บทที่ 2

ตัวสถิติและผลงานวิจัยที่เกี่ยวข้อง

2.1 ความนำ

การทดสอบสมมติฐานเกี่ยวกับความแตกต่างระหว่างอิทธิพลของทรีตเม้นต์ในการวิเคราะห์ความแปรปรวนร่วมจะใช้ F-test โดยต้องการเปรียบเทียบวิธีประมาณพารามิเตอร์ระหว่างวิธีบูตสแตรปกับวิธีกำลังสองน้อยที่สุด ในบทนี้จะกล่าวถึงรายละเอียดเกี่ยวกับตัวสถิติที่ใช้ในการทดสอบสมมติฐานในการวิเคราะห์ความแปรปรวนร่วมของแผนการทดลองแบบสุ่มตลอด (completely randomized design) รวมทั้งวิธีประมาณพารามิเตอร์โดยใช้วิธีกำลังสองน้อยที่สุดและวิธีบูตสแตรป ตลอดจนผลงานวิจัยที่เกี่ยวข้อง

2.2 ตัวสถิติที่ใช้ในการทดสอบสมมติฐาน

การวิเคราะห์ความแปรปรวนร่วมเป็นวิธีที่ผสมผสานหลักการของการวิเคราะห์ความแปรปรวนกับการวิเคราะห์ความถดถอยเข้าด้วยกัน เมื่อมีตัวแปรร่วมที่มากับหน่วยทดลอง ซึ่งตัวแปรร่วมนี้จะเป็นแหล่งความแปรปรวนที่มิได้ควบคุมด้วยการทดลองและมีผลต่อตัวแปรตาม ในการวิเคราะห์ความแปรปรวนร่วมนี้เป็นการปรับผลการทดลองโดยขจัดส่วนที่มิได้ควบคุมอันเป็นผลมาจากตัวแปรร่วมออกไป

ตัวแบบของการวิเคราะห์ความแปรปรวนของแผนการทดลองแบบสุ่มตลอด

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

$j = 1,2, \ldots n_i$

ตัวแบบของการวิเคราะห์ความถดถอย เมื่อตัวแปรร่วมมีความสัมพันธ์เชิงเส้นกับตัวแปรตามคือ

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

จาก (1) และ (2) จะได้ว่า

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

$$j = 1,2,3,\ldots n_i$$

$$k = 1,2,3,\ldots q$$

โดยที่ p เป็นจำนวนทรีตเม้นต์

n_i เป็นขนาดตัวอย่างในทรีตเม้นต์ที่ i

q เป็นจำนวนตัวแปรอิสระ

y_{ij} หมายถึง ค่าสังเกตของตัวแปรตามจากหน่วยทดลองที่ j ทรีตเม้นต์ที่ i

μ หมายถึง ค่าเฉลี่ยทั้งหมดของ y

τ_i หมายถึง อิทธิพลของทรีตเม้นต์ที่ i

β_k หมายถึง ค่าสัมประสิทธิ์ความถดถอยของตัวแปรร่วมที่ k

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

ε_{ij} หมายถึง ค่าความคลาดเคลื่อนของการทดลองจากหน่วยทดลองที่ j ทรีตเม้นต์ที่ i

การทดสอบสมมติฐานเพื่อเปรียบเทียบความแตกต่างระหว่างอิทธิพลของทรีตเม้นต์ในการวิเคราะห์ความแปรปรวนร่วมจะพิจารณาตัวแบบและตัวสถิติดังนี้

ตัวแบบเต็มรูป (full model)

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

ตัวแบบลดรูป (reduced model)

$$y_{i,j} = \mu + \sum_{k=1}^{q} \beta_k (X_{i.j.k} - \bar{X}_{..k}) + \varepsilon_{i,j}$$

