รายการอ้างอิง

ภาษาไทย

จะเด็จ สวรรค์ตรานนท์. การเปรียบเทียบวิธีที่ใช้สำหรับการเลือกสมการถดถอยที่ดีที่สุด. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2530.

จิตติมา ผสมญาติ. การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์ เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2546.

ทรงศิริ แต้สมบัติ. การวิเคราะห์การถดถอย. กรุงเทพมหานคร : มหาวิทยาลัยเกษตรศาสตร์, 2541.

ธีระพร วีระถาวร. การอนุมานเชิงสถิติขั้นกลาง : โครงสร้างและความหมาย. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์, 2536.

ธีระพร วีระถาวร. ความน่าจะเป็นกับการประยุกต์. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : วิทยพัฒน์, 2539.

ธีระพร วีระถาวร. ตัวแบบเชิงเส้น ทฤษฎีและการประยุกต์. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร : วิทยพัฒน์, 2541.

นิทัสน์ สุขสุวรรณ. การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดภายใต้แนวทางของเบส์ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ. วิทยานิพนธ์ปริญญามหาบัณฑิต. ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2545.

มนตรี พิริยะกุล. เทคนิคการวิเคราะห์สมการถดถอย เล่ม 1. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : ศรีเมืองการพิมพ์, 2532.

ภาษาอังกฤษ

Adrian E. Raftery, David Madigan and Jennifer A. Hoeting. Bayesian Model Averaging for Linear Regression Models. Journal of the American Statistical Association 92 , 437(1997) : 179-191.

Barbieri M. M. and Berger J. O. Optimal predictive model selection. Technical Report 02-02 (2003).

George E. I. and Mcculloch R. E. Variable Selection via Gibbs Sampling. Journal of the American Statistical Association 88, 423(1993) : 881-889.

Jose B. and Smith A.F.M. Bayesian Theory. New York : John Wiley & Sons, 1994.

ภาคผนวก

ทฤษฎีของเบส์ (Bayes' Theory)[1]

ให้ $\underset{\sim}{X}$ มีฟังก์ชันความหนาแน่นร่วม $p(\underset{\sim}{x};\underset{\sim}{\theta})$ ซึ่งอาจจะเขียนอยู่ในรูปของฟังก์ชันความหนาแน่นที่มีเงื่อนไขเมื่อกำหนด $\underset{\sim}{\theta}$ กล่าวคือ $p(\underset{\sim}{x}\mid\underset{\sim}{\theta})$

และให้ $\underset{\sim}{\theta}$ เป็นเวกเตอร์สุ่มซึ่งมีฟังก์ชันความหนาแน่นร่วม $\pi(\underset{\sim}{\theta})$ ซึ่งเราเรียกว่าฟังก์ชันความหนาแน่นโดยหลักเกณฑ์ (prior density function) ฟังก์ชันความหนาแน่นโดยหลักเกณฑ์เป็นความเชื่อหรือข้อมูลเกี่ยวกับพารามิเตอร์ก่อนทำการทดลอง เมื่อทำการทดลองหรือสุ่มตัวอย่าง กล่าวคือ $\underset{\sim}{X}=\underset{\sim}{x}$ จะได้ว่าฟังก์ชันความหนาแน่นที่มีเงื่อนไขของ $\underset{\sim}{\theta}$ เมื่อกำหนด $\underset{\sim}{X}=\underset{\sim}{x}$ อยู่ในรูปของ

$$\pi(\underset{\sim}{\theta}\mid\underset{\sim}{x})=\begin{cases}\dfrac{p(\underset{\sim}{x}\mid\underset{\sim}{\theta})\pi(\underset{\sim}{\theta})}{\sum\limits_{\underset{\sim}{\theta}} p(\underset{\sim}{x}\mid\underset{\sim}{\theta})\pi(\underset{\sim}{\theta})}, & \text{ในกรณีไม่ต่อเนื่อง}\\[2ex] \dfrac{p(\underset{\sim}{x}\mid\underset{\sim}{\theta})\pi(\underset{\sim}{\theta})}{\int\limits_{\Theta} p(\underset{\sim}{x}\mid\underset{\sim}{\theta})\pi(\underset{\sim}{\theta})d\underset{\sim}{\theta}}, & \text{ในกรณีต่อเนื่อง}\end{cases}$$

และเราเรียก $\pi(\underset{\sim}{\theta}\mid\underset{\sim}{x})$ ว่าเป็น ฟังก์ชันความหนาแน่นโดยประสบการณ์ (posterior density function) ซึ่งเป็นข้อมูลเกี่ยวกับพารามิเตอร์หลังจากทำการทดลองหรือสุ่มตัวอย่าง เราอาจเขียนได้อีกลักษณะ กล่าวคือ

ฟังก์ชันความหนาแน่นโดยประสบการณ์ α ฟังก์ชันภาวะน่าจะเป็น $\times$ ฟังก์ชันความหนาแน่นโดยหลักเกณฑ์

หรือ $Posterior\ \alpha\ Likelihood \times \Pr ior$

จะเห็นได้ว่าฟังก์ชันการแจกแจงโดยประสบการณ์เป็นสัดส่วนกับผลคูณระหว่างฟังก์ชันภาวะน่าจะเป็นและฟังก์ชันการแจกแจงโดยหลักเกณฑ์ จึงเป็นไปได้ว่าฟังก์ชันการแจกแจงโดยประสบการณ์และฟังก์ชันการแจกแจงโดยหลักเกณฑ์เป็นอาจจะมีรูปแบบเดียวกัน

<u>นิยามที่ 2.9.1</u>

ให้ $X_i,\ i=1,2,\ldots,n$ จ.ม.อ. โดยมีฟังก์ชันความหนาแน่น $f(\underset{\sim}{x};\theta)$ และฟังก์ชันภาวะน่าจะเป็นเขียนได้ในรูปของ $p(\underset{\sim}{x}\mid\underset{\sim}{\theta})$ ถ้ามีวงศ์ของการแจกแจงสำหรับ $\underset{\sim}{\theta}$ ซึ่งทำให้การแจกแจงโดยหลักเกณฑ์และโดยประสบการณ์อยู่ในวงศ์เดียวกัน เราเรียกวงศ์การแจกแจงนั้นว่า วงศ์คู่สังยุค (conjugate family)

---

[1] ธีระพร วีระถาวร, <u>การอนุมานเชิงสถิติขั้นกลาง : โครงสร้างและความหมาย</u>, พิมพ์ครั้งที่ 2 (กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์, 2536), หน้า 119-120.

