



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การวิเคราะห์การถดถอยเชิงพหุ (Multiple Regression Analysis) คือการศึกษาความสัมพันธ์ระหว่างตัวแปรตาม (Dependent Variable) และกลุ่มของตัวแปรอิสระ (Set of Independent variables) ภายใต้รูปแบบที่กำหนด ซึ่งอาจเป็นเส้นตรงหรือเส้นโค้ง จากการวิเคราะห์การถดถอยจะได้สมการถดถอยเชิงพหุ ซึ่งสามารถนำไปใช้ประโยชน์ในการพยากรณ์และอธิบายลักษณะของตัวแปรตามได้ โดยทั่วไปเมื่อได้สมการถดถอยเชิงพหุ โดยวิธีการส่งร่องน้อยที่สุด สามารถวัดได้ว่าสมการถดถอยเชิงพหุ มีความเหมาะสมหรือไม่ โดยใช้ค่าสถิติสัมประสิทธิ์การตัดสินใจ (Coefficient of determination) ซึ่งเป็นอัตราส่วนระหว่างส่วนเบี่ยงเบนเนื่องจากสมการถดถอยและส่วนเบี่ยงเบนทั้งหมด นิยมบอกเป็นเปอร์เซ็นต์ เช่น ค่าสถิติสัมประสิทธิ์การตัดสินใจเป็น 98 เปอร์เซ็นต์ หมายถึงภายใต้รูปแบบที่กำหนดกลุ่มของตัวแปรอิสระ สามารถอธิบายตัวแปรตามได้ 98 ส่วนใน 100 ส่วน ในการวิจัย ผู้วิจัยสนใจสมการถดถอยเชิงพหุที่ให้ค่าสถิติสัมประสิทธิ์การตัดสินใจสูง เนื่องจากได้รูปแบบความสัมพันธ์ที่ดี และสมการถดถอยเชิงพหุที่ได้สามารถใช้อธิบายและพยากรณ์ตัวแปรตามได้ถูกต้องมากขึ้น ปัญหาที่เกิดขึ้นก็คือ ค่าสถิติ สัมประสิทธิ์การตัดสินใจจะมีค่ามากหรือน้อยไม่ได้ขึ้นกับความสัมพันธ์ระหว่างตัวแปรตามและกลุ่มของตัวแปรอิสระเพียงอย่างเดียว แต่มีผลเนื่องมาจากลักษณะของข้อมูลที่นำมาศึกษาด้วย เช่น ข้อมูลของกลุ่มตัวแปรอิสระเป็นเชิงคุณภาพ (Qualitative data) ข้อมูลมีค่าผิดปกติ (Outlier) และในบางกรณีกลุ่มของตัวแปรอิสระ มีความสัมพันธ์กับตัวแปรตามน้อยหรือไม่มีความสัมพันธ์กัน เมื่อเพิ่มเข้าไปในรูปแบบที่กำหนด ค่าสถิติสัมประสิทธิ์การตัดสินใจก็มีค่าเพิ่มขึ้น เมื่อเกิดปัญหาดังกล่าวจึงมีผู้วิจัยพยายามคิดปรับปรุงค่าสถิติสัมประสิทธิ์การตัดสินใจขึ้นใหม่ โดยการปรับค่าสถิติสัมประสิทธิ์การตัดสินใจด้วยค่าคงที่ และการหาค่าสถิติสัมประสิทธิ์การตัดสินใจโดยใช้ค่ามีรยฐาน จึงเป็นที่น่าศึกษาว่า ค่าสถิติสัมประสิทธิ์การตัดสินใจ (R^2) ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ปรับค่า (R_1^2) ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ใช้ค่ามีรยฐาน (R_2^2) ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ใช้มีรยฐานและปรับค่า (R_3^2) ค่าใดใช้วัดความเหมาะสมของสมการถดถอยเชิงพหุได้ดีที่สุด

1.2 วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบค่าสถิติสัมประสิทธิ์การตัดสินใจในสมการถดถอยเชิงพหุ เพื่อหาค่าสถิติสัมประสิทธิ์การตัดสินใจที่ใช้วัดความเหมาะสมของสมการถดถอยเชิงพหุได้ดีที่สุด

1.3 สมมติฐานในการวิจัย

ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ปรับค่า ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ใช้ค่ามัธยฐาน และค่าสถิติสัมประสิทธิ์ การตัดสินใจที่ใช้ค่ามัธยฐานและปรับค่าใช้วัดความเหมาะสมของสมการถดถอยเชิงพหุ ได้ดีกว่าค่าสถิติสัมประสิทธิ์การตัดสินใจ

