ปีที่ 16 ฉบับที่ 48 มกราคม - เมษายน 2545

การเปรียบเทียบตัวประมาณกำลังสองน้อยสุดกับตัวประมาณเบส์สำหรับตัวแบบการถดถอยเชิงเส้นเชิงเดียว

*วีรพา ฐานะปรัชญ์

การเปรียบเทียบค่าประมาณของสัมประสิทธิ์การถดถอยของตัวแบบการถดถอยเชิงเส้นเชิงเดียวด้วยตัวประมาณ 4 ตัว คือ ตัวประมาณกำลังสองน้อยสุด (OLS) ตัวประมาณเบส์เมื่อใช้การแจกแจงก่อนที่ไม่ทราบข้อมูล (UNI) ตัวประมาณเบส์เมื่อใช้การแจกแจงก่อนที่ทราบข้อมูล (NOR) และตัวประมาณเบส์เมื่อใช้การแจกแจงก่อนของเจฟฟรีส์ (JEF) และศึกษาวิธีของ NOR ในกรณีการหาค่า Z ที่เหมาะสมซึ่งทำให้วิธีของ NOR มีประสิทธิภาพใกล้เคียงกับวิธีของ OLS และ UNI มากที่สุด โดยที่ค่า Z เป็นค่าที่ทำให้ค่าเฉลี่ยก่อน (prior mean) เบี่ยงเบนจากค่าสัมประสิทธิ์การถดถอยไป Z เท่าของส่วนเบี่ยงเบนมาตรฐานก่อน (prior standard deviation) จากการศึกษาพบว่า UNI และ JEF มีประสิทธิภาพเท่ากัน OLS และ UNI มีประสิทธิภาพใกล้เคียงกัน โดยที่ OLS มีประสิทธิภาพดีกว่า UNI เล็กน้อย และ OLS มีประสิทธิภาพดีกว่า NOR เมื่อสัมประสิทธิ์ความแปรผันของตัวแปรอิสระมีค่าสูง ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าต่ำและขนาดตัวอย่างใหญ่ส่วน NOR มีประสิทธิภาพดีกว่า OLS เมื่อสัมประสิทธิ์ความแปรผันของตัวแปรอิสระมีค่าต่ำ ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าสูงและขนาดตัวอย่างเล็ก ค่า Z ที่เหมาะสมจะแปรผันตามส่วนเบี่ยงเบนมาตรฐานของตัวแปรตาม แต่แปรผกผันกับสัมประสิทธิ์ความแปรผันของตัวแปรอิสระและขนาดตัวอย่าง โดยค่า Z ที่เหมาะสมมีแนวโน้มสูงขึ้น แสดงว่า NOR มีแนวโน้มของประสิทธิภาพสูงกว่า OLS และ UNI มากขึ้นแต่ในทางกลับกันถ้าค่า Z ที่เหมาะสมมีแนวโน้มต่ำลง NOR จะมีแนวโน้มของประสิทธิภาพสูงกว่า OLS และ UNI ลดลง

ที่มาและความสำคัญของปัญหา

โดยทั่วไปงานวิจัยที่ใช้ตัวแบบการถดถอยเชิงเส้นเชิงเดียว (simple linear regression model) จะใช้วิธีกำลังสองน้อยสุดในการประมาณค่าพารามิเตอร์ในตัวแบบ วิธีการดังกล่าวเป็นวิธีการที่ใช้เพียงข้อมูลปัจจุบันในการประมาณค่าพารามิเตอร์ซึ่งจะทำให้เกิดความคลาดเคลื่อนของการประมาณค่าระดับหนึ่ง ดังนั้นในการลดความคลาดเคลื่อนให้ต่ำลง จึงควรใช้ข้อมูลในอดีตของพารามิเตอร์มาช่วยในการประมาณ กล่าวคือ

*นักวิจัยและอาจารย์ประจำฝ่ายวิจัยและพัฒนา ศูนย์วิจัย มหาวิทยาลัยธุรกิจบัณฑิตย์ : สต.ม. (สถิติศาสตรมหาบัณฑิต) จุฬาลงกรณ์มหาวิทยาลัย