สำหรับรายละเอียดขั้นตอนของวิธีการที่ใช้ในการคัดเลือกตัวแบบการถดถอยเชิงเส้นพหุคูณแต่ละวิธี พร้อมทั้งตัวอย่างแสดงดังต่อไปนี้

1. วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging Method (BMA)) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลเมื่อใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo Model Composition (MC^3))(BMA_{MC^3}) มีขั้นตอนดังนี้

1.1) เริ่มจากการนำเทคนิคมอนติคาร์โลโดยอาศัยโครงสร้างของลูกโซ่มาร์คอฟมาใช้เพื่อหาปริภูมิตัวแบบ ในวิธีนี้ได้ใช้การสุ่มแบบกิบส์ซึ่งเป็นกรณีหนึ่งของโครงสร้างแบบลูกโซ่มาร์คอฟเพื่อสร้างลำดับ ซึ่งลำดับนี้จะสอดคล้องกับการคัดเลือกตัวแปรและการสุ่มค่าพารามิเตอร์สัมประสิทธิ์การถดถอย โดยการกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยจะขึ้นอยู่กับการกำหนดค่า $\frac{\sigma_{\beta_i}}{\tau_i}$ เมื่อกำหนดค่า $\frac{\sigma_{\beta_i}}{\tau_i}$ ให้มีค่ามาก การกระจายของพารามิเตอร์ก็จะมากขึ้นจึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลงและจำนวนของตัวแบบที่ไม่แน่นอนก็จะเกิดขึ้นมากด้วย หลังจากเสร็จสิ้นกระบวนการนี้จะปรากฏตัวแบบทุกตัวแบบที่อยู่ในปริภูมิตัวแบบ

1.2) คำนวณค่าความน่าจะเป็นภายหลังของตัวแบบทุกตัวแบบที่อยู่ในปริภูมิตัวแบบ ซึ่งการหาความน่าจะเป็นภายหลังเกิดจากความน่าจะเป็นก่อนรวมกับฟังก์ชันความควรจะเป็น โดยมีการแจกแจงความน่าจะเป็นก่อนของสัมประสิทธิ์การถดถอยเป็นการแจกแจงแบบแกมมา ซึ่งการแจกแจงความน่าจะเป็นภายหลังของแต่ละตัวแบบก็จะมีการแจกแจงแบบแกมมา

ความน่าจะเป็นภายหลังของตัวแบบ M_k คือ

$$p(M_k \mid D) = \frac{p(D \mid M_k) \cdot p(M_k)}{\sum_{l=1}^{K} p(D \mid M_l) \cdot p(M_l)}$$

โดยที่การแจกแจงภายหลังของตัวแบบการถดถอย M_k คือ $p(D \mid M_k)$ ซึ่งมีการแจกแจงแบบแกมมา

1.3) คำนวณค่าพยากรณ์ของตัวแบบในทุกรูปแบบที่ปรากฏอยู่ในปริภูมิตัวแบบ

1.4) นำค่าพยากรณ์ของตัวแบบในทุกรูปแบบที่ปรากฏอยู่ในปริภูมิตัวแบบมาเฉลี่ยกันโดยใช้ความน่าจะเป็นภายหลังของแต่ละตัวแบบเป็นตัวถ่วงน้ำหนัก ค่าเฉลี่ยของค่าพยากรณ์จะเป็นดังนี้

$$\hat{y} = \sum_{k=1}^{K} \hat{y}_k p(M_k \mid D)$$

จากสมการจะทำให้การเฉลี่ยค่าพยากรณ์ของแต่ละตัวแบบ M_k จะทำให้เราได้ค่าพยากรณ์ $\hat{y}$ ตามที่เหมาะสมซึ่งเกิดจากการเฉลี่ยตัวแบบทุกตัวแบบที่อยู่ในปริภูมิตัวแบบ

ตัวอย่าง 1 กำหนดตัวแปรอิสระจำนวน 4 ตัวแปร คือ X_1, X_2, X_3, X_4 โดยสุ่มตัวอย่างที่มีขนาดเท่ากับ n จะได้จำนวนตัวแบบทั้งหมดจำนวน 16 แบบ ดังนี้

ตัวแบบ	ค่าความน่าจะเป็นภายหลังของตัวแบบ $p(M_k \mid D)$	ค่าพยากรณ์ $\hat{y}_k$
$y_i = \beta_0 + \varepsilon_i$	$p(M_1 \mid D)$	$\hat{y}_1$
$y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon_i$	$p(M_2 \mid D)$	$\hat{y}_2$
$y_i = \beta_0 + \beta_2 x_{2i} + \varepsilon_i$	$p(M_3 \mid D)$	$\hat{y}_3$
$y_i = \beta_0 + \beta_3 x_{3i} + \varepsilon_i$	$p(M_4 \mid D)$	$\hat{y}_4$
$y_i = \beta_0 + \beta_4 x_{4i} + \varepsilon_i$	$p(M_5 \mid D)$	$\hat{y}_5$
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$	$p(M_6 \mid D)$	$\hat{y}_6$
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_3 x_{3i} + \varepsilon_i$	$p(M_7 \mid D)$	$\hat{y}_7$
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \varepsilon_i$	$p(M_8 \mid D)$	$\hat{y}_8$
$y_i = \beta_0 + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$	$p(M_9 \mid D)$	$\hat{y}_9$
$y_i = \beta_0 + \beta_2 x_{2i} + \beta_4 x_{4i} + \varepsilon_i$	$p(M_{10} \mid D)$	$\hat{y}_{10}$
$y_i = \beta_0 + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	$p(M_{11} \mid D)$	$\hat{y}_{11}$
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$	$p(M_{12} \mid D)$	$\hat{y}_{12}$
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_4 x_{4i} + \varepsilon_i$	$p(M_{13} \mid D)$	$\hat{y}_{13}$
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	$p(M_{14} \mid D)$	$\hat{y}_{14}$
$y_i = \beta_0 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	$p(M_{15} \mid D)$	$\hat{y}_{15}$
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	$p(M_{16} \mid D)$	$\hat{y}_{16}$

