บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการวิเคราะห์ทางสถิติเข้ามามีบทบาทอย่างมากในการหาข้อสรุปของปัญหา ไม่ว่าจะเป็นในเรื่องของการเกษตร เศรษฐกิจ สังคม การแพทย์ และการศึกษา โดยมีหลายหน่วยงานได้นำสถิติไปช่วยในการตัดสินใจเพื่อให้เกิดความเสี่ยงน้อยและข้อสรุปที่น่าเชื่อถือที่สุด ซึ่งหลายหน่วยงานก็อาจจะนำเอาวิธีทางสถิติต่าง ๆ ไปใช้ แต่วิธีที่มีบทบาทอย่างมากในการพยากรณ์และนิยมใช้อย่างมากคือ การวิเคราะห์ความถดถอย (Regression Analysis) และการวิเคราะห์อนุกรมเวลา (Time Series Analysis)

การวิเคราะห์ความถดถอยเป็นการศึกษาเกี่ยวกับการหาฟังก์ชันหรือรูปแบบความสัมพันธ์ เพื่อใช้ในการทำนายค่าของตัวแปรที่ต้องการศึกษา ซึ่งตัวแปรต้องการศึกษานี้เรียกว่าตัวแปรตาม (Depentdent Variable) หรือตัวแปรต้น ในการวิเคราะห์ส่วนใหญ่แล้วนิยมให้แทนด้วย Y ซึ่งจะมีเพียงตัวเดียว โดยในการสร้างฟังก์ชันหรือหาความสัมพันธ์นั้นอาศัยความรู้เกี่ยวกับค่าของตัวแปรที่เกี่ยวข้องหนึ่งตัวหรือมากกว่า ซึ่งเรียกว่าตัวแปรอิสระ (Indepentdent Variables) หรือตัวแปรที่ใช้ในพยากรณ์ (Predictor Variables)มักแทนด้วย X ถ้ามีตัวแปรอิสระเพียงตัวเดียว และแทนด้วย X_1 , X_2 , X_3 , ... กรณีที่มีตัวแปรอิสระหลายตัว ฟังก์ชันหรือความสัมพันธ์ที่กล่าวถึงนั้นเรียกว่า ตัวแบบการถดถอย (Regression Model) หรือสมการการถดถอย (Regression Equations) ตัวแบบการถดถอยสามารถเขียนให้อยู่ในรูปทั่วไป ดังนี้

$$Y_t = f(x_t;\beta) + \varepsilon_t$$

เมื่อ $f(x_t;\beta)$ คือฟังก์ชันทางคณิตศาสตร์ของตัวแปรอิสระ p-1 ตัว

$x_t = (x_{t1}, x_{t2}, ..., x_{t(p-1)})'$ คือ เวกเตอร์ของตัวแปรอิสระ p-1 ตัว

$\beta = (\beta_0, \beta_1, ..., \beta_{p-1})'$ คือเวกเตอร์พารามิเตอร์ของสัมประสิทธิ์การถดถอยที่ไม่ทราบค่า

ε_t คือความคลาดเคลื่อนซึ่งเป็นตัวแปรสุ่ม

ทั้งนี้มีข้อสมมติเกี่ยวกับตัวแปรสุ่ม ε_t ในแต่ละค่าของตัวแปร X ที่กำหนดดังนี้

1. ε_t มีการแจกแจงแบบปกติ
2. ค่าเฉลี่ย $E(\varepsilon_t) = 0$
3. ความแปรปรวน $V(\varepsilon_t) = \sigma^2$ ซึ่งเป็นค่าคงที่และไม่ขึ้นกับ t
4. ε_t เป็นอิสระต่อกัน หรือ $Cov(\varepsilon_j, \varepsilon_k) = 0$, $j \neq k$