ต้องการทดสอบ สมมติฐานว่าง H_o : $\tau_1 = \tau_2 = \tau_3 = \ldots = \tau_p = 0$

เทียบกับ H_a : มีทรีตเม้นต์อย่างน้อยที่สุด 2 ทรีตเม้นต์ ที่แตกต่างกัน

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

เมื่อ SSF คือ ผลบวกกำลังสองของการประมาณของตัวแบบเต็มรูป

SSR คือ ผลบวกกำลังสองของการประมาณของตัวแบบลดรูป

SST คือ ผลบวกกำลังสองรวม

n = จำนวนหน่วยตัวอย่างทั้งหมดซึ่งเท่ากับ $\sum_{i=1}^{p} n_i$

จะปฏิเสธสมมติฐานว่าง เมื่อค่าตัวสถิติ $F > F_{\alpha}(p-1, n-p-q)$

จากตัวแบบเต็มรูป

$$y_{i,j} = \mu + \tau_i + \sum_{k=1}^{q} \beta_k (X_{i.j.k} - \bar{X}_{..k}) + \varepsilon_{i,j}$$

เราสามารถเขียนอยู่ในรูปสมการถดถอย โดยการสร้างตัวแปร dummy เพื่อแสดงอิทธิพลของทรีตเม้นต์ได้ดังนี้

$$D_{i,1} = \begin{cases} 1 & \text{เมื่อเป็นทรีตเม้นต์ที่ 1} \\ 0 & \text{เมื่อเป็นทรีตเม้นต์อื่น} \end{cases}$$

$$D_{i,2} = \begin{cases} 1 & \text{เมื่อเป็นทรีตเม้นต์ที่ 2} \\ 0 & \text{เมื่อเป็นทรีตเม้นต์อื่น} \end{cases}$$

$$\begin{matrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{matrix}$$

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

จะได้ตัวแบบเต็มรูปใหม่ดังนี้

$$y_{i,j} = \alpha_0+\alpha_1 D_{i,1}+\alpha_2 D_{i,2}+\ldots+\alpha_{p-1} D_{i,p-1}+\sum_{k=1}^{q} \beta_k (X_{i,j,k}-\bar{X}_{..k})+ \varepsilon_{i,j}$$

เมื่อ $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{p-1}$ คือสัมประสิทธิ์ความถดถอยของตัวแปร dummy ซึ่งจะแสดงอิทธิพลของทรีตเม้นต์ จากตัวแบบเต็มรูปใหม่สามารถเขียนแทนด้วยเมตริกซ์ได้ดังนี้

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

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

X เป็นเมตริกซ์ของตัวแปรอิสระซึ่งมีขนาด n x (p+q)

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

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

$$\underset{\sim}{y} = \begin{array}{c} \begin{bmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ . \\ . \\ \underset{\sim}{y}_p \end{bmatrix} \end{array} \begin{array}{c} n_1 \\ n_2 \\ . \\ . \\ n_p \end{array} \qquad \underset{\sim}{\varepsilon} = \begin{bmatrix} \underset{\sim}{\varepsilon}_1 \\ \underset{\sim}{\varepsilon}_2 \\ . \\ . \\ \underset{\sim}{\varepsilon}_p \end{bmatrix} \begin{array}{c} n_1 \\ n_2 \\ . \\ . \\ n_p \end{array}$$

$$X = \begin{bmatrix} \underset{\sim}{1} & \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{0} & (\underset{\sim}{x}_{1.1} - \bar{x}_{..1}\underset{\sim}{1}) & . \; . \; . & (\underset{\sim}{x}_{1.q} - \bar{x}_{..q}\underset{\sim}{1}) \\ \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{1} & \underset{\sim}{0} & (\underset{\sim}{x}_{2.1} - \bar{x}_{..1}\underset{\sim}{1}) & . \; . \; . & (\underset{\sim}{x}_{2.q} - \bar{x}_{..q}\underset{\sim}{1}) \\ . & . & & & . & & . \\ . & . & & & . & & . \\ \underset{\sim}{1} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & (\underset{\sim}{x}_{p.1} - \bar{x}_{..1}\underset{\sim}{1}) & . \; . \; . & (\underset{\sim}{x}_{p.q} - \bar{x}_{..q}\underset{\sim}{1}) \end{bmatrix} \begin{array}{c} n_1 \\ n_2 \\ . \\ . \\ n_p \end{array}$$

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

ตัวแบบลดรูป คือ

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

ซึ่งสามารถเขียนแทนด้วยเมตริกซ์ได้ดังนี้

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

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

X เป็นเมตริกซ์ของตัวแปรอิสระซึ่งมีขนาด n x (q+1)