ตารางที่ 1 แสดงตัวแบบทุกตัวในปริภูมิตัวแบบ ค่าความน่าจะเป็นภายหลังของแต่ละตัวแบบ และค่าพยากรณ์ของแต่ละตัวแบบ

จากตารางที่ 1 เมื่อนำค่าพยากรณ์ของแต่ละตัวแบบถ่วงน้ำหนักด้วยความน่าจะเป็นภายหลังของแต่ละตัวแบบ M_k ดังสมการต่อไปนี้

$$\hat{y} = \sum_{k=1}^{16} \hat{y}_k \cdot p(M_k \mid D)$$

ดังนั้นค่าพยากรณ์ที่เหมาะสมได้จากวิธีการนี้ คือ $\hat{y}$ ซึ่งเกิดจากการเฉลี่ยตัวแบบทุกตัวแบบที่อยู่ในปริภูมิตัวแบบ

2. วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (Optimal Predictive Model Selection : median probability model (OPM)) มีขั้นตอนดังนี้

2.1) ค้นหาตัวแบบด้วยเทคนิคลูกโซ่มาร์คอฟเชนมอนติคาร์โลโดยให้เคลื่อนไประหว่างตัวแบบด้วยค่าความน่าจะเป็นภายหลังของตัวแบบนั้น ๆ ซึ่งคำนวณค่าความน่าจะเป็นภายหลังได้จากสัดส่วนของจำนวนครั้งที่ตัวแบบอยู่ในลูกโซ่

2.2) ค้นหาตัวแบบที่มีความน่าจะเป็นภายหลังกลางๆ โดยจะพิจารณาตัวแปรอิสระที่มีผลรวมของความน่าจะเป็นภายหลังสำหรับตัวแปรอิสระที่ j ดังสมการ

$$p_j = \sum_{\underset{\sim}{l}:l_j=1} P\left(M_{\underset{\sim}{l}} \mid \underset{\sim}{y} \right)$$

เมื่อ p_j จะมีค่าอย่างน้อยเท่ากับ 0.5 เป็นเกณฑ์ในการเลือกตัวแปรอิสระที่จะอยู่ในตัวแบบ ซึ่งจะถือว่าตัวแบบที่ประกอบด้วยตัวแปรอิสระที่มีผลรวมของความน่าจะเป็นภายหลังอย่างน้อยเท่ากับ 0.5 เป็นตัวแบบที่มีความน่าจะเป็นภายหลังกลางๆ

2.3) ทำการคัดเลือกตัวแบบที่เหมาะสมที่สุดจากกลุ่มของตัวแบบที่มีความน่าจะเป็นภายหลังกลางๆ และเกณฑ์ที่ใช้ในการคัดเลือกตัวแบบที่เหมาะสมที่สุดจากกลุ่มของตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ คือ ความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสองโดยจะเลือกตัวแบบที่มีค่าความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสองต่ำสุด

<u>ตัวอย่าง 2</u> กำหนดตัวแปรอิสระจำนวน 4 ตัวแปร คือ X_1, X_2, X_3, X_4 โดยสุ่มตัวอย่างที่มีขนาดเท่ากับ n จะได้จำนวนตัวแบบที่เป็นไปได้จำนวน 16 ตัวแบบ ดังนี้

ตัวแบบ	ความน่าจะเป็นภายหลัง $P(M_l \mid y)$	ค่าความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสอง $R(M_l)$
$y_i = \beta_0 + \varepsilon_i$	0.000003	2652.44
$y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon_i$	0.000012	1207.04
$y_i = \beta_0 + \beta_2 x_{2i} + \varepsilon_i$	0.000026	854.85
$y_i = \beta_0 + \beta_3 x_{3i} + \varepsilon_i$	0.000002	1864.41
$y_i = \beta_0 + \beta_4 x_{4i} + \varepsilon_i$	0.000058	838.20
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$	0.275484	8.19
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_3 x_{3i} + \varepsilon_i$	0.000006	1174.14
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \varepsilon_i$	0.107798	29.73
$y_i = \beta_0 + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$	0.000229	353.72
$y_i = \beta_0 + \beta_2 x_{2i} + \beta_4 x_{4i} + \varepsilon_i$	0.000018	821.15
$y_i = \beta_0 + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	0.003785	118.59
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$	0.170990	1.21
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_4 x_{4i} + \varepsilon_i$	0.190720	0.18
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	0.159959	1.71
$y_i = \beta_0 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	0.041323	20.42
$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$	0.049587	0.47

<u>ตารางที่ 2</u> แสดงตัวแบบทุกตัวในปริภูมิตัวแบบ ค่าความน่าจะเป็นภายหลังของตัวแปรอิสระ และค่าความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสอง

จากตารางที่ 2 นำค่าความน่าจะเป็นภายหลัง $P(M_l \mid y)$ มาคำนวณผลรวมค่าความน่าจะเป็นภายหลังของตัวแปรอิสระแต่ละตัว ดังต่อไปนี้

$$p_1 = \sum_{l: l_1 = 1} P\left(M_l \mid y\right) = 0.954556$$

$$p_2 = \sum_{l: l_2 = 1} P\left(M_l \mid y\right) = 0.728377$$

$$p_3 = \sum_{\underset{\sim}{l}: l_3 = 1} P\left(M_{\underset{\sim}{l}} \mid \underset{\sim}{y}\right) = 0.425881$$

$$p_4 = \sum_{\underset{\sim}{l}: l_4 = 1} P\left(M_{\underset{\sim}{l}} \mid \underset{\sim}{y}\right) = 0.553248$$

