

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการวิจัยทางด้านธุรกิจ เศรษฐศาสตร์ และสาขาต่างๆ ต้องอาศัยวิธีการทางด้านสถิติ มาช่วยในการวิเคราะห์เพื่อหาคำตอบ โดยเฉพาะอย่างยิ่งการวิจัยเพื่อคาดคะเนเหตุการณ์ล่วงหน้า หรือที่เรียกว่า การพยากรณ์ ซึ่งวิธีที่ใช้ในการพยากรณ์นั้นจะใช้วิธีการวิเคราะห์ความถดถอย แต่เนื่องจากปัจจุบันการพยากรณ์ส่วนใหญ่จะอยู่ในรูปแบบที่มีการขึ้นและลงของเส้นโค้ง ดังนั้นวิธีการวิเคราะห์ที่นำมาใช้ในการพยากรณ์ คือ การวิเคราะห์ความถดถอยพหุนาม(Polynomial Regression Analysis) ซึ่งค่าพยากรณ์ของตัวแบบจะมีความถูกต้องหรือไม่ขึ้นอยู่กับตัวแบบ ความถดถอยพหุนามที่เหมาะสม วิธีการหนึ่งที่น่าสนใจในการพิจารณาหาตัวแบบที่เหมาะสม คือ การคัดเลือกตัวแบบ(Model Selection) การคัดเลือกตัวแบบความถดถอยพหุนามส่วนใหญ่จะใช้วิธีการคัดเลือกตัวแปร(Variable Selection) มาใช้ในการคัดเลือกตัวแบบ แต่บางครั้งการคัดเลือกตัวแปรจะได้ตัวแบบที่ไม่เหมาะสมเพราะบางสถานการณ์ไม่ได้พิจารณาครบทุกตัวแปร ดังนั้นงานวิจัยครั้งนี้ได้เสนอการคัดเลือกตัวแบบที่เป็นไปได้ทั้งหมดบนพื้นฐานของการทดสอบสมมติฐานแบบติดกลุ่ม (Nested Hypothesis) ซึ่งการทดสอบดังกล่าวตัวแบบที่นำมาพิจารณาเป็นตัวแบบติดกลุ่ม โดยทุกพจน์ของตัวแบบแรกจะเป็นส่วนหนึ่งของตัวแบบที่สอง

การคัดเลือกตัวแบบความถดถอยพหุนามแบบติดกลุ่ม จะเป็นการคัดเลือกตัวแบบที่มีค่าพยากรณ์ที่ถูกต้องและเหมาะสมที่สุด โดยเกณฑ์ที่น่าจะนำมาใช้ในการคัดเลือกตัวแบบ คือ เกณฑ์การคัดเลือกตัวแบบโดยสถิติทดสอบเอฟบางส่วน (The partial F-test statistic) ด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง(Backward Elimination(BW)) เพราะเริ่มต้นด้วยการพิจารณาจากตัวแปรอิสระทุกตัวแปร แต่อาจไม่ดีขึ้น ถ้าตัวแปรบางตัวมีความสัมพันธ์ร่วมเชิงพหุ(Multicollinearity) กับตัวแปรตัวอื่นๆ ดังนั้น การคัดเลือกตัวแบบด้วยวิธีการของเบย์ (Bayesian Model Selection) จะเป็นการอนุมานด้วยความน่าจะเป็นภายหลัง โดยกำหนดพารามิเตอร์เป็นตัวแปรสุ่ม และเรียกการแจกแจงของพารามิเตอร์ว่า การแจกแจงก่อน(Prior Distribution) แล้วนำมารวมกับฟังก์ชันภาวะน่าจะเป็น(Likelihood Function) เพื่อหาความน่าจะเป็นภายหลัง(Posterior Probability) มาใช้ในการคัดเลือกตัวแบบ ซึ่งในปี 1978 ชวาร์ซ (Schwarz) ได้เสนอเกณฑ์ในการคัดเลือกตัวแบบโดยการพิจารณาค่าสูงสุดของความน่าจะเป็นภายหลัง(Posterior Distribution) มาเป็นส่วนประกอบในการพิจารณา นั่นคือ เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบย์ (Bayesian Information Criterion(BIC)) และต่อมาในปี 1998 โรเบิร์ต อี เคสส์(Robert E. Kass) และ เอเดิร์น

