บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในการศึกษาการเปลี่ยนแปลงของตัวแปรตาม (dependent variable) ว่ามีผลมาจากตัวแปรอิสระ (independent variables) ชุดหนึ่งอย่างไรนั้น วิธีการหนึ่งที่ใช้ในการวิเคราะห์หาความสัมพันธ์ของตัวแปรเหล่านั้นคือ การวิเคราะห์การถดถอยพหุคูณ (multiple regression analysis) ซึ่งการวิเคราะห์การถดถอยพหุคูณมีหลักเกณฑ์ว่า การใช้ตัวแปรอิสระที่เหมาะสมมากกว่าหนึ่งตัวขึ้นไปมาช่วยในการอธิบายตัวแปรตาม โดยทั่วไปแล้วจะทำให้ผลการประมาณค่าตัวแปรตามมีความถูกต้องมากกว่าการใช้ตัวแปรอิสระเพียงตัวเดียวถ้าตัวแปรอิสระนั้นไม่มีความสัมพันธ์กันและมีอิทธิพลต่อตัวแปรตามมากพอสมควร ตัวแบบแสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามอยู่ในรูปของ

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

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

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

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

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

n เป็นขนาดตัวอย่างของตัวแปรอิสระ

และ p เป็นจำนวนตัวแปรอิสระ

โดยที่ $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$, $\text{cov}(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$

ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณจากตัวแบบดังกล่าว วิธีที่นิยมใช้กันมาก คือ วิธีกำลังสองน้อยสุด (Ordinary Least Squares method) โดยมีรูปแบบตัวประมาณเป็น $\underset{\sim}{\hat{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$ เมื่อ X มีค่าลำดับชั้นเต็ม (full column rank) = p+1 ซึ่งมีคุณสมบัติคือ เป็นตัวประมาณไม่เอนเอียงและมีความแปรปรวนต่ำสุดในบรรดาตัวประมาณไม่เอนเอียงเชิงเส้น แต่ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุดมีข้อ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สมมติที่จำเป็นข้อหนึ่ง คือ ตัวแปรอิสระแต่ละตัวต้องไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรอิสระตัวอื่น ซึ่งในทางปฏิบัติเป็นไปได้น้อยมากเพราะตัวแปรอิสระต่าง ๆ ที่นำมาศึกษาส่วนใหญ่อาจมีความสัมพันธ์กัน กล่าวคือ ตัวแปรอิสระบางตัวมีพหุสัมพันธ์ (multicollinearity) ทำให้การประมาณค่าตัวแปรตามที่ได้ไม่เหมาะสม และมีผลทำให้ความแปรปรวนของค่าประมาณสัมประสิทธิ์การถดถอยพหุคูณมีค่ามากขึ้น นั่นคือ ค่าประมาณสัมประสิทธิ์การถดถอยพหุคูณที่ได้ขาดความเที่ยงตรง (precision) ดังนั้น ในการกำหนดตัวแบบของความสัมพันธ์ว่าตัวแปรตามมีผลมาจากตัวแปรอิสระใดบ้างควรพิจารณาว่าตัวแปรอิสระมีพหุสัมพันธ์กันหรือไม่ ถ้าตัวแปรอิสระมีความสัมพันธ์กันสูงสามารถแก้ไขโดยการตัดตัวแปรอิสระบางตัวออกจากตัวแบบ แต่ในบางครั้งการตัดตัวแปรอิสระตัวใดตัวหนึ่งออกจากตัวแบบเป็นไปได้ยากเนื่องจากความสัมพันธ์ระหว่างตัวแปรอิสระไม่ชัดเจนพอและตัวแปรอิสระทุกตัวมีผลต่อการเปลี่ยนแปลงของตัวแปรตามมากพอสมควร