ควรใช้การแจกแจงก่อน (prior distribution) ของพารามิเตอร์มาพิจารณาด้วย โดยแนวความคิดนี้เป็นการใช้การวิเคราะห์เชิงเบส์ ซึ่งมีส่วนประกอบที่สำคัญ 3 ส่วน คือ ข้อมูลปัจจุบันหรือฟังก์ชั่นความควรจะเป็น (likelihood function) ข้อมูลในอดีตหรือการแจกแจงก่อน (prior distribution) โดยแบ่งออกได้เป็น 2 แบบ คือ การแจกแจงก่อนที่ทราบข้อมูล (informative prior distribution) ซึ่งให้ข้อมูลอย่างแน่ชัด และการแจกแจงก่อนที่ไม่ทราบข้อมูล (noninformative prior distribution) ซึ่งให้ข้อมูลคลุมเคลือ (vague) และส่วนประกอบสุดท้าย คือ ข้อมูลอนาคตหรือการแจกแจงภายหลัง ซึ่งเป็นผลมาจากการทราบข้อมูลปัจจุบันและข้อมูลในอดีต

ทฤษฎีและตัวสถิติที่เกี่ยวข้อง

ตัวแบบการถดถอยเชิงเส้นทั่วไปสามารถเขียนในรูปเมทริกซ์ดังนี้

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

เมื่อ $\underset{\sim}{y}$ คือ เวกเตอร์ค่าสังเกตของตัวแปรตามขนาด n x 1

X คือ เมทริกซ์ตัวแปรอิสระขนาด $n \times (p+1)$ และมีค่าลำดับชั้นเท่ากับ p

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

$\underset{\sim}{\varepsilon}$ คือ เวกเตอร์ความคลาดเคลื่อนขนาด $n\times 1$ โดยที่ $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ และ $COV(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$

n คือ ขนาดตัวอย่าง

และ p คือ จำนวนตัวแปรอิสระ

1. ตัวประมาณกำลังสองน้อยสุด ($\underset{\sim}{b}$)

$$\underset{\sim}{b} = (x'x)^{-1}x'\underset{\sim}{y}$$

ซึ่งมีคุณสมบัติเป็นตัวประมาณไม่เอนเอียงเชิงเส้นที่มีความแปรปรวนต่ำสุด (Best Linear Unbiased Estimator : BLUE) โดยเมทริกซ์ของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\underset{\sim}{b}$ มีดังนี้

$$MSE(\underset{\sim}{b}) = \sigma^2(x'x)^{-1}$$

2. ทฤษฎีของเบส์

$$p(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y}) = \frac{p(\underset{\sim}{y} \mid \underset{\sim}{\beta},\sigma)p(\underset{\sim}{\beta},\sigma)}{p(\underset{\sim}{y})}$$

$$\alpha \;\; p(\underset{\sim}{y} \mid \underset{\sim}{\beta},\sigma)p(\underset{\sim}{\beta},\sigma)$$

เมื่อ $p(\underset{\sim}{y} \mid \underset{\sim}{\beta},\sigma)$ คือ ฟังก์ชันความหนาแน่นร่วมของ $\underset{\sim}{y}$ (joint density function) หรือ

$L(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y})$ คือ ฟังก์ชันความควรจะเป็นร่วม (joint likelihood function) สำหรับ $(\underset{\sim}{\beta},\sigma)$ เมื่อทราบข้อมูล $\underset{\sim}{y}$

$p(\underset{\sim}{\beta},\sigma)$ คือ การแจกแจงก่อนร่วม (joint prior distribution) สำหรับ $(\underset{\sim}{\beta},\sigma)$

และ $p(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y})$ คือ การแจกแจงภายหลังร่วม (joint posterior) สำหรับ $(\underset{\sim}{\beta},\sigma)$

2.1 การแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนที่ทราบข้อมูล กรณีทราบค่า σ^2