อริราฟร์เทอรี (Adrian E. Raftery) ได้นำตัวประกอบของเบส์ (Bayes factor) มาใช้ในการประมาณ ซึ่งอยู่ในรูปของสมการ

$$\text{BIC} = -2 \log (\text{Bayes Factor})$$

$$= -2 \log \left[\frac{p(X | \hat{\beta}_{-k}, M_k)}{p(X | \hat{\beta}_{-j}, M_j)} \right]$$

เมื่อ	$\hat{\beta}_{-j}, \hat{\beta}_{-k}$	แทน	ตัวประมาณภาวะน่าจะเป็นสูงสุด
	M_j, M_k	แทน	ตัวแบบที่ j และ k ที่นำมาเปรียบเทียบโดย $M_k \subset M_j$
และ	$p(X \hat{\beta}_{-k}, M_k)$	แทน	ความน่าจะเป็นภายหลังของตัวแบบที่ k
	$p(X \hat{\beta}_{-j}, M_j)$		และ j

นอกจากนี้ในปี 1973 อาไคเคะ (Akaike) ได้เสนอเกณฑ์ในการคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike's Information Criterion (AIC)) เพื่อหาตัวแบบที่ให้ค่าพยากรณ์มีความถูกต้องมากที่สุด ซึ่งอยู่ในรูปของสมการ

$$\text{AIC} = L - p$$

เมื่อ	L	แทน	ฟังก์ชันภาวะน่าจะเป็นสูงสุด
	p	แทน	จำนวนพารามิเตอร์

และต่อมาอาไคเคะได้ใช้แนวคิดของคูแบล็ค - ไลท์เบอร์ (Kulback-Leibler 1951) มาใช้ในการคัดเลือกตัวแบบที่เหมาะสมที่สุดในระยะอนันต์ (Asymptotically) และมีตัวแปรในตัวแบบจำนวนมาก ดังนั้นเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike's Information Criterion (AIC)) จะอยู่ในรูปของสมการ

$$\text{AIC} = -2 \log(ML_j - ML_k) + 2(p_j - p_k)$$

เมื่อ	ML_j, ML_k	แทน	ฟังก์ชันภาวะน่าจะเป็นสูงสุดของตัวแบบที่ j และ k
	p_j, p_k	แทน	จำนวนพารามิเตอร์ของตัวแบบที่ j และ k

ดังนั้นการวิจัยครั้งนี้ผู้วิจัยได้สนใจทำการศึกษาและเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบความถดถอยพหุนามแบบติดกลุ่ม โดยเกณฑ์ที่ใช้ในการศึกษา คือ เกณฑ์การคัดเลือกตัวแบบโดยสถิติทดสอบเอฟบางส่วน(The partial F-test statistic) ด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง (Backward Elimination(BW)) เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike's Information Criterion(AIC)) และเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (Bayesian Information Criterion(BIC)) เพื่อหาข้อสรุปว่าเกณฑ์การคัดเลือกตัวแบบใดให้ค่าพยากรณ์ถูกต้องและแม่นยำมากที่สุด เพื่อให้การเปรียบเทียบมีความชัดเจนและใช้ประโยชน์ได้อย่างกว้างขวาง ดังนั้น การวิจัยครั้งนี้ได้ทำการทดสอบสมมติฐานแบบติดกลุ่ม (Nested Hypothesis) ซึ่งวิธีการทดสอบสมมติฐานแบบติดกลุ่มนี้ เหมาะสำหรับการตัดสินใจในตัวแบบที่มีความสามารถในการพยากรณ์ใกล้เคียงกันได้ชัดเจน

1.2 วัตถุประสงค์ของการวิจัย

1.เพื่อศึกษาเกณฑ์การคัดเลือกตัวแบบในการวิเคราะห์ความถดถอยพหุนาม โดยมีเกณฑ์ที่ใช้ในการศึกษาดังนี้

- 1.1 เกณฑ์การคัดเลือกตัวแบบโดยสถิติทดสอบเอฟบางส่วน(The partial F-test statistic) ด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง(Backward Elimination(BW))
- 1.2 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike's Information Criterion(AIC))
- 1.3 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(Bayesian Information Criterion(BIC))

2.เพื่อเปรียบเทียบความถูกต้องของการพยากรณ์ด้วยการคัดเลือกตัวแบบจากเกณฑ์ต่างๆ ภายใต้การทดสอบสมมติฐานแบบติดกลุ่ม(Nested Hypothesis)