เนื่องจากวิธีกำลังสองน้อยสุดเป็นวิธีที่นิยมใช้กันมาก แต่มีปัญหาในกรณีที่ตัวแปรอิสระมีพหุสัมพันธ์กันทำให้ค่าประมาณสัมประสิทธิ์การถดถอยพหุคูณที่ได้ขาดความเที่ยงตรง จึงได้มีวิธีพื้นฐานที่สร้างขึ้นเพื่อที่จะปรับปรุงความเที่ยงตรงของตัวประมาณสัมประสิทธิ์กำลังสองน้อยสุดในกรณีที่ตัวแปรอิสระมีพหุสัมพันธ์กันคือ วิธีกำลังสองน้อยสุดที่ถูกจำกัด (Restricted Least Squares method) โดยหลักการของวิธีนี้ยังคงใช้หลักการของวิธีกำลังสองน้อยสุดเหมือนเดิมเพียงแต่เพิ่มข้อมูลข้อจำกัด (restriction) เกี่ยวกับการรวมเชิงเส้น (linear combination) ของสัมประสิทธิ์การถดถอยพหุคูณที่ไม่ทราบค่าดังนี้

$$R\underset{\sim}{\beta} = \underset{\sim}{r}$$

เมื่อ R เป็นเมทริกซ์แสดงรูปแบบความสัมพันธ์ของพารามิเตอร์ $\underset{\sim}{\beta}$ ขนาด q×(p+1) โดยที่ q < (p+1)

$\underset{\sim}{\beta}$ เป็นเวกเตอร์ของสัมประสิทธิ์การถดถอยพหุคูณขนาด (p+1)×1

และ $\underset{\sim}{r}$ เป็นเวกเตอร์ที่ได้จากการรวมเชิงเส้นของสัมประสิทธิ์การถดถอยพหุคูณขนาด q×1

จะได้ว่ารูปแบบของตัวประมาณสัมประสิทธิ์วิธีกำลังสองน้อยสุดที่ถูกจำกัดที่ได้ คือ

$$\underset{\sim}{\beta}^{*} = \underset{\sim}{\hat{\beta}} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(\underset{\sim}{r} - R\underset{\sim}{\hat{\beta}})$$

เมื่อ $\hat{\beta}$ เป็นตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณวิธีกำลังสองน้อยสุด

วิธีกำลังสองน้อยสุดที่ถูกจำกัดจะเป็นตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณที่มีคุณสมบัติเป็นตัวประมาณไม่เอนเอียง (unbiased estimator) ในกรณีที่ข้อจำกัดนั้นเป็นจริงกล่าวคือ $r - R\beta = 0$ แต่ตัวประมาณจะกลายเป็นตัวประมาณเอนเอียง (biased estimator) ในกรณีที่ข้อจำกัดนั้นไม่เป็นจริงกล่าวคือ $r - R\beta \neq 0$ โดยที่ตัวประมาณวิธีกำลังสองน้อยสุดที่ถูกจำกัดจะให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าวิธีกำลังสองน้อยสุดในกรณีที่ข้อจำกัดนั้นเป็นจริงและไม่เป็นจริง

ในปี ค.ศ. 1970 โฮเอิลและเคนนาร์ด (Hoerl and Kennard) ได้ศึกษาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณที่ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าวิธีกำลังสองน้อยสุด โดยให้ชื่อว่า วิธีริดจ์รีเกรสชัน (Ridge Regression method) ซึ่งเป็นวิธีที่แก้ปัญหาตัวแปรอิสระมีพหุสัมพันธ์กัน และวิธีนี้ไม่ต้องตัดตัวแปรอิสระออกจากตัวแบบ หลักการของวิธีนี้คือพิจารณาคุณสมบัติของ $\hat{\beta}$ จากการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุดจะพบว่า ทั้งเมทริกซ์ความแปรปรวนร่วมของตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณซึ่งอยู่ในรูปของ $\text{cov}(\hat{\beta}) = \sigma^2(X'X)^{-1}$ และความแตกต่างระหว่างตัวประมาณสัมประสิทธิ์กำลังสองน้อยสุด $(\hat{\beta})$ และค่าสัมประสิทธิ์การถดถอยพหุคูณ (β) เมื่อพิจารณาค่าเฉลี่ยกำลังสองของความแตกต่างระหว่าง $\hat{\beta}$ และ β มีค่าเท่ากับ $\sigma^2 tr(X'X)^{-1}$ เป็นฟังก์ชันของ $(X'X)^{-1}$ ดังนั้นการที่จะทำให้เมทริกซ์ความแปรปรวนร่วมของ $\hat{\beta}$ และค่าเฉลี่ยกำลังสองของความแตกต่างระหว่าง $\hat{\beta}$ และ β มีค่าลดลงจึงต้องพยายามลดค่า $(X'X)^{-1}$ ให้ต่ำลงซึ่งจะทำได้โดยการบวกค่าคงที่ที่มากกว่าศูนย์กับสมาชิกทุกตัวบนเส้นทแยงมุม จะทำให้ได้ตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีริดจ์รีเกรสชัน คือ