จากข้อสมมติของค่าคลาดเคลื่อนพบว่า ตัวแบบการถดถอยที่เหมาะสม ควรจะมีค่าคลาดเคลื่อนเป็นไปตามข้อสมมติ หากข้อสมมติของค่าคลาดเคลื่อนที่ได้จากตัวแบบการถดถอยไม่เป็นไปตามที่กำหนด เมื่อนำตัวแบบนั้นไปใช้ในการพยากรณ์ก็จะทำให้เกิดข้อผิดพลาดได้มาก ดังนั้นก่อนที่จะนำตัวแบบการถดถอยนั้นไปใช้ควรมีการทดสอบเสียก่อนว่ามีความถูกต้องหรือไม่ โดยการทดสอบเทียบความกลมกลืน (Goodness-of -fit Test) สำหรับตัวแบบการถดถอย โดยตัวสถิติทดสอบที่ใช้กันอย่างแพร่หลายคือ ตัวสถิติทดสอบเอฟ (F Statistic) การทดสอบเทียบความกลมกลืนนี้เป็นการทดสอบการเป็นเชิงเส้นของตัวแบบการถดถอย ซึ่งเป็นการทดสอบสมมติฐานว่างที่ว่า "ตัวแบบการถดถอยเป็นตัวแบบเชิงเส้นในตัวแปรอิสระ" ตัวสถิติทดสอบเอฟนั้นเป็นตัวสถิติทดสอบที่เกิดจากอัตราส่วนระหว่างค่ากำลังสองเฉลี่ยของการเทียบความกลมกลืน (Lack of fit mean square) และค่ากำลังสองเฉลี่ยของค่าคลาดเคลื่อนโดยตรง (Pure error mean square) ซึ่งอัตราส่วนนี้จะมีการแจกแจงแบบเอฟ ด้วยระดับขั้นของความเป็นอิสระ c-p และ n-c เมื่อ n คือขนาดตัวอย่างทั้งหมด และ c คือจำนวนค่าของตัวแปรอิสระ X ที่แตกต่างกัน เนื่องจากตัวสถิติทดสอบเอฟนั้นมีข้อจำกัดคือตัวแปรอิสระนั้นจะต้องมีอย่างน้อยหนึ่งค่าที่ซ้ำกัน ดังนั้นจะมีตัวสถิติทดสอบแบบใดบ้างที่ให้ความเหมาะสมมากพอในสถานการณ์ดังกล่าว ซึ่งเป็นสิ่งที่น่าจะทำการวิจัย

งานวิจัยที่ศึกษาเกี่ยวกับการทดสอบเทียบความกลมกลืนนั้นมีนักวิจัยหลายท่านที่ได้ทำการศึกษา ทิพย์วัลย์ กันทอง (2544) ได้ทำการเปรียบเทียบอำนาจการทดสอบของการทดสอบเทียบความกลมกลืนสำหรับการถดถอย โดยใช้ตัวสถิติทดสอบเอฟ สถิติทดสอบ Kolmogorov-smimov (KS) และตัวสถิติทดสอบ Cramer-von Mises (CvM) ซึ่งในกรณีที่ตัวแปรอิสระมีค่าซ้ำกันตัวสถิติทดสอบเอฟให้อำนาจการทดสอบที่สูง ส่วนในกรณีที่ตัวแปรอิสระมีค่าไม่ซ้ำกันนั้น สถิติทดสอบ KS มีอำนาจการทดสอบที่สูงสุด Jianqing Fan (1996) ได้เสนอการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงของประชากร ด้วยวิธี ตัวสถิติทดสอบ Adaptive Neyman และตัวสถิติทดสอบ Hard and Soft Thresholding ในปี ค.ศ. 1996 Kuchibhatla และ Hart (Jianqing Fan และ Li-Shan Haung , 2001 : 645) ได้เสนอตัวสถิติทดสอบเทียบความกลมกลืนโดยการแปลงค่าความคลาดเคลื่อน (Residaul) ด้วยการแปลงแบบฟูเรียร์ (Fourier Transform) ต่อจากนั้นในปี ค.ศ. 2001 Jianqing Fan และ Li-Shan Haung ได้เสนอการทดสอบเทียบความกลมกลืนสำหรับการถดถอยด้วยสถิติทดสอบ Adaptive Neyman (AN) โดยตัวสถิติทดสอบ AN นั้นได้แปลงค่าเศษเหลือ วิธีเดียวกันกับวิธีของ Kuchibhatla และ Hart (KH) ซึ่งค่าของสถิติทดสอบจะขึ้นอยู่กับตัวสถิติอันดับของค่าเศษเหลือ แต่ไม่มีข้อจำกัดว่าตัวแปรอิสระนั้นจะมีค่าซ้ำหรือไม่ ผู้วิจัยจึงสนใจที่จะนำตัวสถิติดังกล่าวนี้มาทดสอบสมมติฐาน