1.4 ขอบเขตของการวิจัย

1.4.1 ศึกษาสมการถดถอยเชิงพหุซึ่งอยู่ในรูปแบบ

$$Y = X\beta + \epsilon$$

โดยที่

Y คือเมตริกซ์ของตัวแปรตามขนาด $N \times 1$

X คือเมตริกซ์ของกลุ่มตัวแปรอิสระขนาด $N \times (M+1)$

β คือเมตริกซ์ของสัมประสิทธิ์ความถดถอยพหุขนาด $(M+1) \times 1$

ϵ คือเมตริกซ์ของความคลาดเคลื่อนกลุ่มขนาด $N \times 1$

N คือจำนวนค่าสังเกต

M คือจำนวนตัวแปรอิสระ

1.4.2 ขนาดตัวอย่างเมื่อเปรียบเทียบกับขนาดประชากรเป็น 10% 14% และ

20%

1.4.3 กลุ่มของตัวแปรอิสระมีลักษณะดังนี้

- 1.4.3.1 ข้อมูลเชิงคุณภาพ
- 1.4.3.2 ข้อมูลเชิงปริมาณ
- 1.4.3.3 ข้อมูลเชิงคุณภาพและข้อมูลเชิงปริมาณมีสัดส่วนเป็น 1:1
1:2 และ 1:3
- 1.4.4 จำนวนตัวแปรอิสระ (M) เท่ากับ 2 3 และ 4
- 1.4.5 ความคลาดเคลื่อนกลุ่มที่ใช้เป็น
- 1.4.5.1 ความคลาดเคลื่อนกลุ่มที่มีการแจกแจงปกติ (Normal distribution) มีค่าเฉลี่ยเป็นศูนย์ ความแปรปรวนเป็น 0.25 0.64 และ 1.00
- 1.4.5.2 ความคลาดเคลื่อนกลุ่มที่มีการแจกแจงปกติปลอมปน เนื่องจาก
เลกกล (Scale-contaminated Normal distribution) ที่มีค่าเลกกลเป็น (b) 9 25
และ 100 มีเปอร์เซ็นต์ของความปลอมปน (Percent of contaminated) (λ) เป็น 0.01
และ 0.05
- 1.4.5.3 ความคลาดเคลื่อนกลุ่มที่มีการแจกแจงไม่สมมาตร (Asymmetric distribution) โดยให้ความโด่ง เป็น 3 และความเบ้เป็น 0.15 0.50 และ 0.75
- 1.4.6 กลุ่มตัวอย่างใช้วิธี Sequential Sampling without Replacement
- 1.4.7 การสุ่มว่ากลุ่มตัวอย่างมีค่าผิดปกติ ใช้ตัวสถิติ T (studentize residuals)
ทดสอบ
- 1.4.8 จำนวนซ้ำในการวิเคราะห์
- 1.4.8.1 ถ้ามีการตรวจสอบค่าผิดปกติจะทำซ้ำ 50 ครั้ง
- 1.4.8.2 ถ้าไม่มีการตรวจสอบค่าผิดปกติ จะทำซ้ำ 100 ครั้ง
- 1.4.9 เกณฑ์ที่ใช้ในการเปรียบเทียบค่า R^2 R_1^2 R_2^2 และ R_3^2 ใช้ค่า
Relative Mean Square Error ของ R_1^2 R_2^2 R_3^2 เทียบกับ R^2
- 1.4.10 ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ใช้ในการเปรียบเทียบมีดังนี้
- 1.4.10.1 ค่าสถิติสัมประสิทธิ์การตัดสินใจ

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}, \quad i = 1, 2, 3, \dots, N$$

1.4.10.2 ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ปรับค่า

$$R_1^2 = 1 - \frac{(N-1)}{(N-M-1)} \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}, \quad i = 1, 2, 3, \dots, N$$

1.4.10.3 ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ใช้ค่ามัธยฐาน

$$R_2^2 = 1 - \frac{\text{Med} |Y_i - \hat{Y}_i|^2}{\text{Med} |Y_i - \bar{Y}|^2}, \quad i = 1, 2, 3, \dots, N$$

Med หมายถึงมัธยฐาน

1.4.10.4 ค่าสถิติสัมประสิทธิ์การตัดสินใจที่ใช้ค่ามัธยฐานและปรับค่า

$$R_3^2 = 1 - \frac{(N-1)}{(N-M-1)} \frac{\text{Med} |Y_i - \hat{Y}_i|^2}{\text{Med} |Y_i - \bar{Y}|^2}, \quad i = 1, 2, 3, \dots, N$$

Med หมายถึงมัธยฐาน

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ผู้วิเคราะห์สามารถเลือกใช้ตัวสถิติสัมประสิทธิ์การตัดสินใจที่เหมาะสมในการวิเคราะห์การถดถอยเชิงพหุ
2. เป็นแนวทางในการศึกษาค่าสถิติสัมประสิทธิ์การตัดสินใจต่อไป