แผนบริหารการสอนประจำบทที่ 5

วิชา 4114601 การวิเคราะห์สถิติขั้นสูงด้วยโปรแกรมสำเร็จรูปทางสถิติ
แผนบริหารการสอน บทที่ 5 การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis)
เวลา 8 ชั่วโมง

สาระสำคัญ

การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis : DA) เป็นเทคนิคการวิเคราะห์ตัวแปรพหุ (Multivariate Analysis) ที่มีวัตถุประสงค์หลักเพื่อสร้างสมการจำแนกกลุ่มที่ดีที่สุดจากข้อมูล และนำสมการจำแนกกลุ่มมาใช้พยากรณ์หน่วยวิเคราะห์ใหม่ว่าสมควรจัดให้อยู่ในกลุ่มใด ซึ่งตัวแปรตาม (ตัวแปรที่ถูกจำแนกหรือถูกทำนาย) เป็นตัวแปรที่แบ่งเป็นกลุ่มๆ ตั้งแต่ 2 กลุ่มขึ้นไป มีระดับการวัดในมาตรานามบัญญัติ (Nominal Scale) หรือมาตราเรียงอันดับ (Ordinal Scale) และถ้าหากมีข้อมูลอยู่ในระดับอื่นให้แปลงข้อมูลเป็นตัวแปรเชิงกลุ่มก่อนที่จะนำไปวิเคราะห์ ส่วนตัวแปรอิสระ (ตัวแปรที่ใช้ในการจำแนกหรือตัวแปรที่ใช้ในการทำนาย) เป็นตัวแปรต่อเนื่องที่มีระดับการวัดอยู่ในมาตราอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) และถ้าหากมีข้อมูลอยู่ในระดับอื่นให้แปลงให้เป็นตัวแปรหุ่น (Dummy Coding) ก่อนที่จะนำไปวิเคราะห์

ผลการเรียนรู้ที่คาดหวัง

นักศึกษาสามารถเข้าใจลักษณะของตัวแปรที่นำมาวิเคราะห์จำแนกกลุ่มและทราบถึงข้อตกลงเบื้องต้นก่อนที่จะนำข้อมูลมาวิเคราะห์จำแนกกลุ่ม รวมทั้งมองเห็นข้อแตกต่างระหว่างเทคนิคการวิเคราะห์จำแนกกลุ่ม และ การวิเคราะห์การถดถอยโลจิสติค และสามารถแปลความหมายผลลัพธ์ที่ได้จากการวิเคราะห์จำแนกกลุ่มได้อย่างถูกต้อง

จุดประสงค์การเรียนรู้

1. เข้าใจหลักการของการวิเคราะห์จำแนกกลุ่ม
2. เปรียบเทียบความแตกต่างระหว่างเทคนิคการวิเคราะห์จำแนกกลุ่ม กับ การวิเคราะห์การถดถอยโลจิสติคได้
3. สามาถประยุกต์ใช้เทคนิค การวิเคราะห์จำแนกกลุ่ม ในงานด้านต่างๆได้
4. สามารถสรุปผลที่ได้จากการวิเคราะห์ด้วยเทคนิคการวิเคราะห์จำแนกกลุ่มได้อย่างถูกต้อง

กิจกรรมการเรียนการสอน

1. นำเสนอ PowerPoint เนื้อหาเกี่ยวกับการวิเคราะห์จำแนกกลุ่มจากเอกสารประกอบการสอนการวิเคราะห์สถิติขั้นสูงด้วยโปรแกรมสำเร็จรูปทางสถิติ
2. ให้นักศึกษาค้นคว้าข้อมูลงานวิจัยที่เลือกใช้การวิเคราะห์จำแนกกลุ่ม และ การวิเคราะห์การถดถอยโลจิสติคในงานวิจัยมาเปรียบเทียบกันและนำมาอภิปรายถึงความแตกต่างของผลลัพธ์ที่ได้ในชั้นเรียน
3. ให้นักศึกษาทำแบบฝึกหัดท้ายบท

สื่อการเรียนรู้

1. ตัวอย่างงานวิจัย
2. PowerPoint
3. เอกสารประกอบการสอนรายวิชา 4114601 การวิเคราะห์สถิติขั้นสูงด้วยโปรแกรมสำเร็จรูปทางสถิติ

การวัดและประเมินผล

1. ประเมินผลจากแบบฝึกหัด
2. ประเมินผลจากการอภิปรายร่วมกันในชั้นเรียน
3. ประเมินผลจากการทดสอบย่อยรายบท

บทที่ 5
การวิเคราะห์จำแนกกลุ่ม
(Discriminant Analysis)

การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis : DA) เป็นเทคนิคการวิเคราะห์ตัวแปรพหุ (Multivariate Analysis) ที่มีลักษณะของตัวแปรที่นำมาวิเคราะห์เหมือนกับการวิเคราะห์ถดถอยโลจิสติคคือ ตัวแปรตาม (ตัวแปรที่ถูกจำแนกหรือถูกทำนาย) เป็นตัวแปรที่แบ่งเป็นกลุ่มๆ ตั้งแต่ 2 กลุ่มขึ้นไป ส่วนตัวแปรอิสระ (ตัวแปรที่ใช้ในการจำแนกหรือตัวแปรที่ใช้ในการทำนาย) เป็นตัวแปรต่อเนื่องที่มีระดับการวัดอยู่ในมาตราอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) แต่จะมีส่วนที่แตกต่างกันคือ ในเทคนิควิเคราะห์ถดถอยโลจิสติคตัวแปรอิสระ ถ้าไม่ได้อยู่ในมาตราอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) ไม่ควรแปลงเป็นตัวแปรหุ่น (Dummy Coding) เพราะจะทำให้เกิดปัญหาในการวิเคราะห์ แต่ในเทคนิคการวิเคราะห์จำแนกกลุ่ม สามารถแปลงตัวแปรอิสระที่ไม่ได้อยู่ในมาตราการวัดอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) ให้เป็นตัวแปรหุ่น (Dummy Coding) ได้ และการวิเคราะห์จำแนกกลุ่ม ตัวแปรอิสระต่องเปนอิสระและถูกสุมมาจากประชากรที่มีการแจกแจงแบบปกติหลายตัวแปร รวมทั้งกลุ่มแต่ละกลุ่มต้องมีความแปรปรวนและความแปรปรวนร่วมเท่ากัน (Equal variance/covariance matrices) แต่ในการวิเคราะห์ถดถอยโลจิสติคไม่มีข้อกำหนดนี้

