

การเปรียบเทียบรูปแบบฟังก์ชันสมาชิกแบบแซดและแบบระฆังซีกขวา

สำหรับการวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์

A Comparison on Membership Functions between

Z-Shape and Simi-Right Bell Shape

in Robust Fuzzy Discriminant Analysis.

ชุติมา ห้องสวัสดิ์ และแสงหล้า ชัยมงคล

ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ปทุมธานี 12121

บทคัดย่อ

วัตถุประสงค์ของการวิจัยครั้งนี้ คือ เปรียบเทียบความคลุมเครือของการจำแนกกลุ่มสำหรับการวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์ที่มีการใช้ฟังก์ชันสมาชิกแบบแซดและแบบระฆังซีกขวา โดยที่ความคลุมเครือในการจำแนกกลุ่มจะพิจารณาจากค่าเอนโทรปีแบ่งส่วน การศึกษาจะใช้การจำลองข้อมูลแบบมอนติคาร์โลในกรณีที่มีข้อมูลประกอบด้วยประชากร 2 กลุ่มและ 3 กลุ่ม โดยที่ในแต่ละกลุ่มจะกำหนดให้มีขนาดตัวอย่างเท่ากับ 10 และ 30 มีตัวแปรอิสระเป็น 3 และ 10 แต่ละกลุ่มกำหนดให้มีข้อมูลผิดปกติจำนวน 1 และ 3 ค่า การวิเคราะห์จะใช้ขั้นตอนวิธีเจเนติกโดยใช้โปรแกรมสำเร็จรูป MATLAB ในแต่ละกรณีจะทำซ้ำ 100 ครั้งและคำนวณค่าเฉลี่ยของค่าเอนโทรปีแบ่งส่วน เพื่อใช้ในการเปรียบเทียบประสิทธิภาพของรูปแบบฟังก์ชันสมาชิกที่ใช้ ผลการศึกษพบว่า รูปแบบฟังก์ชันสมาชิกแบบระฆังซีกขวาให้ค่าเฉลี่ยเอนโทรปีแบ่งส่วนต่ำกว่าฟังก์ชันสมาชิกแบบแซดในทุกกรณี ยกเว้นในกรณีที่จำนวนตัวอย่างเป็น 10 และมีจำนวนค่าผิดปกติในแต่ละกลุ่มเพียง 1 ค่า สรุปได้ว่า ฟังก์ชันสมาชิกแบบระฆังซีกขวาจึงเป็นรูปแบบที่เหมาะสมสำหรับการวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์มากกว่าฟังก์ชันสมาชิกแบบแซด

คำสำคัญ : การวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์ ฟังก์ชันภาวะสมาชิก ขั้นตอนวิธีเจเนติก เอนโทรปีแบ่งส่วน

Abstract

The objective of this research is to compare on the fuzziness of group classification in robust fuzzy discriminant analysis which uses the membership functions between Z-shape and Semi-Right bell shape. The classification fuzziness is measured by partition entropy. The data are generated by using Monte Carlo simulation. We arbitrarily define that the data contain of 2 and 3 groups and the sample size in each group are 10 and 30. There are 3 and 10 independent variables and there are 1 and 3 outliers in each group. The genetic algorithm implemented by MATLAB software is used with 100 runs for each condition. Then we calculate the mean of the partition entropy which is used to determine the efficiency of the membership function. The finding of this study is the semi-right bell shape membership function produces a lower mean

of partition entropy for every condition than the Z-shape membership function does, except for the case that the sample size is 10 and the outlier is 1. Thus, we conclude that the semi-right bell shape is more suitable for the membership function in the robust fuzzy discriminant analysis than the Z-shape.

Keywords: Robust Fuzzy Discriminant Analysis, Membership Function, Genetic Algorithm, Partition Entropy

1. คำนำ

การวิเคราะห์จำแนกประเภทมีวัตถุประสงค์เพื่อจัดกลุ่มให้กับหน่วยสังเกต โดยพิจารณาจากตัวแปรอื่น ๆ ที่มีผลกับการแบ่งกลุ่ม หรือที่เรียกว่าตัวแปรอิสระ ตัวแปรอิสระเหล่านี้ โดยทั่วไปจะเป็นตัวแปรเชิงปริมาณ วิธีการจัดกลุ่มหน่วยสังเกตอาจใช้วิธีการควรจะเป็นสูงสุด (Maximum Likelihood methods) ฟังก์ชันระยะห่าง (Distance function) หรือ ฟังก์ชันจำแนกประเภทเชิงเส้น (Linear classification function) โดยมีข้อกำหนดว่าหน่วยสังเกตใดๆ จะต้องถูกจัดให้อยู่ในกลุ่มใดกลุ่มหนึ่งเพียงกลุ่มเดียวเท่านั้น แต่ในบางสถานการณ์ หน่วยสังเกตบางหน่วยอาจไม่สามารถระบุกลุ่มได้อย่างแน่ชัด ดังนั้นจึงได้นำแนวความคิดของฟัซซีเซตมาใช้ในการวิเคราะห์จำแนกประเภท ที่เรียกว่าการวิเคราะห์จำแนกประเภทฟัซซี [3] ซึ่งวิธีการนี้มีแนวคิดที่ว่า หน่วยสังเกตใดๆ สามารถเป็นสมาชิกของหลายกลุ่มพร้อมกันได้ด้วยระดับความเป็นสมาชิกที่แตกต่างกันในแต่ละกลุ่ม ในปี ค.ศ. 2004 Chang-Chun Lin and An-Pin Chen [3] ได้นำขั้นตอนวิธีเจเนติกมาช่วยในการหาระดับความเป็นสมาชิกในแต่ละกลุ่มของการวิเคราะห์จำแนกประเภทฟัซซี โดยการใช้ระยะห่างมาฮาလာโนบิสเป็นเกณฑ์ในการวิเคราะห์ ในกรณีที่หน่วยสังเกตใดที่มีค่าระดับความเป็นสมาชิกของทุกกลุ่มเป็นศูนย์ แสดงว่าหน่วยสังเกตนั้นไม่ได้เป็นสมาชิกของกลุ่มใดๆ จึงถือว่าหน่วยนั้น เป็นหน่วยที่ผิดปกติ หรือเรียกว่าค่าผิดปกติ (Outlier)

ในกรณีที่ค่าผิดปกติการใช้ระยะห่างมาฮาလာโนบิสในการคำนวณระดับความเป็นสมาชิกของแต่ละกลุ่มอาจมีความคลาดเคลื่อน จึงมีผู้นำแนวคิดเกี่ยวกับตัว

ประมาณค่าแบบโรบัสมาประยุกต์ใช้กับการคำนวณระยะห่างมาฮาလာโนบิส [7] โดยใช้ค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมแบบโรบัสแทนค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมแบบเดิม [2] วิธี FAST-MCD [5] เป็นวิธีประมาณค่าแบบโรบัสที่มีประสิทธิภาพสูงและใช้เวลาในการคำนวณน้อย และเป็นวิธีที่นิยมใช้กันอย่างแพร่หลายในปัจจุบัน วิธีการคำนวณของ FAST-MCD จะเกี่ยวข้องกับตัวสถิติอันดับ ค่าดีเทอร์มิแนนต์และใช้เทคนิคการเลือกซ้อนซ้ำ (Selection iteration) กับเทคนิคการขยายซ้อน (Nested extensions) โดยจะเลือกกลุ่มประชากรของหน่วยสังเกตที่ให้ค่าดีเทอร์มิแนนต์ของความแปรปรวนร่วมต่ำสุด หลังจากนั้นจะคำนวณค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมจากกลุ่มประชากรที่เลือกได้ และเรียกค่าเหล่านี้ว่า ค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมแบบโรบัส ตามลำดับ

รูปแบบฟังก์ชันสมาชิกฟัซซีเซตมีหลายรูปแบบ [1] เช่น ฟัซซีเซตแบบสี่เหลี่ยมคางหมู ฟัซซีเซตแบบสามเหลี่ยม ฟัซซีเซตแบบเอสและฟัซซีเซตแบบประจักษ์ เป็นต้น การใช้รูปแบบที่ต่างกันอาจจะทำให้ได้ผลลัพธ์ที่แตกต่าง มีผู้ศึกษาอิทธิพลของรูปแบบฟังก์ชันสมาชิกฟัซซีในระบบฟัซซีต่างๆ ตัวอย่างเช่น ในปี ค.ศ. 2001 Sanya Mitaim and Bart Kosko [4] ได้ศึกษาผลกระทบของรูปแบบฟังก์ชันสมาชิกฟัซซีเซตที่มีต่อการดำเนินการระบบฟัซซีในการประมาณฟังก์ชันต่อเนื่องทางวิศวกรรมพบว่าฟังก์ชันสมาชิกฟัซซีเซตแต่ละแบบมีประสิทธิภาพแตกต่างกันขึ้นอยู่กับขอบเขตที่ต้องการศึกษาหรือเป้าหมายที่กำหนด เพื่อวิเคราะห์ปัญหาที่น่าสนใจให้ได้ผลลัพธ์ที่ดีที่สุด อย่างไรก็ตามยังไม่มีผู้ทำการศึกษาอิทธิพลของรูปแบบฟังก์ชันสมาชิกฟัซซีในการวิเคราะห์

จำแนกประเภทฟัซซี ดังนั้นงานวิจัยนี้ ผู้วิจัยจึงสนใจที่จะเปรียบเทียบรูปแบบฟังก์ชันสมาชิกแบบแซด (Z-shape) และแบบระฆังซีกขวา (Semi-right bell shape) ที่ใช้ในการวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์ โดยมีวัตถุประสงค์ที่จะเลือกรูปแบบฟังก์ชันสมาชิกที่ให้ค่าเอ็นโทรปีแบ่งส่วน (Partition Entropy) หรือให้ค่าความคลุมเครือในการแบ่งกลุ่มน้อยที่สุด

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ขั้นตอนวิธีเจเนติก (Genetic Algorithm)

ขั้นตอนวิธีเจเนติกเป็นวิธีการค้นหาคำตอบโดยมีพื้นฐานมาจากกระบวนการคัดเลือกทางธรรมชาติ (Natural Selection) ที่ฮอลแลนด์และคณะ (Holland et al.) ได้พัฒนาขึ้นโดยมีวัตถุประสงค์เพื่อประยุกต์ใช้กับปัญหาการหาคำตอบที่ดีที่สุด (Optimization problem) วิธีนี้เป็นวิธีการที่สามารถใช้แก้ปัญหาที่คลุมเครือได้และคำตอบที่ได้จะเป็นคำตอบที่ใกล้เคียงคำตอบที่ดีที่สุดโดยใช้เวลาและค่าใช้จ่ายในการหาคำตอบไม่มากนัก ขั้นตอนวิธีเจเนติกประกอบด้วยขั้นตอนการดำเนินการเบื้องต้น 6 ขั้นตอน [6] ดังนี้

- 1) การเข้ารหัสและสร้างประชากรเริ่มต้น (Initialization) เป็นการเข้ารหัสหรือแปลงค่าพารามิเตอร์ให้อยู่ในรูปของสตริงที่มีความยาวแน่นอน ซึ่งวิธีการเข้ารหัสนี้ขึ้นอยู่กับรูปแบบของปัญหาแต่ละปัญหา
- 2) การประเมินค่า (Evaluation) จะประเมินค่าความเหมาะสม (Fitness value) จากค่าฟังก์ชันเป้าหมาย (Objective function) ที่ขึ้นอยู่กับชนิดของปัญหา ในกรณีการวิเคราะห์จำแนกประเภทฟัซซีฟังก์ชันเป้าหมาย คือ ผลรวมของความคลาดเคลื่อนกำลังสองในการจำแนกประเภท
- 3) การดำเนินการรีโพรดักชัน (Reproduction) คือ กระบวนการที่สร้างแต่ละตัว ทำการเลียนแบบฟังก์ชันเป้าหมายเพื่อหาค่าความเหมาะสมซึ่งสร้าง

ที่มีความเหมาะสมสูงกว่าจะมีความน่าจะเป็นสูงที่จะนำไปใช้ในการสร้างลูกหลานในรุ่นต่อไป

- 4) การดำเนินการครอสโอเวอร์ (Crossover) คือ การจับคู่สมาชิกในเมตติงพูล (Mating Pool) และทำการไขว้สลับค่าที่อยู่หลังตำแหน่งที่เลือกไว้จากการสุ่ม
- 5) การดำเนินการมิวเตชัน (Mutation) คือ การเปลี่ยนแปลงค่าแต่ละตำแหน่งในสตริงแบบสุ่ม
- 6) การตรวจสอบเงื่อนไข (Termination test) เพื่อใช้เป็นจุดสิ้นสุดกระบวนการ

2.2 การวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์

(Fuzzy Robust Discriminant Analysis)

การวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์เป็นการจำแนกประเภทที่พิจารณาจากระดับความเป็นสมาชิกในแต่ละกลุ่ม โดยการประยุกต์ใช้ตัวประมาณค่าแบบโรบัสต์ในการคำนวณระยะห่างมาฮาลานอบิส นั่นคือ ใช้ค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมแบบโรบัสต์ที่คำนวณจากวิธี FAST-MCD แทนค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมแบบเดิม ซึ่งการวิเคราะห์จำแนกประเภทฟัซซีแบบโรบัสต์ประกอบด้วยขั้นตอนดังต่อไปนี้

1) คำนวณหาระยะห่างมาฮาลานอบิส

1.1) คำนวณหาค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมแบบโรบัสต์ โดยวิธี FAST-MCD ซึ่งมีแนวคิดพื้นฐานที่เกี่ยวข้องกับสถิติลำดับและดีเทอร์มิแนนต์ โดยใช้เทคนิคการเลือกย่อนและการขยายซ้อนใน โดยมีวัตถุประสงค์ คือ หาค่าหน่วยสังเกต h ค่าจากทั้งหมด n ค่าที่มีเมทริกซ์ความแปรปรวนร่วมที่ทำให้ดีเทอร์มิแนนต์ของความแปรปรวนร่วมต่ำสุด แล้วนำหน่วยสังเกต h คำนวณหาค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วม

1.2) คำนวณหาระยะห่างมาฮาลานอบิสจากสมการ

$$D_{ij}^2 = (x_j - c_i)' G(x_j - c_i)$$

เมื่อ D_{ij}^2 แทน กำลังสองของระยะห่างในกลุ่มที่ i หน่วยที่ j

$\underline{x}_j$ แทน เวกเตอร์ขนาด $(p \times 1)$ ของตัวแปรอิสระ p

ตัวแปรของ หน่วยที่ j

C_i แทน เวกเตอร์ขนาด (px1) ของค่าเฉลี่ยของตัว

แปร p ตัวของกลุ่มที่ i

G แทน เมทริกซ์ผกผันของเมทริกซ์ความแปร-

ปรวนร่วมของทุกกลุ่มขนาด (pxp) ที่มี

คุณสมบัติสมมาตร (Symmetric) และ

เป็นบวกแน่นอน (Positive Definite)

โดยที่ i แทน ค่าของกลุ่ม ซึ่ง $i = 1, 2, 3, \dots, m$ และ

m เป็นจำนวนกลุ่มทั้งหมด

j แทน หน่วยสังเกต ซึ่ง $j = 1, 2, 3, \dots, n$ และ n

เป็นจำนวนหน่วยสังเกตทั้งหมด

เมื่อค่า C_i และ G เป็นเวกเตอร์ค่าเฉลี่ยและเมทริกซ์ความแปรปรวนร่วมแบบโรบัสต์

2) ใช้ขั้นตอนวิธีเจเนติกในการคำนวณหา

ขอบเขตความคลุมเครือของฟังก์ชันสมาชิกโดยพิจารณาจากค่าผลรวมของความคลาดเคลื่อนกำลังสองในการจำแนกประเภท

$$E_i = \sum_{j=1}^n (\mu_i(D_{ij}) - y_{ij})^2$$

โดยจะเลือกขอบเขตความคลุมเครือที่ให้ค่าผลรวมของความคลาดเคลื่อนกำลังสองในการจำแนกประเภทต่ำสุดมาใช้ในการคำนวณระดับความเป็นสมาชิกต่อไป

3) คำนวณค่าระดับความเป็นสมาชิกของหน่วยสังเกตทุกค่าในแต่ละกลุ่ม ($\mu_i(D_{ij})$) จะพิจารณาระดับความเป็นสมาชิกของหน่วยสังเกตแต่ละค่า เพื่อจัดหน่วยสังเกตไปยังกลุ่มของประชากร โดยจะจัดให้หน่วยสังเกตอยู่ในกลุ่มที่ k ถ้า $\mu_k(D_{ij}) > \mu_i(D_{ij})$ เมื่อ $i = 1, 2, \dots, m$ และ $i \neq k$ เมื่อ m แทน จำนวนกลุ่มทั้งหมด แต่ถ้า $\mu_i(D_{ij}) = 0$ ทุกค่า i แล้วหน่วยสังเกตนั้นจะถูกกำหนดให้เป็นค่าผิดปกติ

2.3 รูปแบบฟังก์ชันสมาชิก (Membership Function)

รูปแบบฟังก์ชันสมาชิกมีหลายรูปแบบ [4] ในที่นี้จะเสนอเฉพาะรูปแบบที่เกี่ยวข้องกับการวิจัย โดยกำหนดให้ D_{ij} แทน รากที่สองของระยะห่างมาฮาลาโน

บิส a, b และ c คือ ค่าขอบเขตของฟังก์ชันสมาชิกที่ได้จากขั้นตอนวิธีเจเนติก โดยค่าระดับความเป็นสมาชิก ($\mu_i(D_{ij})$) ของรูปแบบฟังก์ชันสมาชิกรูปแบบต่างๆ สามารถเขียนได้ดังสมการข้างล่าง ดังต่อไปนี้

1.) รูปแบบฟังก์ชันสมาชิกแบบแซด

$$\mu_i(D_{ij}) = \begin{cases} 1 & ; D_{ij} \leq a \\ (b - D_{ij}) / (b - a) & ; a < D_{ij} \leq b \\ 0 & ; D_{ij} > b \end{cases}$$

2.) รูปแบบฟังก์ชันสมาชิกแบบเอส

$$\mu_i(D_{ij}) = \begin{cases} 0 & ; D_{ij} \leq a \\ 2[(D_{ij} - a) / (c - a)]^2 & ; a < D_{ij} \leq b \\ 1 - 2[(D_{ij} - a) / (c - a)]^2 & ; b < D_{ij} \leq c \\ 1 & ; D_{ij} > c \end{cases}$$

3.) รูปแบบฟังก์ชันสมาชิกแบบระฆัง

$$\mu_i(D_{ij}) = \begin{cases} S(D_{ij}; c - b, c - b/2, c) & ; D_{ij} \leq c \\ 1 - S(D_{ij}; c, c + b/2, c + b) & ; D_{ij} > c \end{cases}$$

เมื่อ $S(.)$ คือ รูปแบบฟังก์ชันสมาชิกแบบเอส

4.) รูปแบบฟังก์ชันสมาชิกแบบระฆังซิกวา

ค่า $\mu_i(D_{ij})$ ของฟังก์ชันสมาชิกแบบระฆังซิก

วาจะถูกกำหนดโดยผู้วิจัย โดยการใช้ค่า $\mu_i(D_{ij})$ ของรูปแบบฟังก์ชันสมาชิกแบบเอสแทนใน $\mu_i(D_{ij})$ ของรูปแบบฟังก์ชันสมาชิกแบบระฆัง ดังนั้น ค่า $\mu_i(D_{ij})$ ของรูปแบบฟังก์ชันสมาชิกแบบระฆังซิกวาสามารถเขียนได้ดังนี้

$$\mu_i(D_{ij}) = \begin{cases} 1 & ; D_{ij} \leq a \\ 1 - 2[(D_{ij} - a) / (c - a)]^2 & ; a < D_{ij} \leq b \\ 2[(D_{ij} - a) / (c - a)]^2 & ; b < D_{ij} \leq c \\ 0 & ; D_{ij} > c \end{cases}$$

3. วิธีดำเนินการวิจัย

3.1 จำลองข้อมูล (Simulation Data) ใช้เทคนิคมอนติคาร์โลในการจำลองข้อมูล โดยใช้โปรแกรมสำเร็จรูป MATLAB โดยในแต่ละกรณีจะทำการจำลองข้อมูลจำนวน 100 ชุด โดยกำหนดเงื่อนไขของการศึกษาดังต่อไปนี้

- 1.) กลุ่มของประชากรเป็น 2 และ 3 กลุ่ม
- 2.) ขนาดตัวอย่างในแต่ละกลุ่มเท่ากับ 10 และ 30
- 3.) จำนวนตัวแปรอิสระเป็น 3 และ 10 ตัวแปร
- 4.) กรณีที่มีประชากร 2 กลุ่มเวกเตอร์ค่าเฉลี่ยของตัวแปรอิสระของประชากรกลุ่มที่ 1 และ 2 มีค่าเท่ากับ 5 และ 10 ตามลำดับ กรณีที่มีประชากร 3 กลุ่มเวกเตอร์ค่าเฉลี่ยของตัวแปรอิสระของประชากรกลุ่มที่ 1, 2 และ 3 มีค่าเท่ากับ 5, 10 และ 15 ตามลำดับ
- 5.) สหสัมพันธ์ระหว่างตัวแปรอิสระของแต่ละกลุ่มมีค่าเท่ากับ 0.1
- 6.) ค่าผิดปกติมีจำนวนเท่ากับ 1 และ 3 ค่า
- 7.) ค่าผิดปกติมีเวกเตอร์ค่าเฉลี่ยเป็น 5 เท่าของค่าปกติ

3.2 คำนวณค่า D_{ij}^2 โดยใช้วิธี FAST-MCD [5]

3.3 ใช้ขั้นตอนวิธีเจเนติกคำนวณหาขอบเขตความคลุมเครือของฟังก์ชันสมาชิกที่อยู่ในรูปฟัซซี่เซต โดยพิจารณาจากค่าผลรวมของความคลาดเคลื่อนกำลังสองในการจำแนกประเภทหรือฟังก์ชันเป้าหมาย

3.4 เลือกขอบเขตของฟังก์ชันสมาชิกที่ให้ค่าผลรวมของความคลาดเคลื่อนกำลังสองในการจำแนกประเภทหรือฟังก์ชันเป้าหมายต่ำสุดมาเพื่อใช้ในการคำนวณหาระดับความเป็นสมาชิกหรือโอกาสที่หน่วยสังเกตจะเป็นสมาชิกในแต่ละกลุ่ม

3.5 คำนวณค่าเอ็นโทรปีแบ่งส่วน

$$H(\tilde{U}, m) = -\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^m \mu_{ij} \ln(\mu_{ij})$$

โดยที่ $H(\tilde{U}, m)$ แทน ค่าเอ็นโทรปีแบ่งส่วน

μ_{ij} แทน ระดับความน่าจะเป็นที่หน่วยสังเกตที่ j จะอยู่ในกลุ่มที่ i

m แทน จำนวนกลุ่ม

n แทน จำนวนหน่วยสังเกตทั้งหมด

3.6 เปรียบเทียบค่าเอ็นโทรปีแบ่งส่วน โดยที่รูปแบบฟังก์ชันสมาชิกที่ให้ค่าเอ็นโทรปีแบ่งส่วนต่ำกว่าจะเป็นรูปแบบที่เหมาะสมสำหรับการวิเคราะห์จำแนกประเภทฟัซซี่แบบโรบัสต์มากกว่า เนื่องจากเป็นรูปแบบที่มีความคลุมเครือในการจำแนกประเภทน้อยกว่า

4. ผลการวิจัย

จากตารางที่ 1 สามารถสรุปผลได้ดังต่อไปนี้

- 1.) ในกรณีที่ขนาดตัวอย่างในแต่ละกลุ่มมีขนาดเท่ากับ 10 และแต่ละกลุ่มมีค่าผิดปกติ 1 ค่า แล้วค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วนของฟังก์ชันสมาชิกแบบบรีซังซิกวาจะมีค่าสูงกว่าค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วนของฟังก์ชันสมาชิกแบบแซด
- 2.) ในกรณีที่ขนาดตัวอย่างในแต่ละกลุ่ม มีขนาดเท่ากับ 30 ค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วนของฟังก์ชันสมาชิกแบบบรีซังซิกวา จะมีค่าต่ำกว่าค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วนของฟังก์ชันสมาชิกแบบแซดในทุกกรณี ยกเว้นในกรณีที่จำนวนกลุ่มของประชากร และจำนวนตัวแปรอิสระเป็น 3 และมีค่าผิดปกติในแต่ละกลุ่มเท่ากับ 1
- 3.) ในกรณีที่ขนาดตัวอย่างในแต่ละกลุ่มเป็น 10 และมีจำนวนตัวแปรอิสระ 10 ตัว ค่าเฉลี่ยเอ็นโทรปีแบ่งส่วนของฟังก์ชันสมาชิกแบบแซด และแบบบรีซังซิกวา จะมีค่าสูงกว่ากรณีอื่นๆ ไม่ว่าจะเป็นค่าผิดปกติ 1 หรือ 3 ค่า

5. สรุปผลการวิจัย

ฟังก์ชันสมาชิกแบบบรีซังซิกวาจะให้ค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วนต่ำกว่าฟังก์ชันสมาชิกแบบแซดในทุกกรณี ยกเว้นในกรณีที่จำนวนตัวอย่างและค่าผิดปกติในแต่ละกลุ่มน้อย โดยที่จำนวนกลุ่มของประชากร ไม่มีผลต่อ

ค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วน นอกจากนี้ยังสามารถสรุปได้ว่า ค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วนจะมีค่ามาก ถ้ามีจำนวนตัวแปรอิสระมากแต่มีขนาดตัวอย่างในแต่ละกลุ่มน้อย ทั้งนี้เป็นเพราะขนาดตัวอย่าง และจำนวนตัวแปรอิสระไม่มีความสมดุลกัน โดยทั่วไปขนาดตัวอย่างในแต่ละกลุ่มควรจะแปรผันตามจำนวนตัวแปรอิสระ

สรุปได้ว่า รูปแบบฟังก์ชันสมาชิกแบบระฆังชีกขวาให้ค่าเฉลี่ยเอ็นโทรปีแบ่งส่วนต่ำกว่า ฟังก์ชันสมาชิกแบบแซดเกือบทุกกรณีที่ศึกษา นั่นคือ ฟังก์ชันสมาชิกแบบระฆังชีกขวาเป็นรูปแบบที่เหมาะสมสำหรับการวิเคราะห์จำแนกประเภทฟัซซี่แบบโรบัสตมากกว่าฟังก์ชันสมาชิกแบบแซด

ตารางที่ 1 ค่าเฉลี่ยของเอ็นโทรปีแบ่งส่วนของรูปแบบฟังก์ชันสมาชิกแบบแซดและแบบระฆังชีกขวา

จำนวนกลุ่ม	จำนวนตัวแปรอิสระ	ขนาดตัวอย่างในแต่ละกลุ่ม	จำนวนค่าผิดปกติในแต่ละกลุ่ม	ค่าเฉลี่ยเอ็นโทรปีแบ่งส่วน	
				ฟังก์ชันสมาชิกแบบแซด	ฟังก์ชันสมาชิกแบบระฆังชีกขวา
2	3	10	1	0.00770	0.01080
			3	0.00460	0.00013
			1	0.00760	0.00370
	10	30	3	0.00600	0.00510
			1	0.01630	0.01070
			3	0.01710	0.01340
		10	1	0.00360	0.00340
			3	0.00670	0.00340
			3	0.00670	0.00340
3	3	10	1	0.00370	0.00470
			3	0.00190	0.00045
			1	0.00410	0.00610
	10	30	3	0.00570	0.00320
			1	0.01660	0.01940
			3	0.01840	0.01550
		10	1	0.00860	0.00650
			3	0.00380	0.00250
			3	0.00380	0.00250

6.เอกสารอ้างอิง

- [1] Chi, Z.; Yan, H., and Pham, T., Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition., NJ: World Scientific, 225 p, 1996.
- [2] Hubert, M., and Van Driessen, K., Fast and Robust Discriminant Analysis, Computational Statistic & Data Analysis. Vol. 45 ; pp. 301-320, 2004
- [3] Lin, C. C., and Chen, A. P., Fuzzy Discriminant Analysis with Outlier Detection by Genetic Algorithm, Computers & Operations. Vol. 31 ; pp. 877-888, 2004.
- [4] Mitaim, S., and Kosko, B., The Shape of Fuzzy Sets in Adaptive Function Approximation, IEEE Transactions on Fuzzy System. Vol. 9 ; pp. 637-656, 2001.
- [5] Rousseeuw, P. J. and Van Driessen, K., A Fast Algorithm for the Minimum Covariance Determinant Estimator, Technometrics. Vol. 41 ; pp.212-223, 1999.
- [6] Sakawa, M., Genetic Algorithms and Fuzzy Multiobjective Optimization., Boston: Kluwer, 288 p, 2002.
- [7] Santos-Pereira, C. M., and Pires, A. M., Detection of Outliers in Multivariate Data: A method based on Clustering and Robust Estimators, Downloadable from <http://citeseer.ist.psu.edu/545222.html>. [2006,Nov.30]