เพื่อเปรียบเทียบตัวสถิติทดสอบในการทดสอบเปรียบเทียบความกลมกลืนสำหรับตัวแบบการถดถอย โดยพิจารณาตัวสถิติทดสอบเอฟและกรณีศึกษาต่าง ๆ เพิ่มเติม ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยได้ทำการศึกษาเปรียบเทียบอำนาจการทดสอบของการทดสอบเทียบความกลมกลืนสำหรับตัวแบบการถดถอยโดยใช้ตัวสถิติทดสอบ 3 ตัวด้วยกันคือ

1. สถิติทดสอบเอฟ (F)
2. สถิติทดสอบ Adaptive Neyman (AN)
3. สถิติทดสอบ Kuchibhatla และ Hart (KH)

โดยจะศึกษาทั้งในกรณีที่ตัวแปรอิสระมีค่าซ้ำกันและไม่ซ้ำ ในการวิจัยนั้นจะทำการจำลองข้อมูลขนาดตามและลักษณะที่ต้องการด้วยเทคนิคมอนติคาร์โล (Monte Carlo Simulation Technique) เกณฑ์ที่นำมาเปรียบเทียบการทดสอบวิธีต่าง ๆ นั้นคือ ค่าความผิดพลาดประเภทที่ 1 และค่าอำนาจการทดสอบ

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาและเปรียบเทียบหาตัวสถิติทดสอบที่เหมาะสม ในการทดสอบเทียบความกลมกลืนของตัวแบบการถดถอย โดยเกณฑ์ที่นำมาใช้พิจารณาคือความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบ ตัวสถิติที่นำมาวิจัยครั้งนี้ประกอบด้วยตัวสถิติทดสอบ 3 ตัวด้วยกันคือ

- สถิติทดสอบเอฟ (F)
- สถิติทดสอบ Adaptive Neyman (AN)
- สถิติทดสอบ Kuchibhatla และ Hart (KH)

2. เพื่อศึกษาหาตัวแบบที่เหมาะสมในกรณีตัวแปรอิสระมีค่าไม่ซ้ำกัน

1.3 สมมติฐานการวิจัย

สมมติฐานของการวิจัยมีดังนี้

กรณีที่ตัวแปรอิสระมีค่าซ้ำกันสถิติทดสอบ F จะให้อำนาจการทดสอบสูงกว่าสถิติทดสอบ AN และ KH ส่วนในกรณีที่ตัวแปรอิสระไม่มีค่าซ้ำกันนั้นสถิติทดสอบ KH จะให้อำนาจการทดสอบสูงกว่าตัวสถิติแบบอื่น

1.4 ขอบเขตของการวิจัย

1. ศึกษาความสารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทดสอบสำหรับการวิเคราะห์ความถดถอย