จากตัวแบบการถดถอยเชิงเส้นปกติ (normal linear regression model) เราจะได้ฟังก์ชันความควรจะเป็นสำหรับ $\underset{\sim}{\beta}$ คือ $L(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y})$ และการแจกแจงก่อนสังยุค (conjugate prior distribution) สำหรับ $\underset{\sim}{\beta}$ คือ $P(\underset{\sim}{\beta})$ แล้วจะหาการแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนที่ทราบข้อมูล กรณีทราบค่า σ^2 ดังนี้

$$p(\underset{\sim}{\beta} \mid \underset{\sim}{y}) \;\; \alpha \;\; l(\underset{\sim}{\beta} \mid \underset{\sim}{y})p(\underset{\sim}{\beta})$$

$$\alpha \;\; \exp\left\{-\frac{1}{2}(\underset{\sim}{\beta}-\underset{\sim}{\bar{\beta}})'\bar{\Sigma}_{\underset{\sim}{\beta}}^{-1}(\underset{\sim}{\beta}-\underset{\sim}{\bar{\beta}})\right\}$$

$$\sim \;\; N(\underset{\sim}{\bar{\beta}},\bar{\Sigma}_{\underset{\sim}{\beta}})$$

เมื่อ $\bar{\Sigma}_{\underset{\sim}{\beta}}^{-1} = \bar{\Sigma}_{\underset{\sim}{\beta}}^{-1} + X'X/\sigma^2$

และ $\underset{\sim}{\bar{\beta}} = \left[\bar{\Sigma}_{\underset{\sim}{\beta}}^{-1} + X'X/\sigma^2\right]^{-1}\left[\bar{\Sigma}_{\underset{\sim}{\beta}}^{-1}\underset{\sim}{\beta} + (X'X/\sigma^2)\underset{\sim}{b}\right]$

โดยที่ $\bar{\underset{\sim}{\beta}}$ คือ เวกเตอร์ค่าเฉลี่ยก่อน (prior mean)

และ $\bar{\Sigma}_{\underset{\sim}{\beta}}$ คือ เมทริกซ์ความแปรปรวนร่วมก่อน (prior covariance) ของ $\bar{\underset{\sim}{\beta}}$ ซึ่ง $\bar{\Sigma}_{\underset{\sim}{\beta}}^{-1} = A/\sigma^2$ จากการ

กำหนดให้ $A = \sigma^2 \bar{\Sigma}_{\underset{\sim}{\beta}}^{-1}$

โดยเมทริกซ์ของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณเบส์เมื่อใช้การแจกแจงก่อนที่ทราบข้อมูล ($\bar{\underset{\sim}{\beta}}$) มีดังนี้

$$\begin{aligned} MES(\bar{\underset{\sim}{\beta}}) &= COV(\bar{\underset{\sim}{\beta}}) + bias(\bar{\underset{\sim}{\beta}}).bias(\bar{\underset{\sim}{\beta}})' \\ &= \sigma^2 WX'XW' + WA\underset{\sim}{\delta}\underset{\sim}{\delta}'A'W' \end{aligned}$$

เมื่อ bias ($\bar{\underset{\sim}{\beta}}$) คือ เวกเตอร์ความเอนเอียงของ $\bar{\underset{\sim}{\beta}}$

$W = (A + X'X)^{-1}$

และ $\underset{\sim}{\delta} = \bar{\underset{\sim}{\beta}} - \underset{\sim}{\beta}$

2.2 การแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนที่ไม่ทราบข้อมูล กรณีทราบค่า σ^2

การแจกแจงก่อนที่ไม่ทราบข้อมูลสำหรับ $\underset{\sim}{\beta}$ เมื่อทราบค่า σ^2 คือ

$$p(\underset{\sim}{\beta}) = p(\beta_0).p(\beta_1)...p(\beta_p)$$

α ค่าคงที่

ฟังก์ชันความควรจะเป็น คือ

$$\begin{aligned} L(\underset{\sim}{\beta} \mid \underset{\sim}{y}) &\alpha\ L(\underset{\sim}{\beta} \mid \underset{\sim}{y})p(\underset{\sim}{\beta}) \\ &\alpha\ \exp\left\{\frac{-1}{2\sigma^2}(\underset{\sim}{\beta} - \underset{\sim}{b})'X'X(\underset{\sim}{\beta} - \underset{\sim}{b})\right\} \\ &\sim\ N(\underset{\sim}{b}, \sigma^2(X'X)^{-1}) \end{aligned}$$

2.3 การแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนร่วมที่ทราบข้อมูล กรณีไม่ทราบค่า σ^2