$$\hat{\beta}_R(k) = (X'X + kI)^{-1}X'y \qquad , \quad k > 0$$

เราสามารถจัดรูปแบบตัวประมาณริดจ์ใหม่โดยใช้สมการปกติของการวิเคราะห์วิธีกำลังสองน้อยสุดซึ่งมีรูปแบบคือ $X'X\hat{\beta} = X'y$ จะทำให้ได้รูปแบบของตัวประมาณดังสมการข้างล่างนี้

$$\hat{\underset{\sim}{\beta}}_R(k) = (X'X + kI)^{-1} X'X \hat{\underset{\sim}{\beta}}$$
$$= [I + k(X'X)^{-1}]^{-1} \hat{\underset{\sim}{\beta}} \quad , \quad k > 0$$

ในปี ค.ศ. 1992 ซาคาร์ (Sarkar) ได้เสนอวิธีประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณในกรณีที่ตัวแปรอิสระมีพหุสัมพันธ์กัน โดยได้นำข้อดีของวิธีริดจ์รีเกรสชันและวิธีกำลังสองน้อยสุดที่ถูกจำกัดมาผสมผสานกันทำให้ตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณมีประสิทธิภาพและความเที่ยงตรงมากขึ้นเรียกวิธีการนี้ว่าวิธีริดจ์รีเกรสชันที่ถูกจำกัด (Restricted Ridge Regression method) โดยมีรูปแบบของตัวประมาณนี้ คือ

$$\underset{\sim}{\beta}^*_R(k) = [I + k(X'X)^{-1}]^{-1} \underset{\sim}{\beta}^* \quad , \quad k > 0$$

เมื่อ $\underset{\sim}{\beta}^*$ เป็นตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณวิธีกำลังสองน้อยสุดที่ถูกจำกัด

ตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณวิธีนี้ได้ปรับปรุงประมาณวิธีกำลังสองน้อยสุดที่ถูกจำกัดโดยใช้หลักการของตัวประมาณริดจ์ ทำให้ตัวประมาณที่ได้จะมีรูปแบบตัวประมาณคล้ายตัวประมาณริดจ์โดยตัวประมาณที่ได้มีคุณสมบัติคือ เป็นตัวประมาณเอนเอียงและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าวิธีกำลังสองน้อยสุดที่ถูกจำกัดและวิธีริดจ์รีเกรสชันในกรณีที่ข้อจำกัดนั้นเป็นจริง นอกจากนี้ในกรณีที่ข้อจำกัดนั้นไม่เป็นจริงได้มีการกำหนดเงื่อนไขของค่า k ที่เป็นไปได้ที่ทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าวิธีกำลังสองน้อยสุดที่ถูกจำกัดและวิธีริดจ์รีเกรสชัน

ในปี ค.ศ. 1993 ลิว คีเจียน (Liu Kejian) ได้เสนอวิธีประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณในกรณีที่ตัวแปรอิสระเกิดพหุสัมพันธ์กัน โดยได้นำข้อดีของวิธีริดจ์รีเกรสชันและวิธีของสไตน์ (Stein) มาผสมผสานกันเรียกวิธีนี้ว่าวิธีลิว ตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณที่ได้มีรูปแบบดังนี้

$$\hat{\underset{\sim}{\beta}}_d(d) = (X'X + I)^{-1}(X' \underset{\sim}{y} + d \hat{\underset{\sim}{\beta}}) \quad , \quad 0 < d < 1$$

เมื่อ $\hat{\underset{\sim}{\beta}}$ เป็นตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณวิธีกำลังสองน้อยสุด