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

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

$$
X = \begin{array}{c}
\left[\begin{array}{cccc}
\underset{\sim}{1} & (\underset{\sim}{x}_{1.1}-\bar{x}_{..1}\,\underset{\sim}{1}) & \cdots & (\underset{\sim}{x}_{1.q}-\bar{x}_{..q}\,\underset{\sim}{1}) \\
\underset{\sim}{1} & (\underset{\sim}{x}_{2.1}-\bar{x}_{..1}\,\underset{\sim}{1}) & \cdots & (\underset{\sim}{x}_{2.q}-\bar{x}_{..q}\,\underset{\sim}{1}) \\
\cdot & \cdot & & \cdot \\
\cdot & \cdot & & \cdot \\
\underset{\sim}{1} & (\underset{\sim}{x}_{p.1}-\bar{x}_{..1}\,\underset{\sim}{1}) & \cdots & (\underset{\sim}{x}_{p.q}-\bar{x}_{..q}\,\underset{\sim}{1})
\end{array}\right]
\end{array}
\begin{array}{c} n_1 \\ n_2 \\ \cdot \\ \cdot \\ n_p \end{array}
$$

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

2.3 วิธีการประมาณพารามิเตอร์

2.3.1 วิธีกำลังสองน้อยที่สุด (Least Square Method)

วิธีประมาณพารามิเตอร์วิธีนี้มีรากฐานมาจากทฤษฎีการประมาณเชิงเส้น (Theory of Linear Estimation) โดย คาร์ล เฟดริก เกาส์ (Karl Friedrick Gauss) ในปี ค.ศ.1777-1855 และอังเดร แอนดรีวิช มาร์คอฟ (Andrei Andrewich Markov) ในปี ค.ศ.1856-1922 ซึ่งมีหลักการในการประมาณพารามิเตอร์คือ ทำให้ผลบวกกำลังสองของความคลาดเคลื่อนมีค่าน้อยที่สุดซึ่งแสดงรายละเอียดได้ดังนี้

จากตัวแบบ

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

เมื่อ $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ และ $Cov(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$

ตัวประมาณกำลังสองน้อยที่สุดของ $\underset{\sim}{\beta}$ คือ $\hat{\underset{\sim}{\beta}}$ ซึ่งทำให้ผลบวกกำลังสองของความคลาด-เคลื่อน (sum square error) หรือ SSE มีค่าน้อยที่สุด

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

การหาค่าน้อยที่สุดของผลบวกกำลังสองของความคลาดเคลื่อนทำได้โดยหาอนุพันธ์ (differentiate) เทียบกับ $\hat{\beta}_i$ แล้วกำหนดให้เท่ากับ 0

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

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

2.3.2 วิธีบูตสแตรป (Bootstrap Method)

วิธีการหาตัวประมาณของพารามิเตอร์วิธีนี้เป็นวิธีที่เสนอขึ้นโดย แบรดเลย์ เอฟรอน (Bradley Efron) ในปี ค.ศ.1979 โดยมีหลักเกณฑ์ดังนี้คือ ข้อมูลที่เก็บรวบรวมมาจะทำการสุ่มตัวอย่างแบบใส่คืน (with replacement) ขนาดเท่ากับจำนวนตัวอย่างหรือข้อมูลที่มีอยู่ เพื่อสร้างข้อมูลชุดใหม่แล้วนำมาใช้ในการประมาณค่าพารามิเตอร์ที่สนใจ

สำหรับการหาตัวประมาณวิธีนี้จะใช้เทคนิคมอนติคาร์โล โดยนำเครื่องคอม-พิวเตอร์มาใช้เป็นเครื่องมือช่วยในการสุ่มตัวอย่าง มีขั้นตอนการทำงานดังนี้

1. สร้างตัวเลขสุ่ม เพื่อนำไปใช้ในการสุ่มตัวอย่างแบบใส่คืน

2. จากตัวอย่างที่ได้แต่ละชุด นำมาหาค่าประมาณของพารามิเตอร์หรือค่าตัวสถิติที่สนใจ

ตัวอย่างของวิธีการสุ่มตัวอย่างแบบใส่คืนและโปรแกรมใช้งานจะเสนอรายละเอียดไว้ในภาคผนวก ก