การแจกแจงก่อนร่วมสำหรับ $(\underset{\sim}{\beta},\sigma)$ มีดังนี้

$$p(\underset{\sim}{\beta},\sigma) = p(\underset{\sim}{\beta}|\sigma)p(\sigma)$$

$$\alpha \quad \sigma^{-(p+1)-\bar{v}-1}\exp\left[\frac{-1}{2\sigma^2}\left\{\bar{v}\bar{s}^2+(\underset{\sim}{\beta}-\underset{\sim}{\bar{\beta}})'A((\underset{\sim}{\beta}-\underset{\sim}{\bar{\beta}})\right\}\right]$$

ฟังก์ชันความควรจะเป็นร่วมสำหรับ $(\underset{\sim}{\beta},\sigma)$ มีดังนี้

$$L(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y}) \quad \alpha \quad \sigma^{-n}\exp\left\{\frac{-1}{2\sigma^2}(\underset{\sim}{y}-X\underset{\sim}{\beta})'(\underset{\sim}{y}-X\underset{\sim}{\beta})\right\}$$

เราหาการแจกแจงภายหลังร่วมเมื่อใช้การแจกแจงก่อนที่ทราบข้อมูล กรณีไม่ทราบค่า σ^2 มีดังนี้

$$p(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y}) \quad \alpha \quad L(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y})p(\underset{\sim}{\beta},\sigma)$$

$$\sigma^{-n-(p+1)-\bar{v}-1}\exp\left[\frac{-1}{2\sigma^2}\left\{\bar{\bar{v}}\bar{\bar{s}}^2+(\underset{\sim}{\beta}-\underset{\sim}{\bar{\bar{\beta}}})'(A+X'X)(\underset{\sim}{\beta}-\underset{\sim}{\bar{\bar{\beta}}})\right\}\right]$$

เมื่อ $\underset{\sim}{\bar{\bar{\beta}}} = (A+X'X)^{-1}(A\underset{\sim}{\bar{\beta}}+X'X\underset{\sim}{\hat{\beta}})$

$$\bar{\bar{v}}\bar{\bar{s}}^2 = \bar{v}\bar{s}^2+\underset{\sim}{y}'\underset{\sim}{y}+\underset{\sim}{\bar{\beta}}'A\underset{\sim}{\bar{\beta}}-\underset{\sim}{\bar{\bar{\beta}}}'(A+X'X)\underset{\sim}{\bar{\bar{\beta}}}$$

และ $\bar{\bar{v}} = n+\bar{v}$

การแจกแจงภายหลังขอบ (marginal posterior distribution) $p(\underset{\sim}{\beta} \mid \underset{\sim}{y})$ หาได้จาก

$$p(\underset{\sim}{\beta} \mid \underset{\sim}{y}) = \int_0^\infty p(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y})d\sigma$$

$$\alpha \left[1+\frac{1}{\bar{\bar{v}}}(\underset{\sim}{\beta}-\underset{\sim}{\bar{\bar{\beta}}})'\frac{(A+X'X)}{\bar{\bar{s}}^2}(\underset{\sim}{\beta}-\underset{\sim}{\bar{\bar{\beta}}})\right]^{-(\bar{\bar{v}}+p+1)/2}$$

$$\sim t(\bar{\bar{v}}, \underset{\sim}{\bar{\bar{\beta}}}, \left[\bar{\bar{v}}/(\bar{\bar{v}}-2)\right]\bar{\bar{s}}^2(A+X'X)^{-1})$$

2.4 การแจกแจงภายหลังร่วมเมื่อใช้การแจกแจงก่อนร่วมที่ไม่ทราบข้อมูล กรณีไม่ทราบค่า σ^2

การแจกแจงก่อนร่วมที่ไม่ทราบข้อมูลที่ใช้ทั่วไป (Conventional joint prior distribution) สำหรับ $(\underset{\sim}{\beta},\sigma)$ มีดังนี้

$$p(\underset{\sim}{\beta} \mid y) \ \alpha \ \frac{1}{\sigma}$$

ฟังก์ชันความควรจะเป็นร่วมสำหรับ $(\underset{\sim}{\beta},\sigma)$ มีดังนี้

$$L(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y})' \ \alpha \ \sigma^{-n} \exp\left[\frac{-1}{2\sigma^2}\left\{v\hat{\sigma}^2 + (\underset{\sim}{\beta}-\underset{\sim}{b})'X'X(\underset{\sim}{\beta}-\underset{\sim}{b})\right\}\right]$$

