beta}}_i^{*}) < Var(\hat{\beta}_i)$ เมื่อ $\hat{\beta}_i$ เป็นตัวประมาณที่ได้จากวิธีกำลังสองน้อยสุด แต่ $Var(\hat{\beta}_i^{*}) > Var(\hat{\beta}_i)$ เราจึงไม่นำค่า $\hat{\beta}^{*}_{\sim}$ มาพิจารณาในการประมาณค่าของ $\underset{\sim}{\beta}$ แม้ว่า $\hat{\beta}^{*}_{\sim}$ จะเป็นตัวประมาณไม่เอนเอียงของ $\underset{\sim}{\beta}$ เหมือนกับ $\bar{\hat{\beta}}^{*}_{\sim}$ ก็ตาม

นอกจากนี้จะเห็นว่า ε_i^{*} เป็นตัวอย่างที่สุ่มได้โดยวิธีบูตสแตรป (Bootstrap Sample) ซึ่งมีคุณสมบัติดังนี้ คือ

1. $E(\varepsilon_i^{*}) = 0$
2. $Cov(\varepsilon_i^{*}, \varepsilon_j^{*}) = 0$
3. $Var(\varepsilon_i^{*}) = \frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^{2}$