



## 1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันข้อมูลข่าวสารมีความสำคัญและจำเป็นต่อการดำเนินงานในหลาย ๆ ด้าน เช่น ธุรกิจ อุตสาหกรรม เกษตรกรรม การแพทย์ การศึกษา การเมือง การตัดสินใจหรือการวางแผน ในกิจกรรมต่าง ๆ ที่จะต้องดำเนินงานมักต้องใช้ข้อมูลประกอบ เช่นทางการศึกษา จะต้องใช้ข้อมูลคาดคะเนเพื่อวางแผนจัดการศึกษาให้ตรงกับความต้องการของเศรษฐกิจและสังคม ในทางธุรกิจ จำเป็นต้องมีการใช้ข้อมูลเพื่อนำมาพัฒนาธุรกิจของตนเองให้สามารถแข่งขันกับคู่แข่งในโลกของธุรกิจได้ เป็นต้น

การวิเคราะห์ความถดถอย (Regression Analysis) เป็นวิธีการทางสถิติที่นำมาใช้คาดคะเนหรือพยากรณ์ค่าของตัวแปรที่สนใจศึกษาซึ่งเรียกว่า ตัวแปรตาม (Dependent Variable) ใช้สัญลักษณ์แทนด้วย  $Y$  โดยที่ค่าของตัวแปรตามต้องอาศัยความสัมพันธ์กับตัวแปรอื่นที่เกี่ยวข้องซึ่งมีอิทธิพลต่อตัวแปรตามที่เรียกว่า ตัวแปรอิสระ (Independent Variable) ใช้สัญลักษณ์แทนด้วย  $X$  เมื่อพิจารณาความสัมพันธ์ระหว่างตัวแปรตาม 1 ตัวกับ ตัวแปรอิสระ 1 ตัว เรียกว่าการวิเคราะห์ความถดถอยอย่างง่าย (Simple Regression Analysis) ซึ่งในสภาพความเป็นจริงนั้น ตัวแปรอิสระที่มีอิทธิพลต่อตัวแปรตามอาจจะมีได้มากกว่า 1 ตัว การวิเคราะห์กรณีนี้เรียกว่า การวิเคราะห์ความถดถอยเชิงพหุ (Multiple Regression Analysis) ซึ่งเป็นการวิเคราะห์ความถดถอยแบบหนึ่งซึ่งซับซ้อนกว่าการวิเคราะห์การถดถอยอย่างง่ายที่มีตัวแปรอิสระเพียงตัวแปรเดียวเท่านั้น รวมไปถึงในความเป็นจริงแล้วการคาดการณหรือการพยากรณ์เชิงปริมาณใดก็ตาม ตัวแปรอิสระที่เกี่ยวข้องพัวพันกับเหตุการณ์ที่สนใจนั้นไม่ได้มีเพียงแค่ตัวแปรเดียว ดังนั้นเพื่อการวิเคราะห์ที่อยู่ในรูปแบบที่ใกล้เคียงกับความเป็นจริงผู้วิจัยจึงใช้ตัวแบบความถดถอยเชิงพหุแบบในรูปเชิงเส้นสำหรับการวิจัยในครั้งนี้

โดยทั่วไปแล้วทดสอบสมมติฐาน  $H_0: A\beta = c$  กับ  $H_1: A\beta \neq c$  หรือในกรณีที่ไม่ต้องกำหนดค่าพารามิเตอร์อื่นซึ่งเป็นพารามิเตอร์ที่ไม่ต้องกำหนดค่า มักใช้ในการทดสอบอัตราส่วนความควรจะเป็นแบบปกติ (Regular likelihood ratio test) เนื่องด้วยเป็นวิธีการที่ใช้กันอย่างแพร่หลาย ซึ่งการทดสอบดังกล่าวมีขั้นตอนและการคำนวณที่ยุ่งยาก ต้องใช้ระยะเวลาในการวิเคราะห์

การทดสอบอัตราส่วนความควรจะเป็นแบบปกติ (Regular likelihood ratio test) ซึ่งเป็นวิธีการที่เสนอโดยเนย์แมนและเพียร์สัน (Neyman and Pearson) ตั้งแต่ปี ค.ศ. 1928 ซึ่งมีบทบาทที่สำคัญในการทดสอบสมมติฐาน โดยหาตัวประมาณพารามิเตอร์จากวิธีความควรจะเป็นสูงสุดจากทฤษฎีการประมาณค่า ซึ่งการทดสอบด้วยอัตราส่วนความควรจะเป็นจะใช้กันอย่างกว้างขวางเพราะสามารถให้การแจ่มแจ้งความน่าจะเป็นประมาณการทดสอบได้แต่อาจจะเกิดความผิดพลาดในการยอมรับหรือปฏิเสธสมมติฐานดังกล่าว