เมื่อ $v\hat{\sigma}^2 = (y - X\underset{\sim}{b})'(\underset{\sim}{y}-X\underset{\sim}{b})$ คือ ผลบวกกำลังสองของการประมาณด้วยวิธีกำลังสองน้อยสุดและ $v = n-(p+1)$ คือ ระดับความเป็นเสรี

การแจกแจงภายหลังร่วม (joint posterior distribution) สำหรับ $(\underset{\sim}{\beta},\sigma)$ กรณีไม่ทราบค่า σ^2 มีดังนี้

$$\begin{aligned} p(\underset{\sim}{\beta},\sigma \mid \underset{\sim}{y}) \ & \alpha \ L(\underset{\sim}{\beta},\sigma \mid y)p(\underset{\sim}{\beta},\sigma) \\ & \alpha \ \sigma^{-(n+1)} \exp\left[\frac{-1}{2\sigma^2}\left\{v\hat{\sigma}^2 + (\underset{\sim}{\beta}-\underset{\sim}{b})'X'X(\underset{\sim}{\beta}-\underset{\sim}{b})'\right\}\right] \end{aligned}$$

การแจกแจงภายหลังขอบ $p(\underset{\sim}{\beta} \mid \underset{\sim}{y})$ จากการอินทิกรัล σ ออกจาก $p(\underset{\sim}{\beta},\sigma \mid y)$ มีดังนี้

$$\begin{aligned} p(\underset{\sim}{\beta} \mid y) \ & \alpha \ \left[1+\frac{1}{v\hat{\sigma}^2}(\underset{\sim}{\beta}-\underset{\sim}{b})'X'X((\underset{\sim}{\beta}-\underset{\sim}{b})\right]^{-(p+1+v)/2} \\ & \sim \ t(v,\underset{\sim}{b},[v/(v-2)]\hat{\sigma}^2(X'X)^{-1}) \end{aligned}$$

2.5 การแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนของเจฟฟรีส์

การแจกแจงก่อนของเจฟฟรีส์ (Jeffereys's prior distribution) เป็นการแจกแจงก่อนที่ไม่ทราบข้อมูลชนิดหนึ่ง ซึ่งมีคุณสมบัติไม่แปรเปลี่ยนภายใต้การแปลงพารามิเตอร์ (parameterization invariance) โดยการแจกแจงก่อนของเจฟฟรีส์ในกรณีตัวแบบเชิงเส้น มีดังนี้

$$p(\underset{\sim}{\beta},\sigma) \ \alpha \ \sigma^{-1}$$

จะเห็นได้ว่าการแจกแจงก่อนของเจฟฟรีส์เหมือนกับการแจกแจงก่อนที่ไม่ทราบข้อมูลที่ใช้ทั่วไป (Conventional noninformative prior) ในกรณีไม่ทราบค่า σ^2 และการแจกแจงก่อนของเจฟฟรีส์จะเป็นสัดส่วนกับค่าคงที่ ($p(\underset{\sim}{\beta},\sigma)\alpha$ ค่าคงที่) ในกรณีที่ทราบค่า σ^2 ซึ่งเหมือนกันกับการแจกแจงก่อนที่ไม่ทราบข้อมูลที่ใช้ทั่วไปในกรณีทราบค่า σ^2 ดังนั้นการแจกแจงภายหลัง

เมื่อใช้การแจกแจงก่อนของเจฟฟรีส์เหมือนกันกับการแจกแจงภายหลังเมื่อใช้การแจกแจงก่อนที่ไม่ทราบข้อมูล

แผนการทดลอง

1. สถานการณ์ต่าง ๆ ที่ศึกษา มีดังนี้

1.1 ขนาดตัวอย่างเท่ากับ 10 30 50 และ 100

1.2 กลุ่มตัวอย่างของความคลาดเคลื่อนสุ่มจากประชากรที่มีการแจกแจงปกติ ซึ่งมีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.1 0.3 0.5 0.7 และ 0.9