เนื่องจากผลรวมค่าความน่าจะเป็นภายหลังของตัวแปรอิสระ X_1, X_2 และ X_4 มีค่ามากกว่า 0.5 ดังนั้นกลุ่มของตัวแบบที่มีความน่าจะเป็นภายหลังกลางๆ คือ ตัวแบบที่ประกอบด้วยตัวแปรอิสระ X_1, X_2 และ X_4 นั่นคือตัวแบบที่ 2 , 3 ,5 , 6, 8 ,10 ,13 และ 16 ทำการคัดเลือกตัวแบบที่เหมาะสมที่สุดโดยใช้เกณฑ์ความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสองโดยจะเลือกตัวแบบที่มีค่าดังกล่าวต่ำสุด ซึ่งจากตารางจะพบว่าตัวแบบที่มีค่าความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสองต่ำสุด คือตัวแบบที่ 13 ซึ่งมีตัวแปรอิสระ X_1, X_2 และ X_4 อยู่ในตัวแบบและจะนำตัวแบบที่นี้มาใช้ในการหาค่าพยากรณ์ต่อไป

3. วิธีการถดถอยแบบขั้นบันได (Stepwise Regression Method (SR))

วิธีการถดถอยแบบขั้นบันไดนี้เป็นวิธีที่นิยมใช้กันอย่างแพร่หลายโดยเลือกตัวแปรอิสระเข้าในสมการถดถอยครั้งละหนึ่งตัวแปรตัวแปรอิสระใดที่อยู่ในสมการแล้วจะต้องมีการทดสอบว่าตัวแปรนั้นยังมีส่วนในการอธิบายความผันแปรของตัวแปรตาม ขณะที่มีตัวแปรอิสระอื่นอยู่ในสมการถดถอยหรือไม่ นั่นคือตัวแปรอิสระใดที่เข้าอยู่ในสมการถดถอยแล้วอาจจะถูกตัดออกภายหลังได้ โดยมีขั้นตอนดังต่อไปนี้

3.1) เลือกตัวแปรอิสระที่ให้ค่าสัมประสิทธิ์สหสัมพันธ์ที่สูงที่สุด

3.2) ทำการตรวจสอบนัยสำคัญของสัมประสิทธิ์การถดถอยดังกล่าว โดยตัวสถิติทดสอบ t หรือ F ถ้ายอมรับสมมติฐาน นั่นคือไม่พบนัยสำคัญที่ระดับนัยสำคัญ 0.05 ชี้ให้เห็นว่าตัวแปรอิสระนี้ไม่มีส่วนในการอธิบายความผันแปรของตัวแปรตาม แต่ถ้าปฏิเสธสมมติฐาน นั่นคือตัวแปรอิสระนี้มีนัยสำคัญที่ระดับนัยสำคัญ 0.05 ชี้ให้เห็นว่าตัวแปรอิสระนี้ อยู่ในสมการถดถอยได้ จะทำขั้นตอนต่อไปเพื่อหาตัวแปรอิสระใหม่เข้าในสมการถดถอย

3.3) เลือกตัวแปรอิสระที่ให้ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนสูงที่สุดแล้วทำการตรวจสอบความมีนัยสำคัญเทียบกับระดับนัยสำคัญ 0.05 ถ้าผลการทดสอบพบว่าไม่มีนัยสำคัญทางสถิติแล้ว ตัวแปรอิสระตัวนี้ก็จะไม่ถูกเลือกเข้าสมการแต่ถ้าผลการทดสอบพบนัยสำคัญทางสถิติ ก็จะเลือกตัวแปรอิสระเข้าสู่สมการ

3.4) จากสมการรูปแบบเต็ม ตรวจสอบนัยสำคัญของสัมประสิทธิ์การถดถอยแต่ละตัวที่อยู่ในสมการ โดยการทดสอบแบบ t บางส่วนหรือ F บางส่วน พิจารณาว่าตัวแปรอิสระตัวใดควรจะถูกตัดออกจากสมการ ถ้ายอมรับสมมติฐานแสดงว่าสัมประสิทธิ์การถดถอยไม่มีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.10 ชี้ให้เห็นว่าตัวแปรอิสระตัวนั้นไม่มีส่วนในการอธิบายความผันแปรของตัวแปรตาม ดังนั้นตัวแปรอิสระตัวนั้นจะถูกตัดออกจากสมการถดถอยและถ้าปฏิเสธสมมติฐานแสดงว่าสัมประสิทธิ์การถดถอยมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.10 ตัวแปรอิสระตัวนั้นจะอยู่ในสมการถดถอย

3.5) ทำซ้ำข้อ 3.3 – 3.4 กระบวนการเลือกแบบขั้นบันไดจะหยุดเมื่อไม่มีตัวแปรอิสระถูกนำเข้าสมการ หรือตัวแปรอิสระถูกตัดออกจากสมการได้

<u>ตัวอย่าง 3</u> กำหนดตัวแปรอิสระจำนวน 4 ตัวแปร คือ X_1, X_2, X_3, X_4 โดยสุ่มตัวอย่างที่มีขนาดเท่ากับ n โดยที่วิธีการถดถอยแบบขั้นบันไดไม่ได้ทำการหาตัวแบบทุกตัวที่เป็นไปได้ดังเช่นในตัวอย่าง 2 ตัวอย่างข้างต้น แต่จะใช้เพียงแค่ตัวแบบเดียวในการหาค่าพยากรณ์ ซึ่งวิธีการนี้จะมีขั้นตอน ดังนี้

1. เริ่มจากการเลือกตัวแปรอิสระตัวที่มีค่าสัมประสิทธิ์สหสัมพันธ์กับตัวแปรตาม y สูงสุด จากการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรแต่ละตัวพบว่า X_1 มีค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดกับ y