นอกจากการทดสอบด้วยวิธีดังกล่าวนี้ มีอีกวิธีการหนึ่งคือการทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โล (Monte Carlo Likelihood ratio test) ซึ่งวิธีมอนติคาร์โล เป็นเทคนิคหนึ่งที่ใช้แก้ปัญหาทางคณิตศาสตร์เพื่อนำมาใช้ในการทดสอบ โดยสร้างข้อมูลของตัวอย่างและทำการทดสอบสมมติฐานของสัมประสิทธิ์ความถดถอยด้วยอัตราส่วนความควรจะเป็นที่อาศัยการคำนวณด้วยคอมพิวเตอร์เพื่ออนุมานทางสถิติที่สนใจ ซึ่งวิธีนี้ถือเป็นแนวคิดใหม่ที่ใช้วิธีการทดสอบมอนติคาร์โล (Monte Carlo test) มาประยุกต์ใช้ในการวิเคราะห์ความถดถอยเชิงพหุ คือจะทำการสุ่มข้อมูลเป็นค่าพารามิเตอร์ตั้งต้นในตัวแบบและมีการกระทำซ้ำเพื่อทดสอบสมมติฐานของอิทธิพลโดยรวมของสัมประสิทธิ์ความถดถอย

สำหรับการเปรียบเทียบการทดสอบทั้งสองวิธีดังกล่าวได้มีผู้ทำการศึกษาภายใต้สถานการณ์ต่าง ๆ ดังนี้

ศศิธร เจษฎาฐิติกุล (2545) ได้ศึกษาการเปรียบเทียบการทดสอบความเป็นอิสระระหว่างตัวแปร 2 ตัวแปรเมื่อตัวแปรอิสระมีการแจกแจงพหุนามและอยู่ในตารางการณัจร 2 ทางซึ่งการทดสอบที่ใช้ในการเปรียบเทียบมี 3 วิธี คือ วิธีการทดสอบไคกำลังสองเพียร์สัน(CP) วิธีการทดสอบด้วยอัตราส่วนควรจะเป็น (MLR) และวิธีการมอนติคาร์โล (MC) สามารถสรุปผลการวิจัยได้ดังนี้ 1. ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 วิธีการทดสอบความเป็นอิสระระหว่าง 2 ตัวแปร วิธีการทดสอบทั้ง 3 วิธี สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ทุกวิธี 2. อำนาจการทดสอบ โดยทั่วไปวิธีการมอนติคาร์โลจะให้อำนาจการทดสอบสูงที่สุด รองลงมาคือวิธีการทดสอบด้วยอัตราส่วนควรจะเป็น ส่วนวิธีการทดสอบไคกำลังสองเพียร์สันมีอำนาจการทดสอบต่ำสุด กรณีที่รูปแบบของตารางการณัจรเป็นแบบจัตุรัสวิธีการทดสอบทั้ง 3 วิธีจะให้อำนาจการทดสอบใกล้เคียงกัน แต่ในกรณีที่รูปแบบของตารางการณัจรไม่เป็นจัตุรัส กล่าวคือ จำนวนแถวกับสดมภ์ไม่เท่ากัน อำนาจการทดสอบมีแนวโน้มลดลงเมื่อความแตกต่างระหว่างแถวกับสดมภ์เพิ่มขึ้น อำนาจการทดสอบของวิธีการทดสอบทั้ง 3 วิธี แปรผันตามขนาดตัวอย่าง ระดับความสัมพันธ์ของข้อมูล และระดับนัยสำคัญ โดยเรียงลำดับจากมากไปน้อยตามลำดับ

อรไท สงวนสินธ์ (2545) ได้ศึกษาการเปรียบเทียบวิธีการทดสอบสมมติฐานเกี่ยวกับความแตกต่างระหว่างอิทธิพลของทรีทเมนต์ กรณีที่ขนาดตัวอย่างในแต่ละทรีทเมนต์เท่ากัน 2 วิธี คือ การทดสอบเอฟและการทดสอบมอนติคาร์โลด้วยอัตราส่วนภาวะน่าจะเป็น เมื่อความแตกต่างระหว่างอิทธิพลของทรีทเมนต์แตกต่างกันน้อย ตัวสถิติทดสอบมอนติคาร์โลด้วยอัตราส่วนภาวะน่าจะเป็นจะให้อำนาจการทดสอบสูงสุด เมื่อความแตกต่างระหว่างอิทธิพลของทรีทเมนต์แตกต่างกันปานกลาง โดยส่วนใหญ่ตัวสถิติทดสอบมอนติคาร์โลด้วยอัตราส่วนภาวะน่าจะเป็นจะให้อำนาจการทดสอบสูงสุด แต่เมื่อจำนวนทรีทเมนต์และขนาดตัวอย่างในแต่ละทรีทเมนต์เพิ่มขึ้น สัมประสิทธิ์ความแปรผันสูงขึ้น ตัวสถิติทดสอบเอฟจะให้อำนาจการทดสอบสูงสุด และเมื่อความแตกต่างระหว่างอิทธิพลของทรีทเมนต์แตกต่างกันมาก ตัวสถิติทดสอบทั้ง 2 วิธีจะให้อำนาจการทดสอบเท่ากันหรือใกล้เคียงกัน