1. วิธีตรวจสอบข้อตกลงเบื้องต้น

1.1 ตัวแปรอิสระมีการแจกแจงปกติหลายตัวแปร (Normality of Independent Variables) ตรวจสอบได้จากการแจกแจงปกติแบบตัวแปรเดียว (Univariate Normal Distribution) โดยแยกทาการตรวจสอบตัวแปรอิสระทีละตัว หากพบว่าตัวแปรอิสระทุกตัวมีการแจกแจงปกติ ก็มีความน่าจะเป็นสูงที่จะมีการแจกแจงแบบปกติหลายตัวแปรด้วย หรืออาจจะตรวจสอบการแจกแจงแบบปกติหลายตัวแปรด้วยข้อมูลสุดโต่งแบบหลายตัวแปร (Multivariate Outliers) จากการวิเคราะห์ Mahalanobis Distances และถ้าพบว่าไม่มีข้อมูลสุดโต่งแบบหลายตัวแปร ก็มีความน่าจะเป็นสูงที่จะมีการแจกแจงแบบปกติหลายตัวแปรด้วย (ทรงศักดิ์ ภูสีอ่อน. 2554: 297)

1.2 เมตริกซ์ความแปรปรวนความแปรปรวนร่วมของตัวแปรอิสระของกลุ่มตัวอย่างต้องเท่ากัน (Equal Dispersion Matrices) ตรวจสอบได้ด้วยสถิติ Box's M (ทรงศักดิ์ ภูสีอ่อน. 2554: 297)

1.3 มีความสัมพันธ์เชิงเส้นตรง (Linearity of Relationships) ตรวจสอบได้จากสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรของเพียร์สัน (Pearson Product Moment Correlation: r_{xy})

1.4 ตัวแปรอิสระไม่มีความสัมพันธ์ร่วมเชิงพหุเชิงเส้น (Multicollinearity) ในการตรวจสอบความสัมพันธ์ร่วมเชิงพหุเชิงเส้น ตรวจสอบได้จากวิธีใช้สถิติสหสัมพันธ์เพียร์สัน (Pearson's Product

Moment Correlation) หรือใช้วิธีตรวจสอบโดยใช้สถิติ Collinearity โดยดูจากค่า Tolerance และ Variance Inflation Factor (VIF) หากค่า Tolerance เข้าใกล้ 0 มากเท่าใด ก็แสดงว่าระดับความสัมพันธ์ร่วมเชิงพหุเชิงเส้นของตัวแปรมีปัญหามาก ส่วนค่า VIF หากมีค่า เข้าใกล้ 10 มากเท่าใดก็แสดงว่าระดับความสัมพันธ์ร่วมเชิงพหุเชิงเส้นของตัวแปรมีปัญหามาก (ทรงศักดิ์ ภูสีอ่อน. 2554 : 297)

2. ลักษณะข้อมูลและการเตรียมข้อมูลที่จะนำมาวิเคราะห์

2.1 แบ่งกลุ่มประชากรหรือกลุ่มตัวอย่างออกเป็นกลุ่มอย่างน้อย 2 กลุ่ม (ต้องทราบมาก่อนการวิเคราะห์ว่าจะแบ่งเป็นกี่กลุ่ม)

2.2 เลือกตัวแปรอิสระที่คาดว่าจะทำให้กลุ่มที่แบ่งไว้ในข้อ 1 แตกต่างกัน

2.3 ตัวแปรตาม (ตัวแปรที่ถูกจำแนกหรือถูกทำนาย) เป็นตัวแปรที่แบ่งเป็นกลุ่มๆ ตั้งแต่ 2 กลุ่มขึ้นไป มีระดับการวัดในมาตรานามบัญญัติ (Nominal Scale) หรือมาตราเรียงอันดับ (Ordinal Scale) และถ้าหากมีข้อมูลอยู่ในระดับอื่นให้แปลงข้อมูลเป็นตัวแปรเชิงกลุ่มก่อนที่จะนำไปวิเคราะห์

2.4 ตัวแปรอิสระ (ตัวแปรที่ใช้ในการจำแนกหรือตัวแปรที่ใช้ในการทำนาย) เป็นตัวแปรต่อเนื่องที่มีระดับการวัดอยู่ในมาตราอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) และถ้าหากมีข้อมูลอยู่ในระดับอื่นให้แปลงให้เป็นตัวแปรหุ่น (Dummy Coding) ก่อนที่จะนำไปวิเคราะห์

3. วิธีการสร้างสมการจำแนกกลุ่ม

การสร้างสมการจำแนกกลุ่ม มี 2 วิธี คือ

3.1 วิธีตรง (Direct Method) เป็นวิธีการที่ผู้วิเคราะห์ต้องการนำตัวแปรทุกตัวตามที่ระบุไว้เข้าไปในสมการด้วยเหตุผลทางทฤษฎีว่าจะแบ่งแยกได้กี่สมการ มีลักษณะอย่างไร และเพื่อพิสูจน์ว่าตัวแปรที่คิดว่ามีความสำคัญต่อการจำแนกที่ระบุไว้ตามทฤษฎีนั้นแท้จริงแล้วมีความสำคัญตามทฤษฎีกล่าวไว้หรือไม่

3.2 วิธีแบบขั้นตอน (Stepwise Method) เป็นวิธีการที่เลือกตัวแปรทีละตัวเข้ามาในสมการโดยนำตัวแปรที่ดีที่สุดในการจำแนกเข้ามาในสมการเป็นตัวแรก จากนั้นก็จะหาตัวแปรที่ดีที่สุดตัวที่สองเข้ามาในสมการเพื่อปรับปรุงแก้ไขทำให้สมการจำแนกกลุ่มดีขึ้น และในขั้นตอนต่อๆไปก็จะเป็นการนำตัวแปรที่ดีที่สุดแต่ละตัวที่เหลือเข้ามาในสมการต่อไปเพื่อจะได้สมการจำแนกกลุ่มที่ดีที่สุด