1.3 ตัวแปรอิสระเป็นค่าคงที่ซึ่งสุ่มมาจากประชากรที่มีการแจกแจงปกติด้วยค่าเฉลี่ยเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.1 0.15 0.2 0.25 0.3 0.5 0.7 และ 0.9 นั่นคือ เปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผันของตัวแปรอิสระเท่ากับ 10% 15% 20% 25% 30% 50% 70% และ 90% ตามลำดับ

1.4 กำหนดค่าสัมประสิทธิ์การถดถอยในตัวแบบ คือ β_0 เท่ากับ 0.5 และ β_1 เท่ากับ 0.5

1.5 จากข้อ 1 ข้อ 3 และข้อ 4 เราจะได้ว่ากลุ่มตัวอย่างของตัวแปรตามมีการแจกแจงปกติ ซึ่งมีค่าเฉลี่ยเท่ากับ 1 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.1 0.3 0.5 0.7 และ 0.9

2. การหาค่า Z ที่เหมาะสม

เนื่องจากค่าเฉลี่ยก่อน (prior mean) ของการแจกแจงก่อนที่ทราบข้อมูลมีผลต่อตัวประมาณเบส์เมื่อใช้การแจกแจงดังกล่าว ดังนั้นผู้วิจัยจึงกำหนดค่าในเมทริกซ์ความแปรปรวนร่วมก่อน (prior covariance matrix) ให้เป็นค่าคงที่ใด ๆ และกำหนดค่าของเวกเตอร์ค่าเฉลี่ยก่อนให้เป็นฟังก์ชันของค่าปกติมาตรฐาน และทำการหาค่า Z ที่เหมาะสมที่ทำทราบค่าเฉลี่ยความเคลื่อนกำลังสองของตัวประมาณสัมประสิทธิ์การถดถอยด้วยวิธีเบส์เมื่อใช้การแจกแจงก่อนที่ทราบข้อมูลมีค่าสูงสุดเท่าที่เป็นไปได้ แต่มีค่าต่ำกว่าค่าเฉลี่ยความคลาดเคลื่อนกำลังสองด้วยวิธีกำลังสองน้อยสุด โดยค่า Z ที่มีค่ายิ่งสูงขึ้นจะทำให้ค่าเฉลี่ยก่อนยิ่งมีค่าห่างจากค่าสัมประสิทธิ์การถดถอย ซึ่งจะส่งผลให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองมีค่าสูงขึ้น

ดังนั้นวิธีการข้างต้นจึงเป็นการหาตัวประมาณเบส์เมื่อใช้การแจกแจงก่อนที่ทราบข้อมูลที่ไม่ดีที่สุดเท่าที่เป็นไปได้ แต่ดีกว่าตัวประมาณกำลังสองน้อยสุดและตัวประมาณเบส์เมื่อใช้การแจกแจงก่อนที่ไม่ทราบข้อมูล

3. ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง

การคำนวณค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square of Error : MSE) ของค่าประมาณสัมประสิทธิ์การถดถอยของแต่ละวิธี โดยกระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์ตามแผนการทดลอง แล้วคำนวณค่าเฉลี่ยของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Average of Mean Square of Error : AMSE) ของแต่ละวิธีโดยการคำนวณเปอร์เซ็นต์ของอัตราส่วนผลต่างค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (DIFF) ดังนี้

$$\mathrm{DIFF} = \frac{\mathrm{AMSEi} - \mathrm{AMSEmin}}{\mathrm{AMSEmin}} \times 100 \ , \ i = 1, 2, 3$$

เมื่อ AMSEi คือ AMSE ของแต่ละวิธี

และ AMSEmin คือ AMSE ของวิธีที่ให้ค่า AMSE ต่ำสุด

โดยที่ i = 1 หมายถึง วิธีกำลังสองน้อยสุด

i = 2 หมายถึง วิธีเชิงเบส์เมื่อใช้การแจกแจงก่อนที่ไม่ทราบข้อมูล

i = 3 หมายถึง วิธีเชิงเบส์เมื่อใช้การแจกแจงก่อนที่ทราบข้อมูล

สรุปผลการวิจัย

1. ผลสรุปจากการเปรียบเทียบตัวประมาณสัมประสิทธิ์การถดถอย เมื่อศึกษากรณีเฉพาะของตัวประมาณเบส์ที่ใช้การแจกแจงก่อนที่ทราบข้อมูล