จากข้อมูลดังกล่าวข้างต้น ในการวิจัยครั้งนี้ผู้วิจัยจึงสนใจศึกษาและเปรียบเทียบการทดสอบอัตราส่วนความควรจะเป็นแบบปกติ และการทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โลสำหรับการวิเคราะห์ความถดถอยเชิงพหุแบบเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ

## 1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและเปรียบเทียบการทดสอบสมมติฐานของสัมประสิทธิ์ความถดถอยเชิงพหุจากการทดสอบ 2 วิธีคือ

- 1.2.1 การทดสอบอัตราส่วนความควรจะเป็นแบบปกติ
- 1.2.2 การทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โล

## 1.3 สมมติฐานการวิจัย

การทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โลให้ค่าอำนาจการทดสอบสูงกว่าการทดสอบอัตราส่วนความควรจะเป็นแบบปกติ

## 1.4 ขอบเขตของการวิจัย

- 1.4.1 วิธีการทดสอบสมมติฐานที่ทำการศึกษา คือ
  - 1.4.1.1 การทดสอบอัตราส่วนความควรจะเป็นแบบปกติ
  - 1.4.1.2 การทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โล
- 1.4.2 ตัวแบบความถดถอยเป็นตัวแบบที่อยู่ภายใต้ข้อตกลงเบื้องต้น
- 1.4.3 ประชากรที่ศึกษาสร้างมาจากตัวแบบ  $y = X\beta + \varepsilon$

1.4.4 กำหนดจำนวนตัวแปรอิสระที่ทำการศึกษา (P) คือ 2 3 4 และ 5 โดยสร้างเมทริกซ์ตัวแปรอิสระจากการจำลองข้อมูลการแจกแจงแบบปกติมาตรฐานหลายตัวแปรพร้อมทั้งกำหนดความสัมพันธ์ระหว่างตัวแปรให้มีค่าเป็นศูนย์

1.4.5 กำหนดขนาดตัวอย่างของแต่ละชุดค่าสังเกต (n) คือ 10 25 50 และ 100

1.4.6 กำหนดค่าพารามิเตอร์ตั้งต้นของสมการถดถอยเป็นดังนี้

$$1.4.6.1 \beta_p = (10)_{(p+1) \times 1} \text{ เมื่อ } p=2, \dots, 5 \text{ สำหรับการหาค่าความ}$$

คลาดเคลื่อนประเภทที่ 1 (Type I error)

$$1.4.6.2 \beta_p \neq (10)_{(p+1) \times 1} \text{ บางค่า สำหรับการหาค่าอำนาจการทดสอบ}$$

(Power of the test) โดยกำหนดความแตกต่างระหว่างสัมประสิทธิ์ความถดถอยแบ่งออกเป็น 3 ระดับ คือ

1.4.6.2.1 ความแตกต่างระหว่างสัมประสิทธิ์ความถดถอย มีความแตกต่างที่แปรผัน 5% ของส่วนเบี่ยงเบนมาตรฐาน

1.4.6.2.2 ความแตกต่างระหว่างสัมประสิทธิ์ความถดถอย มีความแตกต่างที่แปรผัน 50% ของส่วนเบี่ยงเบนมาตรฐาน

1.4.6.2.3 ความแตกต่างระหว่างสัมประสิทธิ์ความถดถอย มีความแตกต่างที่แปรผัน 90% ของส่วนเบี่ยงเบนมาตรฐาน