4. สถิติสำคัญของการวิเคราะห์จำแนกกลุ่ม

ทรงศักดิ์ ภูสีอ่อน (2551: 298) ได้กล่าวถึงสถิติที่สำคัญสำหรับการวิเคราะห์จำแนกกลุ่ม ไว้ดังนี้คือ

4.1 ค่าไอเก้น (Eigenvalue) เป็นค่าที่แสดงอัตราส่วนการผันแปรระหว่างกลุ่ม ต่อการผันแปรภายในกลุ่ม ถ้าค่าไอเก้นมีค่าสูง ก็แสดงว่าสมการดีหรือมีค่าจำแนกสูงหรือกล่าวอีกนัยหนึ่งได้ว่า Eigenvalue ก็คือ Variance ของคะแนนแปลงรูป Y ที่แปลงมาจาก X_1, X_2 ,... X_p นั่นเอง

4.2 ค่าสหสัมพันธ์คาโนนิคอล (Canonical Correlation) เป็นค่าที่แสดงความสัมพันธ์ของสมการจำแนกกับกลุ่ม สามารถชี้ให้เห็นว่า การเป็นสมาชิกกลุ่มดังกล่าวมีความสัมพันธ์กับสมการที่หามาได้มากน้อยเพียงใด ดังนั้นถ้าค่าความสัมพันธ์คานิคอลมีค่าสูง (เข้าใกล้ 1.00) แสดงให้เห็นว่าการเป็นสมาชิกของกลุ่มสามารถอธิบายความผันแปรของตัวแปรกับสมการจำแนกได้มาก

4.3 ค่าวิลค์แลมบ์ดา (Wilks' Lambda) เป็นสถิติที่ใช้ทดสอบความแตกต่างระหว่างกลุ่มและเป็นมาตรวัดอำนาจในการจำแนกกลุ่มของตัวแปรด้วย โดยสิ่งสำคัญและควรคำนึงในการพิจารณาค่าวิลค์แลมป์ดาก็คือ ถึงแม้ว่าค่าดังกล่าวจะมีนัยสำคัญทางสถิติ แต่ค่าดังกล่าวก็ไม่อาจบอกอัตราความมีประสิทธิภาพในการจำแนกของสมการได้ เพียงเป็นสถิติทดสอบว่า ค่าเฉลี่ยของประชากรของกลุ่มต่างๆแตกต่างกันหรือไม่ ซึ่งถ้าค่าวิลค์แลมบ์ดามีค่ามาก ตัวแปรจะอธิบายการเป็นสมาชิกของกลุ่มได้น้อย ถ้าค่าวิลค์แลมบ์ดา มีค่าน้อย ตัวแปรจะอธิบายการเป็นสมาชิกของกลุ่มได้มาก

5. การวิเคราะห์จำแนกกลุ่มด้วยโปรแกรมสำเร็จรูปทางสถิติ

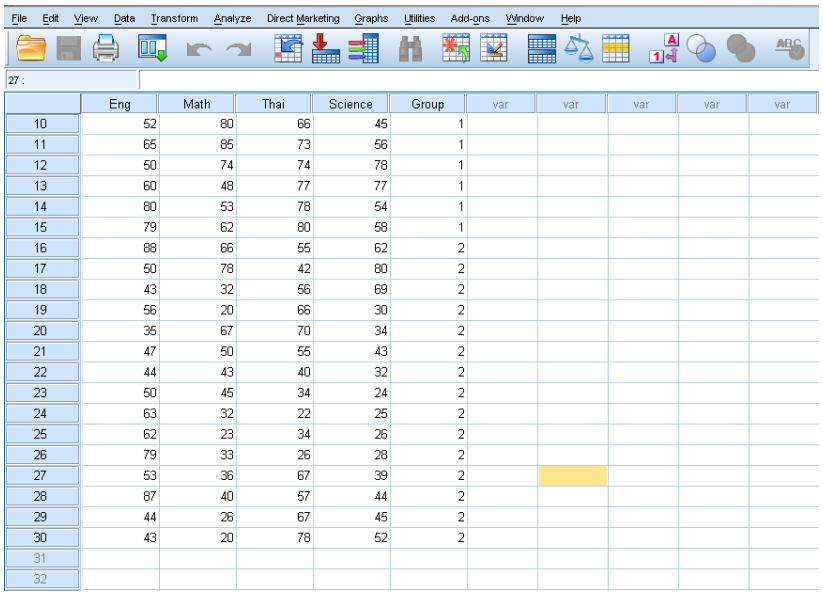
คุณครูผู้สอนมัธยมศึกษาท่านหนึ่งต้องการแบ่งกลุ่มนักเรียนที่เข้าใหม่จึงรวบรวมผลการเรียนวิชาภาษาอังกฤษ คณิตศาสตร์ ภาษาไทย และวิทยาศาสตร์ จากเด็กนักเรียนเก่า 2 กลุ่มคือ กลุ่มเก่ง กับกลุ่มอ่อน นำมากลุ่มละ 15 คน ได้ข้อมูลดังนี้

คะแนนวิชาภาษาอังกฤษ	คะแนนวิชาคณิตศาสตร์	คะแนนวิชาภาษาไทย	คะแนนวิชาวิทยาศาสตร์	กลุ่มเด็กนักเรียน (1 = อ่อน , 2 = เก่ง)
44	78	50	45	1
55	82	45	67	1
68	88	67	80	1
72	77	55	88	1
89	76	66	79	1
60	58	76	80	1
53	83	47	45	1
76	62	58	34	1
43	55	67	26	1
52	80	66	45	1
65	85	73	56	1
50	74	74	78	1
60	48	77	77	1
80	53	78	54	1
79	62	80	58	1
88	66	55	62	2
50	78	42	80	2