ตัวประมาณนี้ได้ใช้หลักการของตัวประมาณริดจ์คือ การบวกค่าคงที่ค่าหนึ่งเข้ากับสมาชิกในแนวทแยงมุมของเมทริกซ์ $X'X$ ซึ่งค่าคงที่มีค่าเท่ากับหนึ่งและใช้รูปแบบของตัวประมาณสไตน์ $\underset{\sim}{\hat{\beta}}_s = d\,\underset{\sim}{\hat{\beta}}$ บวกเข้าไปในเทอมหลังเพื่อจะแก้ปัญหาการเกิดพหุสัมพันธ์ให้กับเทอมหลัง เมื่อจัดรูปแบบตัวประมาณใหม่โดยใช้สมการปกติของการวิเคราะห์วิธีกำลังสองน้อยสุดซึ่งมีรูปแบบคือ $X'X\,\underset{\sim}{\hat{\beta}} = X'\,\underset{\sim}{y}$ จะทำให้ได้รูปแบบของตัวประมาณดังสมการข้างล่างนี้

$$\begin{aligned}\underset{\sim}{\hat{\beta}}_d(d) &= (X'X + I)^{-1}(X'X\,\underset{\sim}{\hat{\beta}} + d\,\underset{\sim}{\hat{\beta}}) \\ &= (X'X + I)^{-1}(X'X + dI)\,\underset{\sim}{\hat{\beta}} \qquad , \quad 0 < d < 1\end{aligned}$$

ข้อดีของตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีของลิวคือ ตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณวิธีของลิวเป็นฟังก์ชันของ d ทำให้การคำนวณหาค่า d ที่เหมาะสมสะดวกกว่าการคำนวณหาค่า k ที่เหมาะสมจากวิธีริดจ์รีเกรสชัน ต่อมาในปี ค.ศ. 1999 เซลาห์ฮัททิน คาซิแรนลาร์ (Selahattin Kaciranlar) , ซาดุลเลาะห์ ซาคัลลิโอกลู (Sadullah Sakallioglu) และ ฟิคลิ แอคเดนนิส (Fikri Akdeniz) ได้เสนอวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณใหม่โดยได้นำข้อดีของวิธีลิวและวิธีกำลังสองน้อยสุดที่ถูกจำกัดมาผสมผสานกัน เรียกวิธีการนี้ว่าวิธีลิวที่ถูกจำกัด (Restricted Liu method) ซึ่งรูปแบบของตัวประมาณ คือ

$$\underset{\sim}{\beta}^*_d(d) = (X'X + I)^{-1}(X'X + dI)\,\underset{\sim}{\beta}^*$$

ตัวประมาณใหม่นี้ได้ปรับปรุงตัวประมาณวิธีกำลังสองน้อยสุดที่ถูกจำกัดโดยใช้หลักการของตัวประมาณลิว ทำให้ตัวประมาณที่ได้มีรูปแบบคล้ายตัวประมาณลิวโดยตัวประมาณที่ได้จะมีประสิทธิภาพและความแม่นยำมากขึ้น โดยตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณวิธีนี้มีคุณสมบัติคือ เป็นตัวประมาณเอนเอียงและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าวิธีกำลังสองน้อยสุดที่ถูกจำกัดและวิธีลิวในกรณีที่ข้อจำกัดนั้นเป็นจริง นอกจากนี้ได้กำหนดเงื่อนไขของค่า d ที่ทำให้วิธีลิวที่ถูกจำกัดมีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าวิธีกำลังสองน้อยสุดที่ถูกจำกัดและวิธีลิวในกรณีที่ข้อจำกัดนั้นไม่เป็นจริง

จากที่กล่าวมาทั้งหมดข้างต้นผู้วิจัยจึงสนใจเปรียบเทียบวิธีประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณ 4 วิธีคือ วิธีกำลังสองน้อยสุด วิธีกำลังสองน้อยสุดที่ถูกจำกัด วิธีริดจ์รีเกรสชันที่ถูก

จำกัด และวิธีลิวที่ถูกจำกัด เพื่อศึกษาว่าการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีใดจะให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองน้อยที่สุดเมื่อเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ

วัตถุประสงค์ของการวิจัย

การวิจัยนี้มีวัตถุประสงค์ที่จะเปรียบเทียบการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณเมื่อเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ โดยวิธีกำลังสองน้อยสุด (Ordinary Least Squares (OLS) method) วิธีกำลังน้อยสุดที่ถูกจำกัด (Restricted Least Squares (RLS) method) วิธีริดจ์รีเกรสชันที่ถูกจำกัด (Restricted Ridge Regression (RRR) method) และวิธีลิวที่ถูกจำกัด (Restricted Liu (RL) method) เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติในกรณีที่ข้อจำกัดเป็นจริงและไม่เป็นจริง

สมมติฐานของการวิจัย

เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติและตัวแปรอิสระมีพหุสัมพันธ์กันมาก ๆ วิธี RRR และวิธี RL จะให้ค่าประมาณสัมประสิทธิ์การถดถอยพหุคูณที่มีความถูกต้องมากกว่าวิธี RLS และวิธี OLS ภายใต้ขนาดตัวอย่าง จำนวนตัวแปรอิสระ ส่วนเบี่ยงเบนมาตรฐาน และความคลาดเคลื่อนของข้อจำกัดเดียวกัน

ขอบเขตของการวิจัย

1. การแจกแจงของความคลาดเคลื่อนจะทำการศึกษาในกรณีที่ความคลาดเคลื่อนมีการแจกแจงปกติ (normal distribution) คือ $\underset{\sim}{\varepsilon} \sim N_n(\underset{\sim}{0}, \sigma^2 I_n)$ โดยศึกษาที่พารามิเตอร์ค่าเฉลี่ย (μ) มีค่าเท่ากับ 0 ความแปรปรวน (σ^2) เท่ากับ 1 , 9 และ 25 ตามลำดับ
2. จำนวนตัวแปรอิสระที่ใช้ในการศึกษามี 2 ระดับ คือ 3 และ 5
3. ขนาดตัวอย่างที่ศึกษามี 3 ระดับ คือ 30 , 50 และ 100 ตามลำดับ
4. ความคลาดเคลื่อนของข้อจำกัดที่ใช้ในการศึกษาคือ 5% , 10% และ 15% ตามลำดับ[1]

---

[1] **ความคลาดเคลื่อนของข้อจำกัด**หมายความว่า ผลรวมเชิงเส้นของข้อจำกัดมีความคลาดเคลื่อนไปจากผลรวมเชิงเส้นของค่าพารามิเตอร์จริงกี่เปอร์เซ็นต์ เช่น ข้อจำกัด $\beta_1 + \beta_2 + \beta_3 = 3$ เมื่อพิจารณาผลรวมเชิงเส้นของค่าพารามิเตอร์จริงเป็น 3.15 จึงทำให้ข้อจำกัดมีความคลาดเคลื่อน 5%

5. ค่าสัมประสิทธิ์การถดถอยพหุคูณที่ใช้ในการศึกษาจะกำหนดให้เป็นค่าคงที่ใด ๆ ชุดหนึ่งเพื่อสร้างค่า $\underset{\sim}{y}$ ขึ้นจากตัวแบบ $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ เมื่อ $X \sim N_n(\underset{\sim}{0}, I)$

6. ระดับพหุสัมพันธ์ของตัวแปรอิสระ แบ่งเป็น 3 ระดับคือ

ระดับต่ำ ค่า ρ มีค่าอยู่ในช่วง 0.1 ถึง 0.3

ระดับปานกลาง ค่า ρ มีค่าอยู่ในช่วง 0.4 ถึง 0.6

ระดับสูง ค่า ρ มีค่าอยู่ในช่วง 0.7 ถึง 0.9

6.1 <u>กรณีตัวแปรอิสระเท่ากับ 3</u> ระดับพหุสัมพันธ์ของตัวแปรอิสระที่ศึกษาคือ

ระดับต่ำ ρ = (0.1 , 0.2 , 0.3)

ระดับปานกลาง ρ = (0.4 , 0.5 , 0.6)

ระดับสูง ρ = (0.7 , 0.8 , 0.9)