2. ทดสอบสัมประสิทธิ์การถดถอยของตัวแปรมีนัยสำคัญทางสถิติ (ที่ระดับนัยสำคัญ 0.05) จากการทดสอบพบว่าค่าสัมประสิทธิ์การถดถอยของ X_1 มีนัยสำคัญทางสถิติภายใต้ระดับนัยสำคัญ 0.05 นั่นคือ X_1 อยู่ในสมการถดถอย

3. เลือกตัวแปรอิสระเข้าสมการถดถอยโดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนที่สูงที่สุดแล้วตรวจสอบนัยสำคัญของสัมประสิทธิ์การถดถอยของตัวแปรที่นำเข้า ถ้าพบนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 จะเลือกตัวแปรอิสระตัวนั้นเข้าสมการ จากการพิจารณาข้างต้นพบว่า X_2 มีค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนกับ y สูงสุด และทดสอบสัมประสิทธิ์การถดถอยของตัวแปรมีนัยสำคัญทางสถิติ จากการทดสอบพบว่าค่าสัมประสิทธิ์การถดถอยของ X_2 มีนัยสำคัญทางสถิติภายใต้ระดับนัยสำคัญ 0.05 นั่นคือ X_2 อยู่ในสมการถดถอย

4. พิจารณาจากตัวแปรอิสระที่มีค่าสัมประสิทธิ์การถดถอยไม่มีนัยสำคัญทางสถิติ (ที่ระดับนัยสำคัญ 0.10) ตัวแปรอิสระตัวนั้นจะถูกตัดออกจากสมการถดถอย เมื่อทำการทดสอบสัมประสิทธิ์การถดถอยของตัวแปรมีนัยสำคัญทางสถิติ (ที่ระดับนัยสำคัญ 0.05) และพบว่าค่าสัมประสิทธิ์การถดถอยของ X_1 และ X_2 มีนัยสำคัญทางสถิติภายใต้ระดับนัยสำคัญ 0.05 ซึ่งหมายความว่า X_1 และ X_2 ยังอยู่ในสมการถดถอย

5. เลือกตัวแปรอิสระเข้าสมการถดถอยโดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนที่สูงที่สุดแล้วตรวจสอบนัยสำคัญของสัมประสิทธิ์การถดถอยของตัวแปรที่นำเข้า ถ้าพบนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 จะเลือกตัวแปรอิสระตัวนั้นเข้าสมการ จากการพิจารณาข้างต้นพบว่า X_4 มีค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนกับ y สูงสุด และทดสอบสัมประสิทธิ์การถดถอยของตัวแปรมีนัยสำคัญทางสถิติ (ที่ระดับนัยสำคัญ 0.05) จากการทดสอบพบว่าค่าสัมประสิทธิ์การถดถอยของ X_4 ไม่มีนัยสำคัญทางสถิติ ดังนั้นจึงไม่มีตัวแปรอิสระตัวใดถูกนำเข้าสมการทำให้วิธีการถดถอยแบบขั้นบันไดสิ้นสุดลง ตัวแบบถดถอยที่ได้อยู่ในรูปของ

$$\hat{y} = b_0 + b_1 X_1 + b_2 X_2$$

ดังนั้นค่าพยากรณ์ที่ได้จากวิธีการนี้ คือ $\hat{y}$ ซึ่งเกิดจากใช้ตัวแบบเพียงตัวแบบเดียว

รายละเอียดของโปรแกรมที่ใช้ในการวิจัย

ในการวิจัยครั้งนี้ได้ใช้โปรแกรม S-plus 2000 สำหรับการสร้างตัวแบบด้วยวิธี BMA_{MC^3} วิธี OPM และวิธี SR เนื่องจากในงานวิจัยของ ราฟเทอร์รี่ เมดิแกน และโฮเอ็ททิง (Raftery, Madigan and Hoeting, 1997) ซึ่งเป็นผู้นำเสนอวิธีการเฉลี่ยตัวแบบของเบส์ได้มีการนำเสนออัลกอริทึมของวิธีการเฉลี่ยตัวแบบของเบส์โดยใช้โปรแกรม S-plus 2000 ดังนั้นผู้วิจัยจึงได้ใช้โปรแกรม S-plus 2000 ในการเขียนโปรแกรมเพื่อสร้างตัวแบบด้วยวิธี BMA_{MC^3} วิธี OPM และวิธี SR ดังกล่าว

สำหรับรายละเอียดทั้งหมดของโปรแกรมที่ใช้ในการวิจัยมีดังนี้

การทำงานของโปรแกรมที่ใช้ในการวิจัย สำหรับการสร้างตัวแบบด้วยวิธี BMA_{MC^3} โดยใช้โปรแกรม S-plus 2000

โปรแกรมที่ใช้ในการสร้างตัวแบบด้วยวิธี BMA_{MC^3} นั้นสร้างขึ้นโดยใช้โปรแกรม S-plus 2000 โดยรายละเอียดของโปรแกรมมีดังนี้