คะแนนวิชาภาษาอังกฤษ	คะแนนวิชาคณิตศาสตร์	คะแนนวิชาภาษาไทย	คะแนนวิชาวิทยาศาสตร์	กลุ่มเด็กนักเรียน (1 = อ่อน , 2 = เก่ง)
43	32	56	69	2
56	20	66	30	2
35	67	70	34	2
47	50	55	43	2
44	43	40	32	2
50	45	34	24	2
63	32	22	25	2
62	23	34	26	2
79	33	26	28	2
53	36	67	39	2
87	40	57	44	2
44	26	67	45	2
43	20	78	52	2

การหาสมการจำแนกที่เหมาะสมเพื่อนำไปใช้ในการจัดกลุ่มนักเรียนที่เข้าใหม่ สามารถวิเคราะห์ตามขั้นตอนได้ดังนี้

1. สร้างข้อมูลดังตารางลงในโปรแกรมจากนั้นให้บันทึกข้อมูลโดยใช้ชื่อ file ว่า Discriminant Analysis ดังภาพ

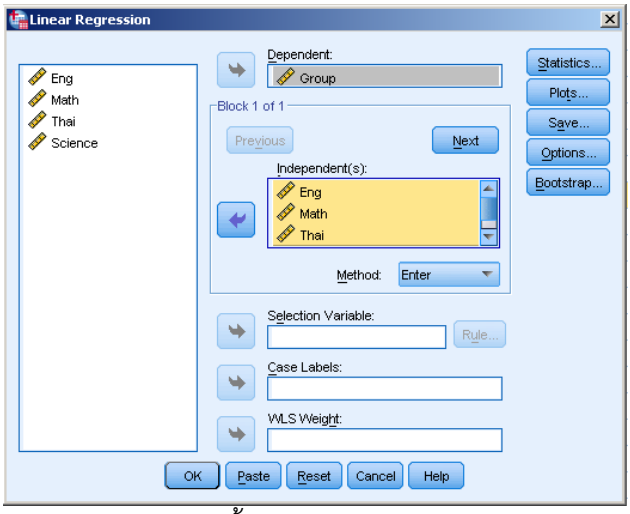


	Eng	Math	Thai	Science	Group	var	var	var	var	var
10	52	80	66	45	1					
11	65	85	73	56	1					
12	50	74	74	78	1					
13	60	48	77	77	1					
14	80	53	78	54	1					
15	79	62	80	58	1					
16	88	66	55	62	2					
17	50	78	42	80	2					
18	43	32	56	69	2					
19	56	20	66	30	2					
20	35	67	70	34	2					
21	47	50	55	43	2					
22	44	43	40	32	2					
23	50	45	34	24	2					
24	63	32	22	25	2					
25	62	23	34	26	2					
26	79	33	26	28	2					
27	53	36	67	39	2					
28	87	40	57	44	2					
29	44	26	67	45	2					
30	43	20	78	52	2					
31										
32										

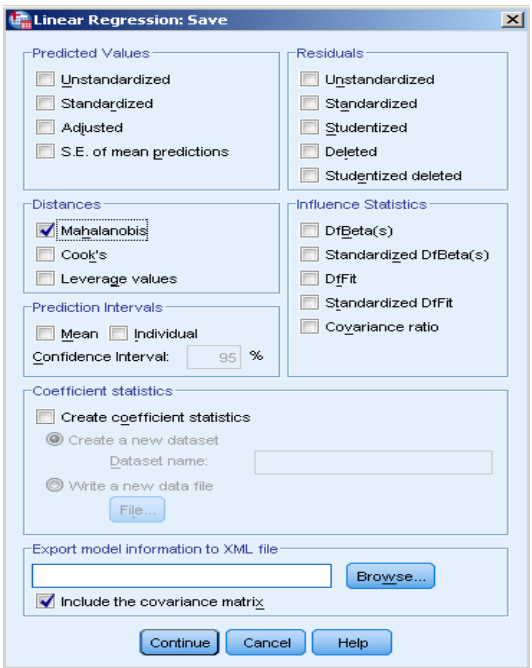
2. ตรวจสอบข้อตกลงเบื้องต้น ดังนี้

2.1 ตรวจสอบว่าตัวแปรอิสระมีการแจกแจงปกติหลายตัวแปร (Normality of Independent Variables) โดยตรวจสอบจากข้อมูลสุดโต่งแบบหลายตัวแปร (multivariate Outliers) ด้วยการวิเคราะห์ Mahalanobis Distances ดังนี้

2.1.1 Click Analyze ⟶Regression ⟶Linear จะมีหน้าต่าง Linear Regression เปิดขึ้นมานำตัวแปรอิสระทั้งหมดเข้าไปในช่อง Independent(s) และเอาตัวแปรตามเข้าไปไว้ที่ช่อง Dependent ดังภาพ



2.2.2 จากนั้น Click ⟶Save จะปรากฏหน้าต่าง Linear Regression: Save ขึ้นมา ในช่อง Distances ให้เลือก Mahalanobis ดังภาพ



2.2.3 Click Continue ⟶ OK จะได้ผลลัพธ์ดังนี้

ตารางที่ 5.1 ตารางแสดงค่า Mahalanobis Distances

Residuals Statistics[a]