1.3 สมมติฐานการวิจัย

ในการอนุมานแบบติดกลุ่ม(Nested Inference)การคัดเลือกตัวแปรด้วยเกณฑ์การคัดเลือกตัวแปรโดยข้อสนเทศของเบส์(Bayesian Information Criterion(BIC)) สามารถคัดเลือกตัวแปรที่ให้ค่าพยากรณ์มีความถูกต้องและเหมาะสมกว่า เกณฑ์การคัดเลือกตัวแปรโดยสถิติทดสอบเอฟบางส่วน(The partial F-test statistic) ด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง(Backward Elimination(BW)) และเกณฑ์การคัดเลือกตัวแปรโดยข้อสนเทศของอาไคเคะ(Akaike's Information Criterion(AIC)) เนื่องจากเกณฑ์การคัดเลือกตัวแปรโดยข้อสนเทศของเบส์(Bayesian Information Criterion(BIC)) จะพิจารณาจากความน่าจะเป็นภายหลังของตัวแปรที่เหมาะสม

1.4 ข้อตกลงเบื้องต้น

1.รูปแบบทั่วไปของสมการความถดถอยพหุนามมีรูปแบบดังนี้

$$\underset{\sim}{Y} = \underset{\sim}{X} \underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \quad (1.1)$$

เมื่อ	$\underset{\sim}{Y}$	แทน	เวกเตอร์สุ่มของตัวแปรตามขนาด $n \times 1$
	$\underset{\sim}{X}$	แทน	เมทริกซ์ของตัวแปรอิสระที่ควบคุมให้คงที่ขนาด $n \times (p+1)$
	$\underset{\sim}{\beta}$	แทน	พารามิเตอร์เวกเตอร์ของสัมประสิทธิ์ความถดถอยขนาด $(p+1) \times 1$
	$\underset{\sim}{\varepsilon}$	แทน	เวกเตอร์ความคลาดเคลื่อนสุ่มที่เกิดขึ้นขนาด $n \times 1$ โดย $\underset{\sim}{\varepsilon} \sim N(0, \sigma^2 I)$
	I	แทน	เมทริกซ์เอกลักษณ์ขนาด $n \times n$
	n	แทน	จำนวนข้อมูลที่สังเกต
และ	p	แทน	จำนวนตัวแปรอิสระ

2.ตัวแปรอิสระแต่ละตัวแปรเป็นค่าคงที่

3.ความคลาดเคลื่อนสุ่มเป็นตัวแปรสุ่มที่มีการแจกแจงปกติ $N(0, \sigma^2)$ เหมือนกันและเป็นอิสระซึ่งกันและกัน

4. ตัวประมาณค่าของพารามิเตอร์สัมประสิทธิ์ความถดถอย β ของตัวแบบความถดถอยเชิงเส้นข้างต้น คือ

4.1 ตัวประมาณค่ากำลังสองน้อยสุด(Least Square Estimator: \mathbf{b}) คือ

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}$$

4.2 ตัวประมาณค่าภาวะน่าจะเป็นสูงสุด(Maximum Likelihood Estimator:

\mathbf{b}_{-M}) คือ

$$\mathbf{b}_{-M} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}$$

1.5 คำจำกัดความที่ใช้ในการวิจัย

ตัวแบบติดกลุ่ม(Nested Models) หมายถึง ตัวแบบ 2 ตัวแบบจะติดกัน ถ้าในแต่ละพจน์ของตัวแบบแรกเป็นส่วนหนึ่งของตัวแบบที่สอง ซึ่งตัวแบบที่สองจะมีพจน์มากกว่าตัวแบบแรกอย่างน้อย 1 เทอม ตัวแบบที่สองที่มีความซับซ้อนมากกว่าตัวแบบแรกจะเรียกว่าตัวแบบที่สมบูรณ์(Complete Model) และตัวแบบแรกที่เป็นตัวแบบอย่างง่ายของตัวแบบสอง เรียกว่าตัวแบบลดรูป(Reduced Model)

ตัวแปรตาม(Dependence Variable) หมายถึง ตัวแปรที่ได้ผลกระทบจากตัวแปรอิสระ

ตัวแปรอิสระ(Independence Variable) หมายถึง ตัวแปรที่มีผลกระทบต่อตัวแปรตาม

เกณฑ์(Criteria) หมายถึง ค่าคาดหวังของตัวประมาณ ที่ใช้วัดความแตกต่างระหว่างตัวแบบที่สมบูรณ์(Complete Model) และ ตัวแบบลดรูป(Reduced Model)