เมื่อ k คือร้อยละที่แปรผันของส่วนเบี่ยงเบนมาตรฐานซึ่งมีค่าเท่ากับ 5%SD , 50%SD และ 90%SD

$$\text{กรณีตัวแปรอิสระ 2 ตัว } \beta = \begin{pmatrix} 10 & 10 & 10+k \end{pmatrix}^T$$

$$\text{กรณีตัวแปรอิสระ 3 ตัว } \beta = \begin{pmatrix} 10 & 10 & 10 & 10+k \end{pmatrix}^T$$

$$\text{กรณีตัวแปรอิสระ 4 ตัว } \beta = \begin{pmatrix} 10 & 10 & 10 & 10 & 10+k \end{pmatrix}^T$$

$$\text{กรณีตัวแปรอิสระ 5 ตัว } \beta = \begin{pmatrix} 10 & 10 & 10 & 10 & 10 & 10+k \end{pmatrix}^T$$

1.4.7 กำหนดให้ข้อมูลมีค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation:  $\sigma$ ) ในระดับต่าง ๆ กันคือ 1 3 5 และ 7

1.4.8 ระดับนัยสำคัญของการทดสอบ ( $\alpha$ ) ที่ศึกษาคือ 0.01 0.05 และ 0.1

1.4.9 ในการวิจัยครั้งนี้สร้างแบบจำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล (Monte Carlo simulation) เขียนด้วยโปรแกรม S-PLUS 2000

- 1.4.10 การจำลองในแต่ละสถานการณ์ของการทดลองการทำซ้ำ 1500 รอบ
- 1.4.11 การสร้างตัวอย่างสุ่มในวิธีการทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โลจะกระทำซ้ำ 100 รอบ

1.5 ข้อตกลงเบื้องต้น

1.5.1 สมการถดถอยที่ใช้ในการศึกษาเป็นแบบสมการถดถอยเชิงพหุ โดยมีตัวแบบดังนี้

$$y = X\beta + \epsilon$$

$$\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$\underset{\sim}{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$X = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n-1} & X_{2n-1} & \cdots & X_{pn-1} \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix}_{(n \times (p+1))}$$

- $y$  แทนเวกเตอร์ของค่าสังเกตหรือตัวแปรที่ได้ตามตัวแปรอิสระ ( $X_i$ 's) ที่มีขนาด  $n \times 1$
- $\beta$  แทนเวกเตอร์สัมประสิทธิ์ความถดถอยของตัวแบบความถดถอยขนาด  $(p+1) \times 1$
- $\epsilon$  แทนเวกเตอร์ความคลาดเคลื่อนสุ่มที่มีขนาด  $n \times 1$
- $n$  แทนขนาดตัวอย่าง
- $p$  แทนจำนวนตัวแปรอิสระ
- $X$  แทนเมทริกซ์ตัวแปรอิสระที่มีขนาด  $n \times (p+1)$

1.5.2 ลักษณะของสมการถดถอยต้องเป็นเชิงเส้น (Linearity)

1.5.3  $E\left(\underset{\sim}{\epsilon}\right) = \underset{\sim}{0}$

$$1.5.4 \quad \text{Var} \begin{bmatrix} \varepsilon \\ \sim \end{bmatrix} = E \begin{bmatrix} \varepsilon \varepsilon^T \\ \sim \sim \end{bmatrix} = \sigma^2 I_n$$

$$\text{เมื่อ } E \begin{bmatrix} \varepsilon \varepsilon^T \\ \sim \sim \end{bmatrix} = \begin{bmatrix} E(\varepsilon_1^2) & E(\varepsilon_1 \varepsilon_2) & \cdots & E(\varepsilon_1 \varepsilon_n) \\ E(\varepsilon_2 \varepsilon_1) & E(\varepsilon_2^2) & \cdots & E(\varepsilon_2 \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_n \varepsilon_1) & E(\varepsilon_n \varepsilon_2) & \cdots & E(\varepsilon_n^2) \end{bmatrix}_{n \times n} = \sigma^2 I_n$$

นั่นคือ  $E(\varepsilon_i^2) = \sigma^2 = \text{Var}(\varepsilon_i)$  และ  $E(\varepsilon_i \varepsilon_j) = \text{cov}(\varepsilon_i, \varepsilon_j) = 0; i \neq j$

$$1.5.5 \quad \begin{matrix} \text{iid} \\ \varepsilon \sim N_n \left( \begin{matrix} 0 \\ \sim \end{matrix}, \sigma^2 I_n \right) \\ \sim \end{matrix}$$

1.5.6 เมทริกซ์  $X$  เป็นเมทริกซ์ของค่าคงที่ ซึ่งมีค่าลำดับชั้น (Rank) เท่ากับ  $p+1 < n$

นั่นคือสดมภ์ของ  $X$  แต่ละสดมภ์เป็นอิสระต่อกัน

$$1.5.7 \quad \begin{matrix} \text{iid} \\ y \sim N_n \left( \begin{matrix} X\beta \\ \sim \end{matrix}, \sigma^2 I_n \right) \\ \sim \end{matrix}$$