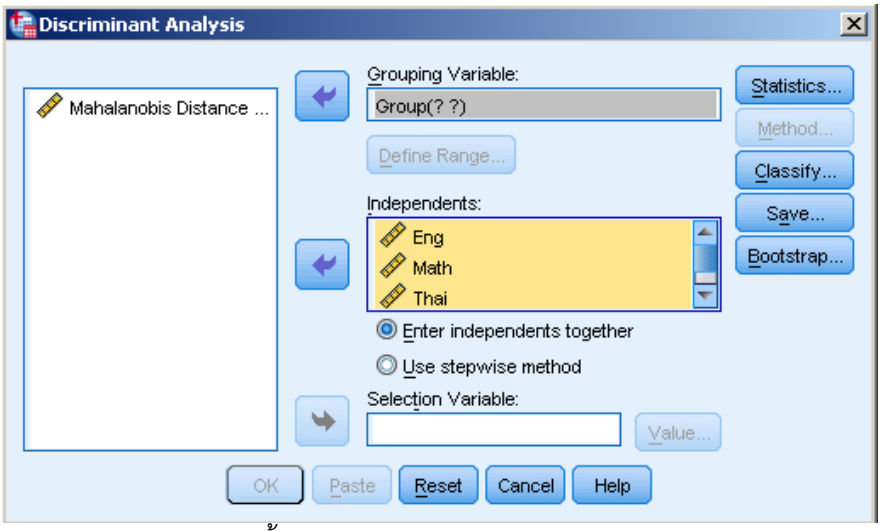
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.87	2.23	1.50	.402	30
Std. Predicted Value	-1.577	1.810	.000	1.000	30
Standard Error of Predicted Value	.082	.171	.136	.020	30
Adjusted Predicted Value	.84	2.28	1.50	.413	30
Residual	-.458	.742	.000	.312	30
Std. Residual	-1.362	2.208	.000	.928	30
Stud. Residual	-1.503	2.403	.003	1.025	30
Deleted Residual	-.557	.917	.003	.381	30
Stud. Deleted Residual	-1.544	2.685	.025	1.084	30
Mahal. Distance	.745	6.506	3.867	1.389	30
Cook's Distance	.000	.334	.046	.082	30
Centered Leverage Value	.026	.224	.133	.048	30

a. Dependent Variable: Group

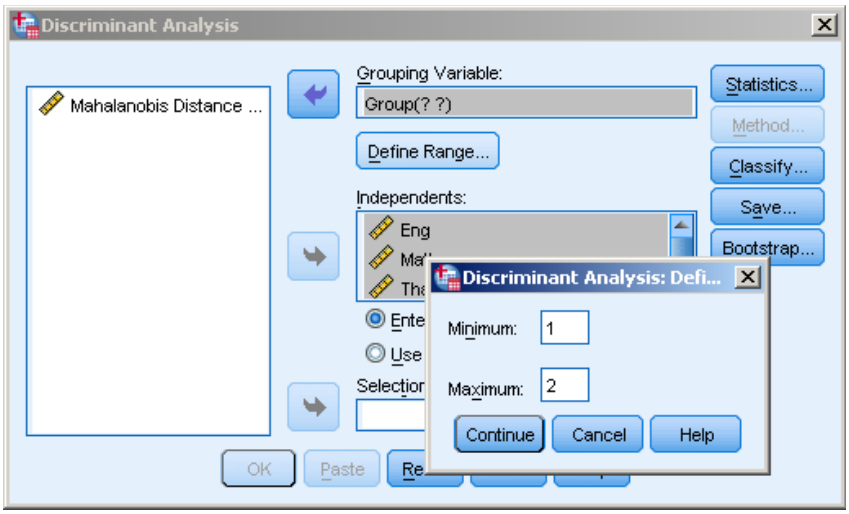
จากตารางพบว่า ค่า Mahal. Distance ที่วิเคราะห์ได้มีค่าสูงสุด 6.506 เมื่อนำไปเทียบกับค่าไคร์สแควร์ (Chi-square) ที่ระดับนัยสำคัญ 0.05 ที่ค่าองศาความเป็นอิสระเท่ากับจำนวนตัวแปรอิสระ พบว่ามีค่าน้อยกว่าค่าวิกฤติ (9.49) แสดงว่าในภาพรวมแล้วไม่มีข้อมูลสุดโต่งแบบหลายตัวแปร จึงสรุปได้ว่าตัวแปรอิสระน่าจะมีการแจกแจงแบบปกติหลายตัวแปร (เป็นไปตามข้อตกลงเบื้องต้น)

2.2 ตรวจสอบความเท่าเทียมกันของเมตริกซ์ความแปรปรวนความแปรปรวนร่วมของแต่ละกลุ่ม (Equal Dispersion Matrices) ตรวจสอบได้ด้วยสถิติ Box's M ดังนี้

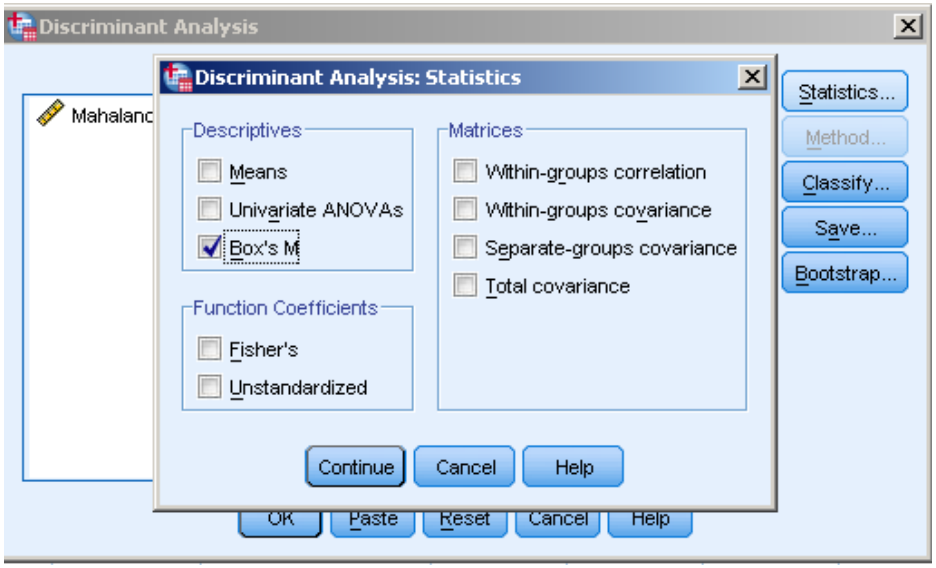
2.2.1 Click Analyze ⟶ Classify ⟶ Discriminant จะมีหน้าต่าง Discriminant Analyze เปิดขึ้นมา นำตัวแปรอิสระทั้งหมดเข้าไปในช่อง Independent(s) และเอาตัวแปรตามเข้าไปในช่อง Grouping Variable ดังภาพ



2.2.2 จากนั้น Click ⟶ Define Range จะปรากฏหน้าต่าง Discriminant Analyze: Define Range ขึ้นมาให้ระบุค่าต่ำสุดและสูงสุดของตัวแปรตาม (group) ในตัวอย่าง Minimum คือ 1 และ Maximum คือ 2 ดังภาพ