โปรแกรมสำหรับสร้างตัวแบบด้วยวิธี BMA_{MC^3}

```
BMA.MC3<<-function(all.y,all.x,num.its,MO.var,MO.out,outs.list ,PI,K,ip,cc,a,b)
{
        Ys<<-scale(all.y)
        Xs<<-scale(all.x)
        MO.var<<-MO.var
        MO.out<<-MO.out
        outs.list<<-outs.list
        PI<<-PI
        K<<-K
        ip<<-ip
        cc<<-cc
        a<<-a
        b<<-b
        flag<<-1
        outcnt<<-sum(outs.list)
        big.list<<-matrix(0,1,4)
big.list[1,1]<<-sum(2^((0:(length(MO.var)-1))[MO.var]))+1
if (sum(MO.out)!=0)
big.list[1,2]<<-sum(2^((0:(length(MO.out)-1))[MO.out]))+1
else big.list[1,2]<<-1
if (outcnt!=0) big.list[1,3]<<-(dim(Ys)[1]-sum(MO.out))*log(1-PI) +
sum(MO.out)*log(PI)+MC3.REG.logpost(Ys,Xs,MO.var,sum(MO.var),ip,cc,a,b)
else big.list[1,3]<<-MC3.REG.logpost(Ys,Xs,MO.var,sum(MO.var),ip,cc,a,b)
i<-1
while (i<=num.its)
{
```

```
if (flag= =1)
{
if (sum(MO.var)!=0)
MO.1<<-sum(2^((0:length(M0.var)-1))[MO.var]))+1
else MO.1<<-1
if (sum(MO.out)!=0)
MO.2<<-sum(2^((0:length(MO.out)-1))[MO.out]))+1
else MO.2<<-1
}
M1<-MC3.REG.choose(MO.var,MO.out)
if (sum(M1$var)!=0)
M1.1<<-sum(2^((0:(length(MO.var)-1))[M1$var]))+1
else M1.1<-1
if (sum(M1$out)!=0)
M1.2<<-sum(2^((0:(length(MO.out)-1))[M1$out]))+1
else M1.2<-1
if (sum(big.list[,1]= =M1.1&big.list[,2]= =m1.2)= =0)
{
        if (M1.1= =1)
        {
        if (outcnt!=0) a<-(dim(Ys)[1]-sum(M1$out))*log(1-PI) +
        sum(M1$out)*log(PI)+MC3.REG.logpost(Ys,Xs,0,0,ip,cc,a,b)
        else a<-MC3.REG.logpost(Ys,Xs,0,0,ip,cc,a,b)
        }
        else
        {
        if (outcnt!=0) a<-(dim(Ys)[1]-sum(M1$out))*log(1-PI)+
        sum(M1$out)*log(PI)+MC3.REG.logpost(Ys,Xs,M1$var,
        sum(M1$var),outs.list[M1$out],ip,cc,a,b)
        else a<-MC3.REG.logpost(Ys,Xs,M1$var,sum(M1$var),
        outs.list[M1$out],ip,cc,a,b)
        }
```

```
big.list<<-rbind(big.list,c(M1.1,M1.2,a,0))
}
BF<-exp(big.list[big.list[,1]==M1.1&big.list[,2]==m1.2,3]-
big.list[big.list[,1]==M0.1&big.list[,2]==m0.2,3])
if (BF>=1) flag<<-1
else flag<<-rbinom(1,1,BF)
if (flag= =1)
{
MO.var<<-M1$var
MO.out<<-M1$out
MO.1<<-M1.1
MO.2<<-M1.2
}
big.list[big.list[,1]= =M0.1&big.list[,2]= =m0.2,4]<<- big.list[big.list[,1]= =M0.1&big.list[,2]=
=m0.2,4]+1
i<-i+1
}
var.vect<<-matrix(as.logical(rep(big.list[,1]1,rep(length(MO.var),length(big.list[,1])))%/%
2^(0:(length(MO.var)-1))%%2),ncol=length(MO.var),byrow=T)
nvar<-length(MO.var)
ndx<-1:n.var
Xn<-rep("X",n.var)
labs<-paste(Xn,ndx,sep="")
dimnames(var.vect)<-list(c(1:length(var.vect[,1])),labs)
postprob<<-matrix((exp(big.list[,3]))/(sum(exp(big.list[,3]))),ncol=1)
dimnames(postprob)[2]_list(c('Post.Mod.Pr."))
visits<<-matrix(big.list[,4],ncol=1)
dimnames(visits[2]<-list(c("#visits"))
if (length(out.list)!=0)
{
```

```
        out.vect<<-matrix(as.logical(rep(big.list[,2]-
        1,rep(length(outs.list),length(big.list[,2])))%/%2^(0:(length(outs.list)-
        1))%%2),ncol=length(outs.list),byrow=T)
        dimnames(out.vect)<-list(c(1:length(out.vect[,1])),c(outs.list))
        model.matrix<<-cbind(var.vect,out.vect,postprob,visits)}
        else model.matrix<<-cbind(var.vect,postprob,visits)}
        colno<-length(MO.var)+length(MO.out)+1
        model.matrix<<- model.matrix[order(-model.matrix[,colon]),]
        return(model.matrix)
}
MC3.REG.choose<-function(MO.var,MO.out)
{
        var<-MO.var
        in.or.out<-sample(c(1:length(MO.var),rep(0,length(MO.out))),1)
        if(in.or.out==0)
        {
                out<-MO.out
                in.or.out2<-sample(1:length(MO.out),1)
                out[in.or.out2]<-!MO.out[in.or.out2]
        }
        else
        {
                var[in.or.out]<-!MO.var[in.or.out]
                out<-MO.out
        }
        return(var,out)
}
MC3.REG.logpost<-function(x,y,a,b)
{
        a<<-a
        b<<-b
        x_x[,model.vect]
```

```
        xtx_t(x)%*%x
        xty_t(x)%*%y
        beta_ginverse(xtx)%*%xty
        beta_rep(1,numx+1)
}
for (i in 1 : numx)
        {
                post_(exp(inverse(beta*(b+numx)))*(beta*(a+(sum(y)[i])-1)
        }
        logpost_log(post,10)
        return(logpost)
}
numx_numx
samplesize_samplesize
sde_sde
numloop_500
x_matrix(scan("x.txt"),ncol=1,byrow=T)
msemc3_matrix(nrow=numloop,ncol=1)
fy_function(samplesize,sde,x,numx)
{
        error_ matrix(scan("error.txt"),ncol=1,byrow=T)
        ones_rep(1,samplesize)
        xones_cbind(ones,x)
        beta_rep(1,numx+1)
        y_(xones%*%beta)+error
        return(y)
}
for(i in 1:numloop)
{
        y_fy(samplesize,sde,x,numx)
        resultmc3_BMA.MC3(y,x,10000,rep(T,numx),NULL,NULL,0,0,ip,cc,a,b)
        rowresult_nrow(resultmc3)
```

```
        msemc3[i]_sum(regmodel$residuals^2)/regmodel$df
}
amsemc3_sum(msemc3)/numloop
print(amsemc3)
stdmc3_matrix(nrow=numloop,ncol=1)
for(indexi in 1:numloop)
{
        stdmc3[indexi]_(msemc3[indexi]-amsemc3)^2
}
stdamsemc3_sqrt(sum(stdmc3)/(numloop-1))
print(stdamsemc3)
```