## 1.6 เกณฑ์ที่ใช้ในการตัดสินใจ

ในการวิจัยครั้งนี้จะพิจารณาสถิติการทดสอบอัตราส่วนความควรจะเป็นแบบปกติและอัตราส่วนความควรจะเป็นแบบมอนติคาร์โลดังนี้

1.6.1 ค่าความคลาดเคลื่อนประเภทที่ 1 (Type I error) คือ นำค่า p-value ของทั้งสองวิธีเทียบกับระดับนัยสำคัญ ( $\alpha$ ) ที่จะศึกษา เมื่อกำหนดชุดข้อมูลให้สอดคล้องกับสมมติฐานว่างโดยการนับชุดข้อมูลที่มีการปฏิเสธสมมติฐานว่าง  $H_0$  ต่อจำนวนชุดข้อมูลทั้งหมด (1,500 ชุดข้อมูล) สามารถสรุปได้ดังนี้

$$\hat{\alpha} = \frac{\text{จำนวนครั้ง}(p\text{-value} \leq \alpha)}{\text{จำนวนรอบภายนอกที่ทำการทดลอง}}$$

1.6.2 ค่าอำนาจการทดสอบ (Power of the test) คือนำค่า p-value ของทั้งสองวิธีเทียบกับระดับนัยสำคัญ ( $\alpha$ ) ที่จะศึกษา เมื่อกำหนดชุดข้อมูลให้สอดคล้องกับสมมติฐานแย้งโดยการนับชุดข้อมูลที่มีการปฏิเสธสมมติฐานว่าง  $H_0$  ต่อจำนวนชุดข้อมูลทั้งหมด (1,500 ชุดข้อมูล) สามารถสรุปได้ดังนี้

$$(1 - \hat{\beta}) = \frac{\text{จำนวนครั้ง}(p\text{-value} \leq \alpha)}{\text{จำนวนรอบภายนอกที่ทำการทดลอง}}$$

## 1.7 คำจำกัดความที่ใช้ในการวิจัย

1.7.1 ความคลาดเคลื่อนประเภทที่ 1 (Type I error) คือ ความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นจริง ความน่าจะเป็นของความผิดพลาดชนิดนี้มีค่าเท่ากับระดับนัยสำคัญของการทดสอบ แทนด้วย  $\alpha$

1.7.2 ความคลาดเคลื่อนประเภทที่ 2 (Type II error) คือ ความผิดพลาดที่เกิดจากการยอมรับสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นไม่จริง ความน่าจะเป็นของความผิดพลาดชนิดนี้แทนด้วย  $\beta$

1.7.3 อำนาจการทดสอบ (Power of the test) คือ ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างนั้นไม่จริง ซึ่งแทนด้วย  $1-\beta$

1.7.4 P – value (มานพ วรภักดี ,2547:349) คือ ความน่าจะเป็นต่ำสุดที่สามารถปฏิเสธสมมติฐานว่าง  $H_0$  เมื่อสมมติฐานว่าง  $H_0$  เป็นจริง

## 1.8 ประโยชน์ที่คาดว่าจะได้รับ

1.8.1 เพื่อใช้เป็นแนวทางในการทดสอบสมมติฐานสำหรับตัวแบบการวิเคราะห์ความถดถอยเชิงพหุโดยใช้การทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โล

1.8.2 เพื่อใช้เป็นแนวทางในการศึกษาการทดสอบสมมติฐานสำหรับตัวแบบการวิเคราะห์ความถดถอยแบบอื่น ๆ

## 1.9 วิธีดำเนินการวิจัย

ศึกษาการทดสอบสมมติฐานสำหรับตัวแบบการวิเคราะห์ความถดถอยเชิงพหุ สรุปขั้นตอนทั้งหมดไว้ดังนี้

1.9.1 ศึกษาและทำความเข้าใจในทฤษฎีที่เกี่ยวข้องกับการทดสอบสมมติฐานของแต่ละวิธี คือ

1.9.1.1 การทดสอบอัตราส่วนความควรจะเป็นแบบปกติ

1.9.1.2 การทดสอบอัตราส่วนความควรจะเป็นแบบมอนติคาร์โล

1.9.2 ศึกษาการเขียนโปรแกรมจำลองตัวแปรอิสระและตัวแปรตามในตัวแบบที่ทำการศึกษา

1.9.3 จำลองข้อมูลตามขอบเขตที่ต้องการศึกษา รวมทั้งเขียนโปรแกรมทดสอบสมมติฐาน

1.9.4 สรุปผลข้อมูลที่ได้จากการจำลอง