2.2.3 Click ⟶ Continue จะปรากฏหน้าต่าง Discriminant Analyze ขึ้นมาอีกครั้ง ให้ Click ⟶ Statistics จะปรากฏหน้าต่าง Discriminant Analyze: Statistics ขึ้นมา ในช่อง Descriptiveให้เลือก Box's M ดังภาพ



2.2.4 Click Continue แล้ว OK จะปรากฏผลการวิเคราะห์สถิติ Box's M ดังตาราง

ตารางที่ 5.2 แสดงผลการวิเคราะห์สถิติ Box's M

Test Results

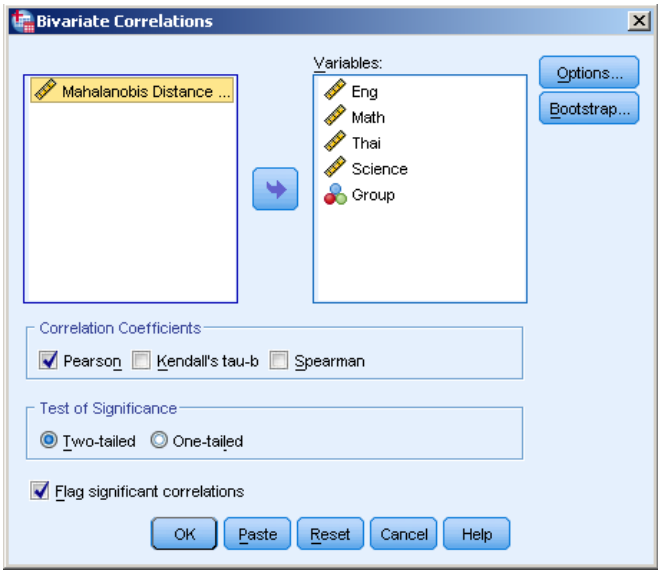
Box's M		11.534
F	Approx.	.973
	df1	10
	df2	3748.207
	Sig.	.464

Tests null hypothesis of equal population covariance matrices.

จากตาราง 5.2 เป็นผลการทดสอบความเท่าเทียมกันของเมตริกซ์ความแปรปรวนความแปรปรวนร่วมของแต่ละกลุ่ม โดยดูจากค่า sig = 0.464 ซึ่งมีค่ามากกว่าค่า α ที่ตั้งไว้ (.05) จึงสรุปว่าเมตริกซ์ความแปรปรวนความแปรปรวนร่วมของแต่ละกลุ่มมีความเท่าเทียมกัน (เป็นไปตามข้อตกลงเบื้องต้น)

2.3 มีความสัมพันธ์เชิงเส้นตรง (Linearity of Relationships) ตรวจสอบได้จากสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรของเพียร์สัน ดังนี้

2.3.1 Click Analyze ⟶ Correlate ⟶ Bivariate จะมีหน้าต่าง Bivariate Correlations เปิดขึ้นมา นำเอาตัวแปรอิสระและตัวแปรตามทั้งหมดเข้าไปในช่อง Variables ดังภาพ



2.3.2 Click Ok จะปรากฏผลการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรของเพียร์สันดังตาราง

ตารางที่ 5.3 ตารางแสดงผลการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรของเพียร์สัน

Correlations

		Eng	Math	Thai	Science	Group
Eng	Pearson Correlation	1	.135	.009	.222	-.224
	Sig. (2-tailed)		.475	.963	.238	.234
	N	30	30	30	30	30
Math	Pearson Correlation	.135	1	.168	.523**	-.705**
	Sig. (2-tailed)	.475		.375	.003	.000
	N	30	30	30	30	30
Thai	Pearson Correlation	.009	.168	1	.422*	-.441*
	Sig. (2-tailed)	.963	.375		.020	.015
	N	30	30	30	30	30
Science	Pearson Correlation	.222	.523**	.422*	1	-.468**
	Sig. (2-tailed)	.238	.003	.020		.009
	N	30	30	30	30	30
Group	Pearson Correlation	-.224	-.705**	-.441*	-.468**	1
	Sig. (2-tailed)	.234	.000	.015	.009	
	N	30	30	30	30	30

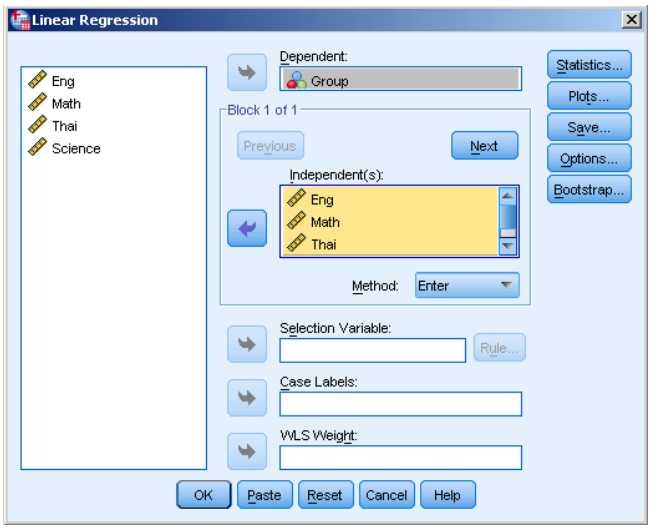
**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

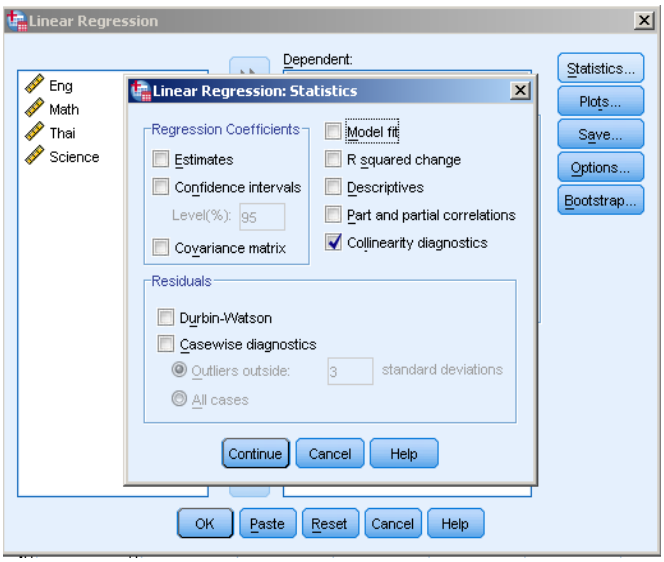
จากตารางที่ 5.3 แสดงผลการวิเคราะห์ความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรตามและตัวแปรอิสระพบว่ามีความสัมพันธ์กันเชิงเส้นตรงโดยดูจากค่าในสัมประสิทธิ์สหสัมพันธ์ของตัวแปรอิสระกับตัวแปรตามไม่มีค่าเป็น 0 (Group กับ Eng , Maths , Thai , Science)