1.6 ขอบเขตของการวิจัย

1. สมมติฐานที่ใช้ในการวิจัย คือ การทดสอบสมมติฐานแบบติดกลุ่ม (Nested Hypothesis) นั่นคือ

$$H_0 : M_j = M_0$$

$$H_1 : M_k \in R^{m-m_0} \quad \text{โดยที่ } M_j \subset M_k$$

เมื่อ	R	แทน	ปริภูมิจำนวนจริง
	m, m_0	แทน	จำนวนพารามิเตอร์ของตัวแบบที่ใช้ในการพิจารณา
	M_j, M_k	แทน	ตัวแบบที่ j และ k ที่ใช้ในการพิจารณา
	M_0	แทน	ตัวแบบลดรูปที่มีทุกพจน์อยู่ในตัวแบบ M_k

2. ตัวแบบของความถดถอยพหุนามเชิงเส้นที่สนใจศึกษาเป็นดังนี้

2.1 เมื่อตัวแบบที่สมบูรณ์ (Complete Model) มีกำลังสูงสุดของตัวแปรอิสระเป็น 6 ตัวแบบอยู่ในรูปของ

$$\begin{aligned}
 y_i = & \beta_0 + \beta_{10}x_{i1} + \beta_{20}x_{i1}^2 + \beta_{30}x_{i1}^3 + \beta_{40}x_{i1}^4 + \beta_{50}x_{i1}^5 + \beta_{60}x_{i1}^6 \\
 & + \beta_{01}x_{i2} + \beta_{02}x_{i2}^2 + \beta_{03}x_{i2}^3 + \beta_{04}x_{i2}^4 + \beta_{05}x_{i2}^5 + \beta_{06}x_{i2}^6 \\
 & + \beta_{11}x_{i1}x_{i2} + \beta_{12}x_{i1}x_{i2}^2 + \beta_{13}x_{i1}x_{i2}^3 + \beta_{14}x_{i1}x_{i2}^4 + \beta_{15}x_{i1}x_{i2}^5 \\
 & + \beta_{21}x_{i1}^2x_{i2} + \beta_{22}x_{i1}^2x_{i2}^2 + \beta_{23}x_{i1}^2x_{i2}^3 + \beta_{24}x_{i1}^2x_{i2}^4 \\
 & + \beta_{31}x_{i1}^3x_{i2} + \beta_{32}x_{i1}^3x_{i2}^2 + \beta_{33}x_{i1}^3x_{i2}^3 \\
 & + \beta_{41}x_{i1}^4x_{i2} + \beta_{42}x_{i1}^4x_{i2}^2 \\
 & + \beta_{51}x_{i1}^5x_{i2} + \varepsilon_i
 \end{aligned}$$

ในกรณีนี้มีตัวแปรเชิงพหุจำนวน 27 ตัวแปร

2.2 เมื่อตัวแบบที่สมบูรณ์ (Complete Model) มีกำลังสูงสุดของตัวแปรอิสระเป็น 5 ตัวแบบอยู่ในรูปของ

$$\begin{aligned}
 y_i = & \beta_0 + \beta_{10}x_{i1} + \beta_{20}x_{i1}^2 + \beta_{30}x_{i1}^3 + \beta_{40}x_{i1}^4 + \beta_{50}x_{i1}^5 \\
 & + \beta_{01}x_{i2} + \beta_{02}x_{i2}^2 + \beta_{03}x_{i2}^3 + \beta_{04}x_{i2}^4 + \beta_{05}x_{i2}^5 \\
 & + \beta_{11}x_{i1}x_{i2} + \beta_{12}x_{i1}x_{i2}^2 + \beta_{13}x_{i1}x_{i2}^3 + \beta_{14}x_{i1}x_{i2}^4 \\
 & + \beta_{21}x_{i1}^2x_{i2} + \beta_{22}x_{i1}^2x_{i2}^2 + \beta_{23}x_{i1}^2x_{i2}^3 \\
 & + \beta_{31}x_{i1}^3x_{i2} + \beta_{32}x_{i1}^3x_{i2}^2 \\
 & + \beta_{s1}x_{i1}^4x_{i2} + \varepsilon_i
 \end{aligned}$$