2. ศึกษาอำนาจการทดสอบโดยใช้ตัวแบบการถดถอยต่าง ๆ ซึ่งในการศึกษาครั้งนี้เป็นการศึกษาว่าตัวแบบถดถอยเป็นตัวแบบเชิงเส้นในตัวแปรอิสระหรือไม่ เนื่องจากค่าสัมประสิทธิ์การถดถอยไม่เกี่ยวข้องกับการไม่เป็นเชิงเส้นในตัวแปรอิสระ ดังนั้นในการวิจัยจึงกำหนดค่าสัมประสิทธิ์ต่างๆ ดังต่อไปนี้

<u>ตัวแบบสมมติฐานว่างกรณีตัวแปรอิสระ 1 ตัว</u>

<u>ตัวแบบที่ 1.</u> ตัวแบบการถดถอยเป็นตัวแบบเชิงเส้นที่มีตัวแปรอิสระ 1 ตัว

$$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i \quad \text{เมื่อ} \quad 1 \le i \le n$$

<u>ตัวแบบสมมติฐานแย้งกรณีตัวแปรอิสระ 1 ตัว</u>

<u>ตัวแบบที่ 2.</u> ตัวแบบการถดถอยเป็น

$$Y_i = \beta_0 + \beta_1 X_{i1}^2 + \varepsilon_i$$

เมื่อ β_1 มีค่าเท่ากับ 0.2 , 0.4 , 0.6 , 0.8 และ 1.0 , $1 \le i \le n$

<u>ตัวแบบที่ 3.</u> ตัวแบบการถดถอยเป็น

$$Y_i = \beta_0 + \cos(\beta_1 X_{i1} \pi) + \varepsilon_i$$

เมื่อ β_1 มีค่าเท่ากับ 0.2 , 0.4 , 0.6 , 0.8 และ 1.0 , $1 \le i \le n$

<u>ตัวแบบที่ 4.</u> ตัวแบบการถดถอยเป็น

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \varepsilon_i$$

เมื่อ β_2 มีค่าเท่ากับ 0.2 , 0.4 , 0.6 , 0.8 และ 1.0 , $1 \le i \le n$

<u>ตัวแบบสมมติฐานว่างกรณีตัวแปรอิสระ 3 ตัว</u>

<u>ตัวแบบที่ 5.</u> ตัวแบบการถดถอยเป็นตัวแบบเชิงเส้นพหุคูณที่มีตัวแปรอิสระสามตัว

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i \quad \text{เมื่อ} \quad 1 \le i \le n$$

<u>ตัวแบบสมมติฐานแย้งกรณีตัวแปรอิสระ 3 ตัว</u>

<u>ตัวแบบที่ 6.</u> ตัวแบบพหุนามระดับขั้นเป็น 2

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}^2 + \beta_3 X_{i3} + \varepsilon_i$$

เมื่อ β_2 มีค่าเท่ากับ 0.2 , 0.4 , 0.6 , 0.8 และ 1.0 , $1 \le i \le n$

<u>ตัวแบบที่ 7.</u> ตัวแบบการถดถอยเป็น

$$Y_i = \beta_0 + \beta_1 X_{i1} + \cos(\beta_2 X_{i2} \pi) + \beta_3 X_{i3} + \varepsilon_i$$

เมื่อ β_2 มีค่าเท่ากับ 0.2 , 0.4 , 0.6 , 0.8 และ 1.0 , $1 \le i \le n$

<u>ตัวแบบที่ 8.</u> ตัวแบบการถดถอยเป็น

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1} X_{i2} + \beta_3 X_{i3} + \varepsilon_i$$