2.4 ตรวจสอบว่าตัวแปรอิสระไม่มีความสัมพันธ์ร่วมเชิงพหุเชิงเส้น (Multicollinearity) โดยใช้สถิติ Collinearity ดังนี้

2.4.1 Click Analyze ⟶ Regression ⟶ Linear จะมีหน้าต่าง Linear Regression เปิดขึ้นมา นำตัวแปรอิสระทั้งหมดเข้าไปในช่อง Independent(s) และนำตัวแปรตามเข้าไปในช่อง Dependent ดังภาพ



2.4.2 Click ⟶ Statistics จะปรากฏหน้าต่าง Linear Regression: Statistics ขึ้นมา ในช่อง Regression Coefficient ให้เลือก Colinearity diagnostics ดังภาพ



2.4.3 Click Continue แล้ว OK จะได้ผลลัพธ์ดังตาราง

ตารางที่ 5.4 แสดงผลการวิเคราะห์ค่า Tolerance และ VIF

Coefficients[a]

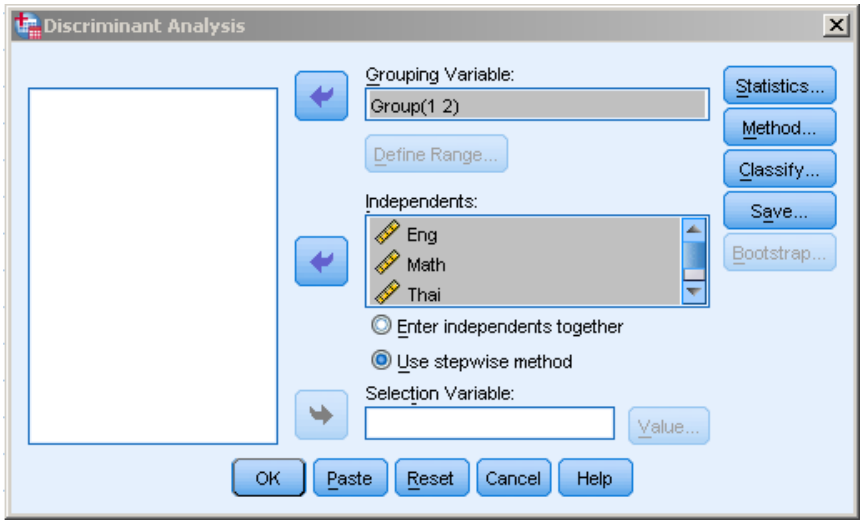
Model		Collinearity Statistics	
		Tolerance	VIF
1	Eng	.942	1.062
	Math	.722	1.384
	Thai	.811	1.233
	Science	.586	1.705

a. Dependent Variable: Group

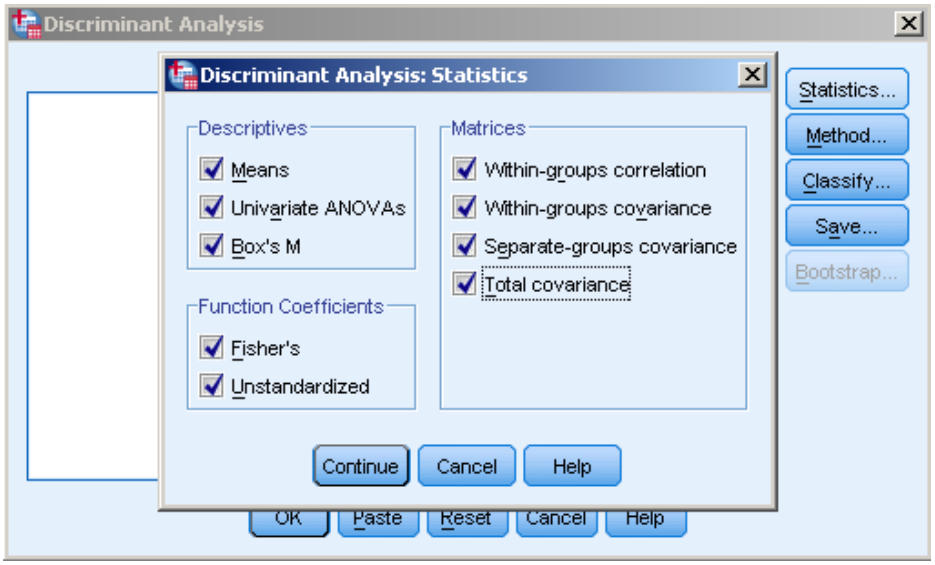
จากตารางที่ 5.4 แสดงผลการวิเคราะห์ค่า Tolerance และ VIF พบว่าค่า Tolerance มีค่าเข้าใกล้ 1 และค่า VIF ไม่ได้เข้าใกล้ 10 จึงได้สรุปว่าตัวแปรอิสระไม่ได้มีความสัมพันธ์ร่วมเชิงพหุเชิงเส้น (Multicollinearity)