การหาตัวประมาณโดยวิธีบูตสแตรป

จากตัวแบบ

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

เมื่อ $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ และ $Cov(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$

ความคลาดเคลื่อนเป็นตัวแปรสุ่มแบบต่อเนื่องที่มีการแจกแจงแบบเดียวกันและเป็นอิสระซึ่งกันและกัน (identically independent distribution)

นั่นคือ $\varepsilon_i \sim F \quad , i = 1,2,\ldots,n$

เมื่อ F เป็นการแจกแจงความน่าจะเป็นที่ไม่ทราบ

การหาตัวประมาณกระทำตามขั้นตอนต่อไปนี้

1. คำนวณหาค่า $\hat{\underset{\sim}{\beta}}$ โดยวิธีกำลังสองน้อยที่สุด

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

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

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

กล่าวคือ $\hat{\varepsilon}_i = y_i - \hat{y}_i$, $i=1,2,3,\ldots n$

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

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

3. นำค่า ε^*_i มาพิจารณาโดยรวมไว้ในสมการจะได้ว่า

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

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

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

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

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

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

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

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

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

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

4. กระทำตามขั้นตอนในข้อ 2-3 ซ้ำ 50 ครั้ง จะได้ $\hat{\underset{\sim}{\beta}}^{*1}, \hat{\underset{\sim}{\beta}}^{*2}, \ldots, \hat{\underset{\sim}{\beta}}^{*50}$

5. คำนวณหาค่า $\bar{\hat{\underset{\sim}{\beta}}} = \sum_{i=1}^{50} \hat{\underset{\sim}{\beta}}^{*i} / 50$

2.4 ผลงานวิจัยที่เกี่ยวข้อง

ในปี ค.ศ.1979 แบรดเลย์ เอฟรอน (Bradley Efron) ได้ศึกษาและนำเอาวิธีบูตสแตรปมาใช้ในการประมาณค่าส่วนเบี่ยงเบนมาตรฐานของสัมประสิทธิ์ความถดถอย $(Var(\hat{\beta}_k^{*}))^{1/2}$ ในสมการการถดถอยที่มีตัวแบบยุ่งยากซับซ้อนและไม่ทราบลักษณะการแจกแจงของความคลาดเคลื่อน เมื่อสมการมีตัวแบบทั่วไป คือ

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

และประมาณสัมประสิทธิ์ความถดถอยโดยวิธีกำลังสองน้อยที่สุด สุ่มตัวอย่างซ้ำด้วยเทคนิคมอน-ติคาร์โลโดยใช้เครื่องคอมพิวเตอร์ ซึ่งจะมีขั้นตอนการประมาณค่าส่วนเบี่ยงเบนมาตรฐานของสัมประสิทธิ์ความถดถอยเช่นเดียวกันกับในหัวข้อการหาตัวประมาณโดยวิธีบูตสแตรปจะได้ว่า

$$Var(\hat{\beta}_k^{*}) = \sum_{i=1}^{N} (\hat{\beta}_k^{*i} - \bar{\hat{\beta}}_k^{*})^2 / N-1 \quad , k = 1,2, \ldots ,p$$

โดยที่ N คือจำนวนครั้งในการทำสุ่มตัวอย่างใหม่ของวิธีบูตสแตรปที่เหมาะสมควรอยู่ในช่วง 50-200 ครั้ง

ต่อมาในปี ค.ศ. 1983 เดวิด เอ ฟรีแมน (David A.Freeman) และ สตีเฟน ซี ปีเตอร์ (Stephen C.Peters) ได้นำวิธีการนี้ไปใช้ในการประมาณค่าส่วนเบี่ยงเบนมาตร-ฐานและพยากรณ์ในการวิเคราะห์ความถดถอยเชิงเส้น เมื่อขนาดของตัวอย่างมีจำนวนจำกัดและไม่เป็นไปตามข้อตกลงเบื้องต้นที่กำหนด สรุปผลได้ว่าวิธีบูตสแตรปสามารถประมาณค่าส่วน-เบี่ยงเบนมาตรฐานของสัมประสิทธิ์ความถดถอยได้ดีและเข้าใกล้ค่าจริงมากกว่าค่าประมาณที่หาได้จากสูตรทั่วไป