การทำงานของโปรแกรมที่ใช้ในการวิจัย สำหรับการสร้างตัวแบบด้วยวิธี OPM โดยใช้โปรแกรม S-plus 2000

โปรแกรมที่ใช้ในการสร้างตัวแบบด้วยวิธี OPM นั้นสร้างขึ้นโดยใช้โปรแกรม S-plus 2000 โดยรายละเอียดของโปรแกรมมีดังนี้

โปรแกรมสำหรับสร้างตัวแบบด้วยวิธี OPM

```
BMA.MC3<<-function(all.y,all.x,num.its,MO.var,MO.out,outs.list ,PI,K,ip,cc,a,b)
{
        Ys<<-scale(all.y)
        Xs<<-scale(all.x)
        MO.var<<-MO.var
        MO.out<<-MO.out
        outs.list<<-outs.list
        PI<<-PI
        K<<-K
        ip<<-ip
        cc<<-cc
        a<<-a
        b<<-b
        flag<<-1
        outcnt<<-sum(outs.list)
        big.list<<-matrix(0,1,4)
big.list[1,1]<<-sum(2^((0:(length(MO.var)-1))[MO.var]))+1
if (sum(MO.out)!=0)
big.list[1,2]<<-sum(2^((0:(length(MO.out)-1))[MO.out]))+1
else big.list[1,2]<<-1
if (outcnt!=0) big.list[1,3]<<-(dim(Ys)[1]-sum(MO.out))*log(1-PI) +
sum(MO.out)*log(PI)+MC3.REG.logpost(Ys,Xs,MO.var,sum(MO.var),ip,cc,a,b)
else big.list[1,3]<<-MC3.REG.logpost(Ys,Xs,MO.var,sum(MO.var),ip,cc,a,b)
i<-1
while (i<=num.its)
{
```

```
if (flag= =1)
{
if (sum(MO.var)!=0)
MO.1<<-sum(2^((0:length(M0.var)-1))[MO.var]))+1
else MO.1<<-1
if (sum(MO.out)!=0)
MO.2<<-sum(2^((0:length(MO.out)-1))[MO.out]))+1
else MO.2<<-1
}
M1<-MC3.REG.choose(MO.var,MO.out)
if (sum(M1$var)!=0)
M1.1<<-sum(2^((0:(length(MO.var)-1))[M1$var]))+1
else M1.1<-1
if (sum(M1$out)!=0)
M1.2<<-sum(2^((0:(length(MO.out)-1))[M1$out]))+1
else M1.2<-1
if (sum(big.list[,1]= =M1.1&big.list[,2]= =m1.2)= =0)
{
        if (M1.1= =1)
        {
        if (outcnt!=0) a<-(dim(Ys)[1]-sum(M1$out))*log(1-PI) +
        sum(M1$out)*log(PI)+MC3.REG.logpost(Ys,Xs,0,0,ip,cc,a,b)
        else a<-MC3.REG.logpost(Ys,Xs,0,0,ip,cc,a,b)
        }
        else
        {
        if (outcnt!=0) a<-(dim(Ys)[1]-sum(M1$out))*log(1-PI)+
        sum(M1$out)*log(PI)+MC3.REG.logpost(Ys,Xs,M1$var,
        sum(M1$var),outs.list[M1$out],ip,cc,a,b)
        else a<-MC3.REG.logpost(Ys,Xs,M1$var,sum(M1$var),
        outs.list[M1$out],ip,cc,a,b)
        }
```

```
            big.list<<-rbind(big.list,c(M1.1,M1.2,a,0))
        }
        BF<-exp(big.list[big.list[,1]==M1.1&big.list[,2]==m1.2,3]-
        big.list[big.list[,1]==M0.1&big.list[,2]==m0.2,3])
        if (BF>=1) flag<<-1
        else flag<<-rbinom(1,1,BF)
        if (flag= =1)
        {
            MO.var<<-M1$var
            MO.out<<-M1$out
            MO.1<<-M1.1
            MO.2<<-M1.2
        }
big.list[big.list[,1]= =M0.1&big.list[,2]= =m0.2,4]<<- big.list[big.list[,1]= =M0.1&big.list[,2]=
=m0.2,4]+1
i<-i+1
}
var.vect<<-matrix(as.logical(rep(big.list[,1]1,rep(length(MO.var),length(big.list[,1])))%/%
2^(0:(length(MO.var)-1))%%2),ncol=length(MO.var),byrow=T)
nvar<-length(MO.var)
ndx<-1:n.var
Xn<-rep("X",n.var)
labs<-paste(Xn,ndx,sep="")
dimnames(var.vect)<-list(c(1:length(var.vect[,1])),labs)
postprob<<-matrix((exp(big.list[,3]))/(sum(exp(big.list[,3]))),ncol=1)
dimnames(postprob)[2]_list(c('Post.Mod.Pr."))
visits<<-matrix(big.list[,4],ncol=1)
dimnames(visits[2]<-list(c("#visits"))
if (length(out.list)!=0)
{
```

```
        out.vect<<-matrix(as.logical(rep(big.list[,2]-
        1,rep(length(outs.list),length(big.list[,2])))%/%2^(0:(length(outs.list)-
        1))%%2),ncol=length(outs.list),byrow=T)
        dimnames(out.vect)<-list(c(1:length(out.vect[,1])),c(outs.list))
        model.matrix<<-cbind(var.vect,out.vect,postprob,visits)}
        else model.matrix<<-cbind(var.vect,postprob,visits)}
        colno<-length(MO.var)+length(MO.out)+1
        model.matrix<<- model.matrix[order(-model.matrix[,colon]),]
        return(model.matrix)
}
MC3.REG.choose<-function(MO.var,MO.out)
{
        var<-MO.var
        in.or.out<-sample(c(1:length(MO.var),rep(0,length(MO.out))),1)
        if(in.or.out==0)
        {
                out<-MO.out
                in.or.out2<-sample(1:length(MO.out),1)
                out[in.or.out2]<-!MO.out[in.or.out2]
        }
        else
        {
                var[in.or.out]<-!MO.var[in.or.out]
                out<-MO.out
        }
        return(var,out)
}
MC3.REG.logpost<-function(x,y,a,b)
{
        a<<-a
        b<<-b
        x_x[,model.vect]
```

```
        xtx_t(x)%*%x
        xty_t(x)%*%y
        beta_ginverse(xtx)%*%xty
        beta_rep(1,numx+1)
}
for (i in 1 : numx)
        {
                post_(exp(inverse(beta*(b+numx)))*(beta*(a+(sum(y)[i])-1)
        }
        logpost_log(post,10)
        return(logpost)
}
numx_numx
samplesize_samplesize
sde_sde
numloop_500
x_matrix(scan("x.txt"),ncol=1,byrow=T)
mseopm_matrix(nrow=loop,ncol=1)
fr_function(regwhich,xtemp,p.5,y,rowwhich,numtemp)
{
      q_t(xtemp)%*%xtemp
      bfull_solve(q)%*%(t(xtemp)%*%y)
      r_matrix(nrow=rowwhich,ncol=numxtemp)
      sumr_matrix(nrow=numxtemp,ncol=1)
      for(j in 1:rowwhich)
      {
            for(k in 1:numxtemp)
            {
                  r[j,k]_((bfull[k,1]^2)*q[k,k])%*%((regwhich[j,k]-p.5[k,1])^2)
            }
      sum[j,1]_sum(r[j,])
      }
```

```
return(sumr)
}
for(i in 1:numloop)
{
	resultmc<-BMA.MC3(y,x,10000,rep(T,numx),NULL,NULL,0,0,ip,cc,a,b)
	rowresult_nrow(resultmc)
	p_matrix(nrow=numx,ncol=1)
	numxtemp_0
	for (j in 1:numx)
	{
		prob_matrix(nrow=rowresult,ncol=1)
	for (k in 1:rowresult)
	{
	prob[k,1]_resultmc[k,j]%*%resultmc[k,numx+1]
	}
	p[j,1]_sum(prob)
		if(p[j,1]>=0.5)
		{
			numxtemp_numxtemp+1
		}
	}
xtemp_matrix(nrow=samplesize,ncol=numxtemp)
p.5_matrix(nrow=numxtemp,ncol=1)
sumx_1
for (l in 1:numx)
{
	if(p[l,1]>=0.5)
	{
		p.5[l,1]_p[l,1]
		xtemp[,sumx]_x[,l]
		sumx_sumx+1
	}
```

```
}
reg_leaps(xtemp,y,rep(1,samplesize),int=T,method="adjr2",keep.int=T,nbest=1,df=samplesize)
numxtemp_ncol(xtemp)
rowwhich_nrow(reg$which)
}
mr_fr(reg$which,xtemp,p.5,y,rowwhich,numxtemp)
minr_min(mr)
for (l in 1:rowwhich)
{
      if (mr[l,1]==minr)
      {
            regmodel_lm(y~xtemp[,reg$which[l,]])
            mseopm[i]_sum(regmodel$residuals^2)/regmodel$df
      }
}
amseopm_sum(mseopm)/numloop
print(amseopm)
stdopm_matrix(nrow=numloop,ncol=1)
for (indexi in 1:numloop)
{
        stdopm[indexi]_(mseopm[indexi]-amseopm)^2
}
stdamseopm_sqrt(sum(stdopm)/(numloop-1))
print(stdamseopm)
```