กรณีทราบค่า σ^2

ค่า AMSE ของแต่ละวิธีแปรผันตามส่วนเบี่ยงเบนมาตรฐานของตัวแปรตาม และแปรผกผันกับเปอร์เซ็นต์ของสัมประสิทธ์ความแปรผันของตัวแปรอิสระและขนาดตัวอย่าง ตามลำดับความสำคัญจากมากไปน้อย

วิธี NOR จะให้ประสิทธิภาพดีเมื่อเปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผันของตัวแปรอิสระมีค่าต่ำ ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าสูงและขนาดตัวอย่างเล็ก และวิธี OLS จะให้ประสิทธิภาพดีเมื่อเปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผันของตัวแปรอิสระมีค่าสูง ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าต่ำและขนาดตัวอย่างใหญ่ ส่วนวิธี UNI จะไม่ให้ประสิทธิภาพดีกว่าวิธีอื่น เนื่องจากค่า AMSE ของวิธี UNI จะมีค่าสูงกว่าค่า AMSE ของวิธี OLS เสมอ โดยวิธี OLS และ NOR จะให้ประสิทธิภาพดีที่สุด ณ สถานการณ์ดังนี้

ตารางที่ 1 วิธีที่ให้ประสิทธิภาพดีที่สุดจำแนกตามสถานการณ์ ในกรณีที่ทราบค่า σ^2

	CV(X) = 10%				CV(X) = 15%				CV(X) = 20%				CV(X) = 25%			
n	10	30	50	100	10	30	50	100	10	30	50	100	10	30	50	100
SD(Y)=0.1	×	×	×	×	×	×	×	×	✓	✓	×	×	✓	✓	✓	✓
SD(Y)=0.3	×	×	×	×	×	×	×	×	×	×	×	×	×	✓	×	✓
SD(Y)=0.5	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
SD(Y)=0.7	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
SD(Y)=0.9	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

ตารางที่ 1 (ต่อ)

	CV(X) = 30%				CV(X) = 50%				CV(X) = 70%				CV(X) = 90%			
n	10	30	50	100	10	30	50	100	10	30	50	100	10	30	50	100
SD(Y)=0.1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.3	×	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.5	×	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.7	×	×	×	✓	×	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.9	×	×	×	×	×	×	✓	✓	×	✓	✓	✓	×	✓	✓	✓

<u>หมายเหตุ</u> ✓ หมายถึงวิธี OLS
× หมายถึงวิธี NOR

กรณีไม่ทราบค่า σ^2

ผลสรุปที่ได้มีลักษณะเดียวกับกรณีทราบค่า σ^2 ซึ่งประสิทธิภาพของวิธี OLS จะเหมือนเดิม แต่ประสิทธิภาพของวิธี UNI และวิธี NOR จะต่ำกว่ากรณีทราบค่า σ^2 ณ สถานการณ์เดียวกัน เนื่องจากค่า AMSE ของวิธี UNI และวิธี NOR ในกรณีไม่ทราบค่า σ^2 จะมีค่าความแปรปรวนจากการประมาณค่าเพิ่มขึ้น และเมื่อขนาดตัวอย่างเพิ่มขึ้น ประสิทธิภาพของทั้งสองกรณีจะใกล้เคียงกันมากขึ้น โดยวิธี OLS และ NOR จะให้ประสิทธิภาพดีที่สุด ณ สถานการณ์ดังนี้

ตารางที่ 2 วิธีที่ให้ประสิทธิภาพดีที่สุดจำแนกตามสถานการณ์ ในกรณีที่ทราบค่า σ^2

	CV(X) = 30%				CV(X) = 50%				CV(X) = 70%				CV(X) = 90%			
n	10	30	50	100	10	30	50	100	10	30	50	100	10	30	50	100
SD(Y)=0.1	×	×	×	×	✓	×	×	×	✓	✓	✓	×	✓	✓	✓	✓
SD(Y)=0.3	×	×	×	×	×	×	×	×	×	×	×	×	×	✓	×	✓
SD(Y)=0.5	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
SD(Y)=0.7	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
SD(Y)=0.9	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