เมื่อ β_2 มีค่าเท่ากับ 0.2 , 0.4 , 0.6 , 0.8 และ 1.0 , $1 \le i \le n$

3. กำหนดขนาดตัวอย่าง n เท่ากับ 10 , 15 , 20 , 25 , 30 , 40 , 60 และ 80

4. กำหนดระดับนัยสำคัญของการทดสอบ 3 ระดับคือ 0.1 , 0.05 และ 0.01

5. ข้อมูลหรือตัวแปรอิสระที่ใช้ในการวิจัยประกอบด้วย 2 ประเภท คือ

4.1 ข้อมูลที่มีค่าซ้ำกัน

4.2 ข้อมูลที่มีค่าไม่ซ้ำกัน

กรณีที่ 4.2 จะไม่ศึกษาตัวสถิติทดสอบเอฟ เนื่องจากตัวสถิติทดสอบเอฟนั้นกรณีที่ตัวแปรอิสระไม่ซ้ำกันไม่สามารถทดสอบได้

5. ตัวแปรอิสระที่นำมาใช้ในการวิจัยในครั้งนี้กำหนดให้มีการแจกแจงแบบปกติ (Normal Distribution) โดยที่ตัวแปรอิสระแต่ละตัวกำหนดให้มีค่าเฉลี่ยและความแปรปรวนดังนี้ $X_{i1} \sim N(10,5)$, $X_{i2} \sim N(20,50)$ และ $X_{i3} \sim N(30,100)$

6. การแจกแจงของความคลาดเคลื่อนที่นำมาใช้ในการวิจัยมีการแจกแจงแบบปกติ คือ

ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\varepsilon-\mu}{\sigma})^2} \quad ; \quad -\infty < \varepsilon < \infty$$

โดยที่ $\sigma > 0$

ในการวิจัยจะกำหนดค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

1.5 ข้อตกลงเบื้องต้น

1. สมการถดถอยที่ใช้ในสมมติฐานว่างเป็นสมการถดถอยที่มีรูปแบบดังนี้

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

เมื่อ $\underset{\sim}{Y}$ คือเวกเตอร์ที่มีขนาด n x 1

X คือเมทริกซ์ที่มีขนาด n x p โดยที่ p คือจำนวนของพารามิเตอร์

$\underset{\sim}{\beta}$ คือเวกเตอร์ของพารามิเตอร์สัมประสิทธิ์การถดถอยที่มีขนาด p x 1

$\underset{\sim}{\varepsilon}$ คือเวกเตอร์ของค่าคลาดเคลื่อนที่มีขนาด n x 1

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

2. ตัวแปรอิสระ $\underset{\sim}{X} = (X_1, X_2, ..., X_m)'$ มีการแจกแจงร่วม คือ การแจกแจงปกติของหลายตัวแปร (Multivariate Normal Distribution) เขียนเป็นสัญลักษณ์ได้เป็น $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ โดยที่ $\underset{\sim}{\mu} = E(\underset{\sim}{X})$ และ $\Sigma = Cov(\underset{\sim}{X})$

3. ตัวแปรอิสระ X และตัวแปรตาม Y เป็นตัวแปรเชิงปริมาณ

4. สมมติฐานของการทดสอบเทียบความกลมกลืนครั้งนี้คือ

H_0: ตัวแบบการถดถอยเป็นตัวแบบเชิงเส้นในตัวแปรอิสระ

H_1 : ตัวแบบการถดถอยไม่เป็นตัวแบบเชิงเส้นในตัวแปรอิสระ

5. ในกรณีที่ให้ตัวแปรอิสระมีค่าซ้ำกันกำหนดให้ตัวแปรอิสระมีค่าซ้ำกันแบ่งออกเป็น 5 ระดับ

6. ค่าสัมประสิทธิ์การถดถอยที่ใช้ในการทดสอบได้จากการประมาณด้วยตัวประมาณกำลังสองน้อยที่สุด (Least Square Estimator : LSE)

1.6 คำจำกัดความที่ใช้ในการวิจัย

1. ความผิดพลาดประเภทที่ 1 (Type I Error : α) คือ ความคลาดเคลื่อนในการตัดสินใจปฏิเสธ H_0 เมื่อ H_0 เป็นจริง

2. ความผิดพลาดประเภทที่ 2 (Type II Error : β) คือ ความคลาดเคลื่อนในการตัดสินใจยอมรับ H_0 เมื่อ H_0 ไม่เป็นจริง