ในกรณีนี้มีตัวแปรเชิงพหุจำนวน 20 ตัวแปร

2.3 เมื่อตัวแบบที่สมบูรณ์ (Complete Model) มีกำลังสูงสุดของตัวแปรอิสระเป็น 4 ตัวแบบอยู่ในรูปของ

$$\begin{aligned}
 y_i = & \beta_0 + \beta_{10}x_{i1} + \beta_{20}x_{i1}^2 + \beta_{30}x_{i1}^3 + \beta_{40}x_{i1}^4 \\
 & + \beta_{01}x_{i2} + \beta_{02}x_{i2}^2 + \beta_{03}x_{i2}^3 + \beta_{04}x_{i2}^4 \\
 & + \beta_{11}x_{i1}x_{i2} + \beta_{12}x_{i1}x_{i2}^2 + \beta_{13}x_{i1}x_{i2}^3 \\
 & + \beta_{21}x_{i1}^2x_{i2} + \beta_{22}x_{i1}^2x_{i2}^2 \\
 & + \beta_{s1}x_{i1}^2x_{i2} + \varepsilon_i
 \end{aligned}$$

ในกรณีนี้มีตัวแปรเชิงพหุจำนวน 14 ตัวแปร

2.4 เมื่อตัวแบบที่สมบูรณ์ (Complete Model) มีกำลังสูงสุดของตัวแปรอิสระเป็น 3 ตัวแบบอยู่ในรูปของ

$$\begin{aligned}
 y_i = & \beta_0 + \beta_{10}x_{i1} + \beta_{20}x_{i1}^2 + \beta_{30}x_{i1}^3 \\
 & + \beta_{01}x_{i2} + \beta_{02}x_{i2}^2 + \beta_{03}x_{i2}^3 \\
 & + \beta_{11}x_{i1}x_{i2} + \beta_{12}x_{i1}x_{i2}^2 \\
 & + \beta_{s1}x_{i1}^2x_{i2} + \varepsilon_i
 \end{aligned}$$

ในกรณีนี้มีจำนวนตัวแปรอิสระ 9 ตัวแปร

2.5 เมื่อตัวแบบที่สมบูรณ์ (Complete Model) มีกำลังสูงสุดของตัวแปรอิสระเป็น 2 ตัวแบบอยู่ในรูปของ

$$\begin{aligned}
 y_i = & \beta_0 + \beta_{10}x_{i1} + \beta_{20}x_{i1}^2 \\
 & + \beta_{01}x_{i2} + \beta_{02}x_{i2}^2 \\
 & + \beta_3x_{i1}x_{i2} + \varepsilon_i
 \end{aligned}$$

ในกรณีนี้มีตัวแปรเชิงพหุจำนวน 5 ตัวแปร

2.6 เมื่อตัวแบบที่สมบูรณ์(Complete Model)มีกำลังสูงสุดของตัวแปรอิสระเป็น 1 ตัวแบบอยู่ในรูปของ

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} + \varepsilon_i$$

ในกรณีนี้มีตัวแปรเชิงพหุจำนวน 3 ตัวแปร

3.การวิจัยครั้งนี้กำหนดให้ $\beta' = (1 \dots 1)_{1 \times (p+1)}$ ในประชากรทุกรูปแบบที่ศึกษาโดยที่ p' เป็นจำนวนตัวแปรอิสระ

4.ขนาดตัวอย่างที่ศึกษา(n) ที่ศึกษา คือ 35 50 75 และ 100

5.จำนวนตัวแปรอิสระเริ่มต้นที่ศึกษาคือ 2 ตัวแปร โดยกำหนดให้ตัวแปรอิสระควบคุมให้เป็นค่าคงที่ ซึ่งอยู่ในรูป $f(y | x_1, x_2, \dots, x_n)$ และค่าของ x มีระยะห่างเท่า ๆ กัน การวิจัยครั้งนี้ได้ทำการแปลงข้อมูลด้วย 2 วิธี คือ การแปลงข้อมูลโดยใช้พหุนามเชิงตั้งฉาก และการแปลงค่าเข้าสู่ศูนย์กลาง และกำหนดระดับความสัมพันธ์ของตัวแปรอิสระ x_1 และ x_2 ซึ่งแบ่งออกเป็น 3 ระดับคือ

5.1. x_1 และ x_2 มีความสัมพันธ์กันในระดับสูง ($\rho(x_1, x_2) = 1$)

5.2. x_1 และ x_2 มีความสัมพันธ์กันในระดับปานกลาง ($\rho(x_1, x_2) = 0.5$)

5.3. x_1 และ x_2 มีความสัมพันธ์กันในระดับต่ำ ($\rho(x_1, x_2) = 0$)

7.ระดับนัยสำคัญของการทดสอบ(α)ที่ศึกษาคือ 0.05 และ 0.01

8.ในการวิจัยครั้งนี้จะศึกษาเมื่อความคลาดเคลื่อนสุ่มมีการแจกแจงปกติ $N(0, \sigma^2)$ โดยกำหนดให้ $\sigma = 5$ 10 15 และ 25