โดย ค่า ρ ในวงเล็บ คือ ความสัมพันธ์ระหว่าง X_1 กับ X_2 , X_1 กับ X_3 และ X_2 กับ X_3 ตามลำดับ

6.2 <u>กรณีตัวแปรอิสระเท่ากับ 5</u> ระดับพหุสัมพันธ์ของตัวแปรอิสระที่ศึกษาคือ

ระดับต่ำ ρ = (0.1 , 0.2 , 0.3 , 0.3)

ระดับปานกลาง ρ = (0.4 , 0.5 , 0.6 , 0.6)

ระดับสูง ρ = (0.7 , 0.8 , 0.9 , 0.9)

โดย ค่า ρ ในวงเล็บ คือ ความสัมพันธ์ระหว่าง X_1 กับ X_2 , X_1 กับ X_3 , X_2 กับ X_3 และ X_4 กับ X_5 ตามลำดับ

เกณฑ์ที่ใช้ในการตัดสินใจ

เกณฑ์ที่ใช้ในการตัดสินใจว่า วิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณวิธีการใดให้ค่าพารามิเตอร์ใกล้เคียงกับค่าจริงมากที่สุดนั้นจะพิจารณาจากค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (AMSE) และส่วนประกอบที่ใช้ในการพิจารณาจะใช้ค่าเปอร์เซ็นต์ของอัตราส่วนผลต่างค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (DIFF) ของการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีต่าง ๆ

$$MSE_j = \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{\beta}_{ij} - \beta_j\right)^2$$

$$AMSE = \frac{1}{p+1}\sum_{j=1}^{p+1} MSE_j$$

$$DIFF = \left[\frac{AMSE_{(i)} - AMSE_{(\min)}}{AMSE_{(\min)}} \right] \times 100 \quad , i = 1,2,3,4$$

เมื่อ β_j แทนตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณตัวที่ j

$\hat{\beta}_{ij}$ แทนตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณตัวที่ j จากการประมาณครั้งที่ i

MSE_j แทนค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าประมาณสำหรับสัมประสิทธิ์การถดถอย β_j

$AMSE$ แทนค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองเฉลี่ยของสัมประสิทธิ์การถดถอยพหุคูณ

$DIFF$ แทนเปอร์เซ็นต์ของอัตราส่วนผลต่างค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีต่างๆ

$AMSE_{\min}$ แทนค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของสัมประสิทธิ์การถดถอยพหุคูณเฉลี่ยที่มีค่าน้อยที่สุด

และ $AMSE_i$ แทนค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของสัมประสิทธิ์การถดถอยพหุคูณเฉลี่ยแต่ละวิธี

วิธีดำเนินการวิจัย

1. สร้างข้อมูลของความคลาดเคลื่อนตามที่กำหนดไว้ในขอบเขตการวิจัย

2. สร้างข้อมูลตัวแปรอิสระ (X) ให้มีระดับพหุสัมพันธ์ที่กำหนดไว้ในขอบเขตการวิจัย และสร้างข้อมูลของตัวแปรตาม (y) จากรูปแบบความสัมพันธ์ $\underset{\sim}{y} = X \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$

3. ประมาณค่าพารามิเตอร์ตามวิธีกำลังสองน้อยสุด วิธีกำลังสองน้อยสุดที่ถูกจำกัด วิธีริดจ์รีเกรสชันที่ถูกจำกัด และวิธีลิวที่ถูกจำกัด

4. คำนวณหาค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง $(AMSE)$ ของแต่ละวิธี

5. เปรียบเทียบเปอร์เซ็นต์ของอัตราส่วนผลต่างค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (DIFF) กับแต่ละวิธี

6. สรุปผลการวิจัยและข้อเสนอแนะ

ประโยชน์ที่คาดว่าจะได้รับ

1. ผลการศึกษาจะเป็นแนวทางในการเลือกใช้วิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณ ในกรณีที่ตัวแปรอิสระมีพหุสัมพันธ์กัน เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ

2. ผลการศึกษาจะเป็นแนวทางในการแก้ไขและปรับปรุงวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณ ในกรณีที่ตัวแปรอิสระมีพหุสัมพันธ์กันได้ต่อไปในอนาคต

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย