

**ชื่อเรื่องวิทยานิพนธ์** การเปรียบเทียบความสามารถในการจำแนกกลุ่มได้ถูกต้องระหว่างวิธีการวิเคราะห์จำแนกกลุ่มกับวิธีการวิเคราะห์การถดถอยลอจิสติก เมื่อใช้กลุ่มตัวอย่างขนาดต่าง ๆ

**ชื่อผู้เขียน** นางสาวเดือนเพ็ญ สนโต

**ศึกษาศาสตรมหาบัณฑิต** สาขาวิชาวิจัยและสถิติการศึกษา

**คณะกรรมการสอบวิทยานิพนธ์**

รองศาสตราจารย์ ดร. ด้าย เชื้องณี

ประธานกรรมการ

รองศาสตราจารย์ อุเทน ปิญโญ

กรรมการ

อาจารย์ ดร. เกียรติสุดา ศรีสุข

กรรมการ

### บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อศึกษาความสามารถในการจำแนกกลุ่มได้ถูกต้องจากวิธีการวิเคราะห์จำแนกกลุ่มกับวิธีการวิเคราะห์การถดถอยลอจิสติก เมื่อใช้กลุ่มตัวอย่างขนาดต่าง ๆ และเพื่อเปรียบเทียบความสามารถในการจำแนกกลุ่มได้ถูกต้องระหว่างวิธีการวิเคราะห์จำแนกกลุ่มกับวิธีการวิเคราะห์การถดถอยลอจิสติก เมื่อใช้กลุ่มตัวอย่างขนาดต่าง ๆ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองด้วยเทคนิคมอนติคาร์โล โดยใช้โปรแกรมเขียนด้วยภาษาเบสิกสุ่มตัวเลขสุ่มเพื่อสร้างประชากรจำนวน 4 กลุ่ม กลุ่มละ 10,000 จำนวน ซึ่งลักษณะของประชากรแต่ละกลุ่มจะประกอบด้วย ตัวแปรอิสระจำนวน 4 ตัวแปร โดยตัวแปรอิสระของแต่ละกลุ่มมีลักษณะการแจกแจงแบบปกติและเมตริกซ์ความแปรปรวนร่วมของตัวแปรอิสระแต่ละกลุ่มเท่ากันและตัวแปรตามจำนวน 1 ตัวแปร แบ่งเป็น 2 กลุ่ม 3 กลุ่ม 4 กลุ่ม และ 5 กลุ่มตามลำดับโดยกำหนดให้ประชากรแต่ละกลุ่มมีความสามารถในการจำแนกกลุ่มได้ถูกต้องประมาณ 80 เปอร์เซ็นต์ และใช้โปรแกรม SPSS for Windows เพื่อสุ่มกลุ่มตัวอย่างขนาดต่าง ๆ คือ 40 80 120 และ 160 และคำนวณด้วยสถิติ 2 วิธี คือวิธีวิเคราะห์จำแนกกลุ่มและวิธีวิเคราะห์การถดถอยลอจิสติก โดยสุ่มตัวอย่างซ้ำจำนวน 500 ครั้ง เพื่อเปรียบเทียบผลการคำนวณทั้ง 2 วิธี ผลการวิจัยสรุปได้ดังนี้

1. ผลของความสามารถในการจำแนกกลุ่มได้ถูกต้องจากวิธีวิเคราะห์จำแนกกลุ่มกับวิธีวิเคราะห์การถดถอยลอจิสติก พบว่า

1.1 เมื่อประชากรประกอบด้วยตัวแปรอิสระ 4 ตัวแปรตัวแปรตาม 1 ตัวแปรแบ่งเป็น 2 กลุ่ม ผลการจำแนกกลุ่มได้ถูกต้องเมื่อสุ่มกลุ่มตัวอย่างขนาด 40 80 120 และ 160 จากวิธีการวิเคราะห์จำแนกกลุ่มโดยเฉลี่ยร้อยละ 79.75 79.91 80.14 และ 80.13 ตามลำดับ และจากวิธีการ

วิเคราะห์การถดถอยลอจิสติกแบบ Binary Logistic โดยเฉลี่ยร้อยละ 79.76 79.53 79.98 และ 79.90 ตามลำดับ

1.2 เมื่อประชากรประกอบด้วยตัวแปรอิสระ 4 ตัวแปร ตัวแปรตาม 1 ตัวแปรแบ่งเป็น 3 กลุ่ม ผลการจำแนกกลุ่มได้ถูกต้อง เมื่อสุ่มกลุ่มตัวอย่างขนาด 40 80 120 และ 160 จากวิธีการวิเคราะห์จำแนกกลุ่มโดยเฉลี่ยร้อยละ 69.27 67.29 66.44 และ 67.03 ตามลำดับ และจากวิธีการวิเคราะห์การถดถอยลอจิสติกแบบ Multinomial Logistic โดยเฉลี่ยร้อยละ 78.29 76.84 76.27 และ 76.53 ตามลำดับ

1.3 เมื่อประชากรประกอบด้วยตัวแปรอิสระ 4 ตัวแปร ตัวแปรตาม 1 ตัวแปรแบ่งเป็น 4 กลุ่ม ผลการจำแนกกลุ่มได้ถูกต้อง เมื่อสุ่มกลุ่มตัวอย่างขนาด 40 80 120 และ 160 จากวิธีการวิเคราะห์จำแนกกลุ่มโดยเฉลี่ยร้อยละ 70.72 67.03 66.00 และ 65.57 ตามลำดับ และจากวิธีการวิเคราะห์การถดถอยลอจิสติกแบบ Multinomial Logistic โดยเฉลี่ยร้อยละ 79.65 76.71 76.77 และ 76.58 ตามลำดับ

1.4 เมื่อประชากรประกอบด้วยตัวแปรอิสระ 4 ตัวแปร ตัวแปรตาม 1 ตัวแปรแบ่งเป็น 5 กลุ่ม ผลการจำแนกกลุ่มได้ถูกต้อง เมื่อสุ่มกลุ่มตัวอย่างขนาด 40 80 120 และ 160 จากวิธีการวิเคราะห์จำแนกกลุ่มโดยเฉลี่ยร้อยละ 75.59 71.22 69.25 และ 68.69 ตามลำดับ และจากวิธีการวิเคราะห์การถดถอยลอจิสติกแบบ Multinomial Logistic โดยเฉลี่ยร้อยละ 84.01 78.75 77.49 และ 77.30 ตามลำดับ

2. ผลการเปรียบเทียบความสามารถในการจำแนกกลุ่มได้ถูกต้อง ระหว่างวิธีการวิเคราะห์จำแนกกลุ่มกับวิธีการวิเคราะห์การถดถอยลอจิสติก เมื่อสุ่มกลุ่มตัวอย่างขนาด 40 80 120 และ 160 พบว่า

2.1 เมื่อประชากรประกอบด้วยตัวแปรอิสระ 4 ตัวแปร ตัวแปรตาม 1 ตัวแปรแบ่งเป็น 2 กลุ่ม เมื่อสุ่มกลุ่มตัวอย่างขนาด 40 วิธีการวิเคราะห์จำแนกกลุ่มกับวิธีการวิเคราะห์การถดถอยลอจิสติกสามารถจำแนกกลุ่มได้ถูกต้องไม่แตกต่างกัน เมื่อสุ่มกลุ่มตัวอย่างขนาด 80 120 และ 160 วิธีการวิเคราะห์จำแนกกลุ่มสามารถจำแนกกลุ่มได้ถูกต้องมากกว่าวิธีการวิเคราะห์การถดถอยลอจิสติกแบบ Binary Logistic อย่างมีนัยสำคัญที่ระดับ .01

2.2 เมื่อประชากรประกอบด้วยตัวแปรอิสระ 4 ตัวแปร ตัวแปรตาม 1 ตัวแปรแบ่งเป็น 3 กลุ่ม 4 กลุ่ม และ 5 กลุ่ม จะพบว่า เมื่อสุ่มกลุ่มตัวอย่างขนาด 40 80 120 และ 160 วิธีการวิเคราะห์การถดถอยลอจิสติกแบบ Multinomial Logistic สามารถจำแนกกลุ่มได้ถูกต้องมากกว่าวิธีการวิเคราะห์จำแนกกลุ่ม อย่างมีนัยสำคัญที่ระดับ .01

<b>Thesis Title</b>	Comparison of Correct Group Classification Abilities Between Discriminant Analysis and Logistic Regression Analysis Methods with Various Sample Sizes	
<b>Author</b>	Miss Duanphen Sonto	
<b>M.Ed.</b>	Research and Statistics in Education	
<b>Examining Committee</b>	Assoc. Prof. Dr. Tay Chiengchee	Chairman
	Assoc. Prof. Uthen Panyo	Member
	Lect. Dr. Kaitsuda Srisuk	Member

### Abstract

The purpose of this research was to study abilities of correct group classification from Discriminant Analysis and Logistic Regression Analysis methods with various sample sizes and to comparison of correct group classification abilities between Discriminant Analysis and Logistic Regression Analysis methods with various sample sizes.

This experimental research was conducted by using Monte-Carlo Method with basic language writing programe to random numbers for creation of four groups of population. Each group contained 10,000 numbers. Characteristic of each consisted of 4 independent variables. The independent variables of each group were normal distribution, and their covariance matrixes were all equal. One dependent variable was divided into 2 groups 3 groups 4 groups and 5 groups respectively. Each group was assigned to have correct group classification abilities of 80 percent. SPSS for Windows programe was used to sample various sizes – 40, 80, 120 and 160. Then, they were calculated by 2 methods - Discriminant Analysis and Logistic Regression Analysis, and repeatedly random sampled 500 times to compare the results of these two methods. The findings were as follows:

1. The result of correct group classification abilities between Discriminant Analysis and Logistic Regression Analysis methods revealed that:

- 1.1 When population consisted of 4 independent variables, 1 dependent variable was divided into 2 groups with sample sizes of 40 80 120 and 160, the results of group classification from Discriminant Analysis were correct with the average percentage of 79.75 79.91 80.18 and

80.13 respectively, and those from Binary Logistic Regression Analysis were 79.76 79.53 79.98 and 79.90 respectively.

1.2 When population consisted of 4 independent variables, 1 dependent variable was divided into 3 groups with sample sizes of 40 80 120 and 160, the results of group classification from Discriminant Analysis were correct with the average percentage of 69.27 67.29 66.44 and 67.03 respectively, and those from Multinomial Logistic Regression Analysis were 78.29 76.84 76.27 and 76.53 respectively.

1.3 When population consisted of 4 independent variables, 1 dependent variable was divided into 4 groups with sample sizes of 40 80 120 and 160, the results of group classification from Discriminant Analysis were correct with the average percentage of 70.72 67.03 66.00 and 65.57 respectively, and those from Multinomial Logistic Regression Analysis were 79.65 76.71 76.77 and 76.58 respectively.

1.4 When population consisted of 4 independent variables, 1 dependent variable was divided into 5 groups with sample sizes of 40 80 120 and 160, the results of group classification from Discriminant Analysis were correct with the average percentage of 75.59 71.22 69.25 and 68.69 respectively, and those from Multinomial Logistic Regression Analysis were 84.01 78.75 77.49 and 77.30 respectively.

2. The results of correct group classification comparison between Discriminant Analysis and Logistic Regression Analysis methods with sample sizes of 40 80 120 and 160 were as follows:

2.1 When population consisted of 4 independent variables, 1 dependent variable was divided into 2 groups with the sample size of 40, the results of group classification from Discriminant Analysis and Binary Logistic Regression Analysis methods were correct, and not different. When the sample sizes were 80, 120 and 160, the results of group classification from Discriminant Analysis was more accurate than those from the Binary Logistic Regression Analysis at the significant level of .01.

2.2 When population consisted of 4 independent variables, 1 dependent variable was divided into 3, 4 and 5 groups with the sample sizes of 40 80 120 and 160, the results of group classification from Multinomial Logistic Regression Analysis was more accurate than those from Discriminant Analysis at the significant level of .01.