9.การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง ทำการสร้างแบบจำลองข้อมูลจากการทำซ้ำจำนวน 500 รอบ โดยใช้เทคนิคการจำลองแบบมอนติคาร์โล(Monte Carlo Simulation) และเขียนด้วยโปรแกรม S-plus 2000

1.7 ประโยชน์ของการวิจัย

1.7.1 เพื่อเป็นแนวทางในการศึกษาและคัดเลือกตัวแบบความถดถอยพหุนามแบบติดกลุ่มกรณีที่มี 2 ตัวแปรอิสระซึ่งเกิดอันตรกิริยา สำหรับการนำไปใช้ในการพยากรณ์

1.7.2 เพื่อหาตัวแบบที่เหมาะสม โดยไม่จำเป็นต้องเป็นตัวแบบจริง ซึ่งจะมีความสะดวกในเรื่อง การประมวลผล การวิเคราะห์ข้อมูล และการแปลความหมายตลอดจนการนำเสนอข้อมูลที่ถูกต้อง

1.8 เกณฑ์ที่ใช้ในการตัดสินใจ

เกณฑ์การตัดสินใจว่าตัวแบบใดจะมีความถูกต้องและเหมาะสมมากที่สุด จะพิจารณาจากเกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Average of Mean Square Error (AMSE)) มีสูตรดังนี้

$$MSE_j = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p}$$

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

เมื่อ	y_i	แทน	ค่าสังเกตที่ i
	\hat{y}_i	แทน	ค่าพยากรณ์ที่ i
	p	แทน	จำนวนของพารามิเตอร์ในตัวแบบ
	n	แทน	ขนาดตัวอย่าง
	MSE_j	แทน	ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการทำซ้ำรอบที่ j
	$AMSE$	แทน	ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการทำซ้ำ
	r	แทน	จำนวนรอบของการทำซ้ำ

โดยเกณฑ์การคัดเลือกตัวแบบใดที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดจะเป็นเกณฑ์ที่ดีที่สุด

1.9 ขั้นตอนในการดำเนินงานวิจัย

1. กำหนดความคลาดเคลื่อนสุ่มจากประชากรที่มีการแจกแจงปกติ มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 5 10 15 และ 25

1.1 กำหนดขนาดตัวอย่างที่ศึกษาเท่ากับ 35 50 75 และ 100

1.2 กำหนดตัวแปรอิสระเริ่มต้นเป็นค่าคงที่

1.3 กำหนดระดับนัยสำคัญ (α) คือ 0.05 และ 0.01

1.4 กำหนดตัวแปรอิสระเริ่มต้นที่ใช้สำหรับการสร้างตัวแปรตามที่เหมาะสมคือ 2 ตัวแปร

- 1.5 กำหนดกำลังสูงสุดของตัวแปรอิสระใช้สำหรับการสร้างตัวแปรตามที่เหมาะสม คือ กำลังหก
2. สร้างข้อมูลของตัวแปรตาม y จากตัวแปรอิสระตามระดับความสัมพันธ์ต่างๆ ที่กำหนดไว้ในขอบเขตของการวิจัย
3. ทำการแปลงข้อมูลให้เป็นไปตามคุณสมบัติพหุนามเชิงตั้งฉาก(Orthogonal Polynomial)และการแปลงเข้าสู่ศูนย์กลาง(Centering)
4. กำหนดตัวแบบสมบูรณ์(Complete Model)ในการคัดเลือกตัวแบบที่เหมาะสม
5. ทำการสร้างตัวแบบจากเกณฑ์การคัดเลือกตัวแบบทั้ง 3 เกณฑ์ คือ
 - 5.1 เกณฑ์การคัดเลือกตัวแบบโดยสถิติทดสอบเอฟบางส่วน(The partial F-test statistic)ด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง(Backward Elimination(BW))
 - 5.2 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike's Information Criterion(AIC))
 - 5.3 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(Bayesian Information Criterion(BIC))
6. คำนวณหาค่าคลาดเคลื่อนกำลังสองเฉลี่ยจากการทำซ้ำจำนวน 500 รอบของตัวแบบที่ได้จากเกณฑ์การคัดเลือกตัวแบบทั้ง 3 เกณฑ์
7. ทำการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวแบบที่ได้จากเกณฑ์การคัดเลือกตัวแบบทั้ง 3 เกณฑ์ โดยสรุปผลในรูปของตารางและรูปภาพ (หน้า 42 - 117)