3. อำนาจการทดสอบ (Power of the Test) คือความน่าจะเป็นที่เกิดจากการปฏิเสธ H_0 เมื่อ H_0 ไม่จริง ซึ่งมีค่าเท่ากับ 1- β

4. การทดสอบเทียบความกลมกลืน (Goodness-of-fit Test) เป็นการทดสอบสมมติฐานทางสถิติที่ใช้ในการทดสอบสมการการถดถอยว่าเป็นไปตามสมมติฐานหรือไม่ ซึ่งในการวิเคราะห์ครั้งนี้จะทำการทดสอบว่าตัวแบบการถดถอยเป็นตัวแบบเชิงเส้นหรือไม่

1.7 เกณฑ์ที่ใช้ในการเปรียบเทียบ

เกณฑ์ที่ใช้เปรียบเทียบในการวิจัยครั้งนี้คือ ค่าความผิดพลาดประเภทที่ 1 และค่าอำนาจการทดสอบ กรณีที่ข้อมูลที่ผลิตขึ้นเป็นข้อมูลที่สอดคล้องกับสมมติฐานว่างที่ว่าตัวแบบนั้นเป็นตัวแบบเชิงเส้น จะได้ว่าสัดส่วนของจำนวนครั้งของการปฏิเสธสมมติฐานว่างก็จะเป็นค่าของความน่าจะเป็นของความผิดพลาดประเภทที่ 1 แต่ถ้าหากข้อมูลที่ผลิตขึ้นไม่สอดคล้องกับสมมติฐานว่างหรือผลิตขึ้นจากข้อมูลที่ไม่เป็นเชิงเส้น จะได้ว่าสัดส่วนของจำนวนครั้งในการปฏิเสธสมมติฐาน

ว่างจะเป็นค่าประมาณของอำนาจการทดสอบ การเปรียบเทียบค่าอำนาจการทดสอบจะทำเฉพาะกรณีที่ตัวสถิติทดสอบนั้นสามารถควบคุมความผิดพลาดประเภทที่ 1 ในสถานการณ์นั้น ๆ ได้เท่านั้น โดยมีรายละเอียดของตรวจสอบดังนี้

1. พิจารณาจากความสามารถในการควบคุมค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ($\hat{\alpha}$) ซึ่งในการทดสอบสมมติฐานจะพยายามควบคุมค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ให้มีค่าน้อย อีกทั้งพยายามทำให้ความผิดพลาดประเภทที่ 2 มีค่าน้อยที่สุด เนื่องจากในทางปฏิบัติแล้วความผิดพลาดประเภทที่ 2 นั้นมีความร้ายแรงกว่าความผิดพลาดประเภทที่ 1 พิจารณาค่าประมาณของความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ด้วยการทดสอบแบบทวินาม (Binomial Test) ที่ระดับนัยสำคัญของการทดสอบแบบทวินาม (α^*) เท่ากับ 0.05 โดยมีรูปแบบของการทดสอบดังนี้

สมมติฐานในการทดสอบ คือ

$$H_0 : \alpha \le \alpha_0$$

$$H_1 : \alpha > \alpha_0$$

ตัวสถิติทดสอบคือ

$$Z = \frac{\hat{\alpha} - \alpha_0}{\sqrt{\frac{\alpha_0(1-\alpha_0)}{n^*}}}$$

จะได้ว่า $$P\left[\frac{\hat{\alpha} - \alpha_0}{\sqrt{\frac{\alpha_0(1-\alpha_0)}{n^*}}} \le Z_{\alpha^*}\right] = 1 - \alpha^*$$

หรือ $$P\left[\hat{\alpha} \le \alpha_0 + Z_{\alpha^*} \cdot \sqrt{\frac{\alpha_0(1-\alpha_0)}{n^*}}\right] = 1 - \alpha^*$$