ตารางที่ 2 (ต่อ)

	CV(X) = 30%				CV(X) = 50%				CV(X) = 70%				CV(X) = 90%			
n	10	30	50	100	10	30	50	100	10	30	50	100	10	30	50	100
SD(Y)=0.1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.5	×	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.7	×	×	×	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SD(Y)=0.9	×	×	×	×	×	×	✓	✓	×	✓	✓	✓	✓	✓	✓	✓

<u>หมายเหตุ</u> ✓ หมายถึงวิธี OLS
× หมายถึงวิธี NOR

2. ผลสรุปจากการหาตัวประมาณเบส์ที่ใช้การแจกแจงก่อนที่ทราบข้อมูลที่ไม่ดีที่สุด แต่ดีกว่าตัวประมาณกำลังสองน้อยสุดและตัวประมาณเบส์ที่ใช้การแจกแจงก่อนที่ไม่ทราบข้อมูล

กรณีทราบค่า σ^2

ค่า Z ที่เหมาะสมแปรผันตามส่วนเบี่ยงเบนมาตรฐานของตัวแปรตาม แต่แปรผกผันกับเปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผันของตัวแปรอิสระและขนาดตัวอย่างตามลำดับความสำคัญจากมากไปน้อย

ค่า Z ที่เหมาะสมจะมีค่าสูงขึ้นเมื่อเปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผันของตัวแปรอิสระมีค่าต่ำ ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าสูงและขนาดตัวอย่างเล็ก ส่วนค่า Z ที่เหมาะสมจะมีค่าต่ำเมื่อเปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผันของตัวแปรอิสระมีค่าสูง ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามมีค่าต่ำและขนาดตัวอย่างใหญ่

กรณีไม่ทราบค่า σ^2

ผลสรุปที่ได้มีลักษณะเดียวกับกรณี σ^2 ทราบค่า แต่ค่า Z ที่เหมาะสมในกรณีไม่ทราบค่า σ^2 จะมีค่าต่ำกว่าค่า Z ที่เหมาะสมในกรณีทราบค่า σ^2 ณ สถานการณ์เดียวกัน และเมื่อขนาดตัวอย่างเพิ่มขึ้นค่า Z ที่เหมาะสมของทั้งสองกรณีจะมีค่าใกล้เคียงกันมากขึ้น

วิธีต่าง ๆ ที่ใช้ในการประมาณจะมีประสิทธิภาพสูงขึ้น เมื่อเปอร์เซ็นต์ของสัมประสิทธิ์ความแปรผันของตัวแปรอิสระและขนาดตัวอย่างเพิ่มขึ้น และจะมีประสิทธิภาพลดลงเมื่อส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามเพิ่มขึ้น

ดังนั้น การเลือกวิธีการประมาณค่าสัมประสิทธิ์การถดถอยของตัวแบบการถดถอยที่เหมาะสมในเบื้องต้นสามารถทำได้โดยการวัดการกระจายสัมพัทธ์ของตัวแปรอิสระ (ค่าสัมประสิทธิ์ความแปรผันของตัวแปรอิสระ) หากค่าวัดการกระจายมีค่าค่อนข้างสูง (ตั้งแต่ 50% ขึ้นไป) เราจะเลือกวิธีกำลังสองน้อยสุด และค่าวัดการกระจายมีค่าต่ำ (น้อยกว่า 50%) เราจะเลือกวิธีเชิงเบส์เมื่อใช้การแจกแจงก่อนที่ทราบข้อมูล ❑

บรรณานุกรม

Birkes, David and Dodge, yadolah. **Alternative methods of regression.** New York : John Wiley & Sons, 1993.

Box, George E. P. and Tiao, George C. **Bayesian inference in statistical analysis.** Reading, Mass : Addison-Wesley, 1973.

Press, S. James. **Bayesian statistics : principles, models, and applications.** New York : John Wiley & Sons, 1989.

Vinod, Hrishikesh D. and Ullah, Aman. **Recent advances in regression methods.** New York : Dekker, 1981.