การทำงานของโปรแกรมที่ใช้ในการวิจัย สำหรับการสร้างตัวแบบด้วยวิธี SR โดยใช้โปรแกรม S-plus 2000

โปรแกรมที่ใช้ในการสร้างตัวแบบด้วยวิธี SR นั้นสร้างขึ้นโดยใช้โปรแกรม S-plus 2000 โดยรายละเอียดของโปรแกรมมีดังนี้

โปรแกรมสำหรับสร้างตัวแบบด้วยวิธี SR

```
numx_numx
samplesize_samplesize
sde_sde
numloop_500
x_matrix(scan("x.txt"),ncol=1,byrow=T)
msesr_matrix(nrow=numloop,ncol=1)
fy_function(samplesize,sde,x,numx)
{
        error_ matrix(scan("error.txt"),ncol=1,byrow=T)
        ones_rep(1,samplesize)
        xones_cbind(ones,x)
        beta_rep(1,numx+1)
        y_(xones%*%beta)+error
        return(y)
}
for(i in 1:numloop)
{
    y_fy(samplesize,sde,x,numx)
    resultsr_stepwise(x,y,intercept=T,tolerance=1.e-07,method="ex",nbest=3)
    regmodel_lm(y~x[,sr$which[3,]])
    msesr[i]_sum(regmodel$residuals^2)/regmodel$df
}
amsesr_sum(msesr)/numloop
print(amsesr)
stdsr_matrix(nrow=numloop,ncol=1)
for (indexi in 1:numloop)
```

```
{
	stdsr[indexi]_(msesr[indexi]-amsesr)^2
}
stdamsesr_sqrt(sum(stdsr)/(numloop-1))
print(stdamsesr)
```

ประวัติผู้เขียนวิทยานิพนธ์

นางสาววรรษมน จันทร์โอกุล เกิดเมื่อวันที่ 3 มกราคม พ.ศ. 2526 สำเร็จการศึกษาระดับปริญญาตรีวิทยาศาสตรบัณฑิต สาขาวิชาเอกสถิติ จากคณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร เมื่อ พ.ศ. 2546 และเข้าศึกษาต่อในระดับปริญญามหาบัณฑิต สาขาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย เมื่อ พ.ศ. 2547