ดังนั้นช่วงของการยอมรับสมมติฐานว่างคือ $\left(0, \alpha_0 + Z_{\alpha^*} \cdot \sqrt{\frac{\alpha_0(1-\alpha_0)}{n^*}}\right)$

โดยที่ $\hat{\alpha}$ คือ ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 จากการทดสอบด้วยตัวสถิติทดสอบ

α^* คือ ค่าระดับนัยสำคัญของการทดสอบสมมติฐาน $H_0 : \alpha \le \alpha_0$, $H_1 : \alpha > \alpha_0$ ซึ่งกำหนดให้เท่ากับ 0.05

α_0 คือ ระดับนัยสำคัญที่ใช้การทดสอบครั้งนี้

n^* คือ จำนวนรอบของการทดลองครั้งนี้ซึ่งใช้ 1,000 รอบ

วิธีการคำนวณเกณฑ์การทดสอบความน่าจะเป็นของความผิดพลาดประเภทที่ 1 คือ กำหนดให้มีการทดสอบทวินามที่ระดับนัยสำคัญ $\alpha^* = 0.05$ ดังนั้นจะได้ $Z_{\alpha^*} = 1.645$

1. ที่ระดับนัยสำคัญ $\alpha_0 = 0.01$

$$\hat{\alpha} \text{ จะอยู่ในช่วง } \left[0, 0.01 + 1.645\sqrt{\frac{0.01(0.99)}{1,000}}\right] = [0, 0.015]$$

2. ที่ระดับนัยสำคัญ $\alpha_0 = 0.05$

$$\hat{\alpha} \text{ จะอยู่ในช่วง } \left[0, 0.05 + 1.645\sqrt{\frac{0.05(0.95)}{1,000}}\right] = [0, 0.061]$$

3. ที่ระดับนัยสำคัญ $\alpha_0 = 0.10$

$$\hat{\alpha} \text{ จะอยู่ในช่วง } \left[0, 0.10 + 1.645\sqrt{\frac{0.10(0.90)}{1,000}}\right] = [0, 0.116]$$

ดังนั้นช่วงของการยอมรับเป็นดังนี้

- กรณี $\alpha_0 = 0.01$ ช่วงของการยอมรับคือ [0 , 0.015]
- กรณี $\alpha_0 = 0.05$ ช่วงของการยอมรับคือ [0 , 0.061]
- กรณี $\alpha_0 = 0.10$ ช่วงของการยอมรับคือ [0 , 0.116]

จากช่วงของการยอมรับ ในสถานการณ์ใดๆ หากค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ของตัวสถิติทดสอบใดตกอยู่ในช่วงของการยอมรับ นั่นแสดงว่าตัวสถิติทดสอบนั้นสามารถควบคุมความผิดพลาดประเภทที่ 1 ในสถานการณ์นั้นได้

2. ค่าอำนาจการทดสอบ ในสถานการณ์ใด ๆ หากตัวสถิติทดสอบสามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ เราก็จะพิจารณาค่าอำนาจการทดสอบของตัวสถิติทดสอบนั้นต่อ โดยที่ตัวสถิติทดสอบใดมีค่าอำนาจการทดสอบสูงสุดในสถานการณ์ใดก็ถือว่าตัวสถิติทดสอบนั้นมีความเหมาะสมในสถานการณ์นั้น

1.8 ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่คาดว่าจะได้รับในการวิจัยครั้งนี้คือ

1. เพื่อเป็นแนวทางในการเลือกตัวสถิติทดสอบที่เหมาะสมมาใช้ในการทดสอบเทียบความกลมกลืนสำหรับตัวแบบถดถอยในแต่ละสถานการณ์ในการใช้ในทางปฏิบัติ

2. เพื่อเป็นแนวทางในการศึกษาเปรียบเทียบการทดสอบเทียบความกลมกลืนภายใต้สถานการณ์ต่าง ๆ ต่อไป