3. วิเคราะห์ข้อมูลเพื่อหาสมการจำแนกโดยใช้ Discriminant Analysis

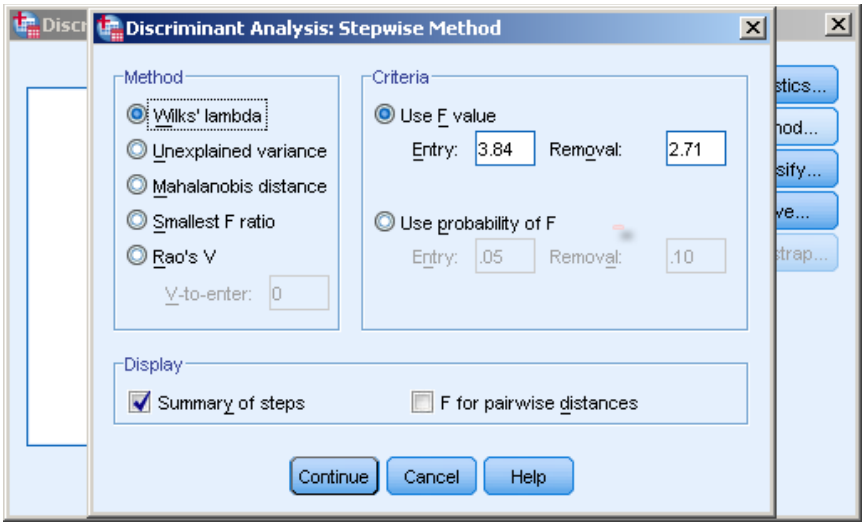
3.1 เลือก Analyze ⟶ Classify ⟶ Discriminant จากนั้นจะปรากฏหน้าต่าง Discriminant Analysis ขึ้นมา ให้เลือกตัวแปร Group เข้าช่อง Grouping Variable ทำการ Define Range ในส่วนของ Minimum ใส่เลข 1 ในส่วนของ Maximum ใส่เลข 2 จากนั้นให้เลือกตัวแปร Eng , Maths , Thai , Science เข้าช่อง Independents แล้วคลิกเลือก Use Stepwise Method (เลือกสมการทำนายที่ดีที่สุด) ดังภาพ



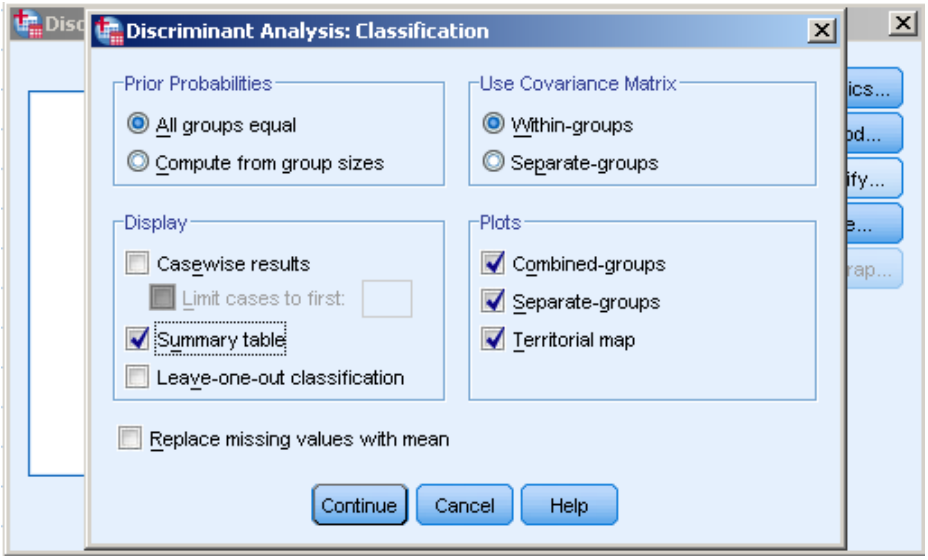
3.2 Click Statistics จะปรากฏหน้าต่าง Discriminant Analysis Statistics ขึ้นมาและมีกลุ่มสถิติให้เลือก 3 กลุ่ม ได้แก่ Descriptive (สถิติพรรณนา), Function Coefficients (สัมประสิทธิ์ของสมการ) และ Matrices (เมตริกซ์) ให้ทำเครื่องหมาย เลือกสถิติที่ต้องการแต่ละกลุ่มเป็นทางเลือกที่จะเลือกอย่างน้อย 1 อย่าง ในที่นี้ให้เลือกทั้งหมดดังภาพ



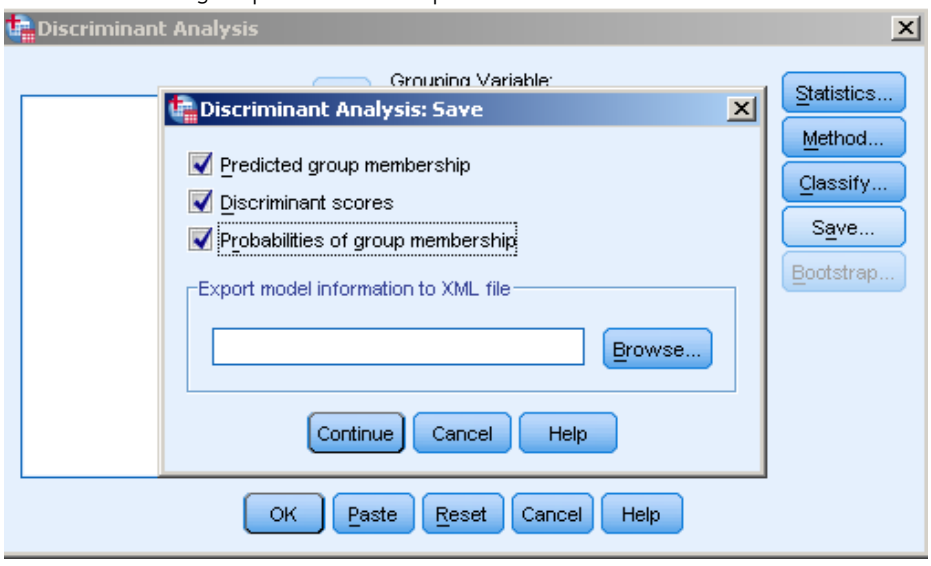
3.2 คลิก Continue จะกลับมาที่หน้าต่าง Discriminant Analysis ให้เลือก Method ในส่วนของเมนูย่อย Method มีเกณฑ์ในการคัดเลือกตัวแปรอยู่หลายวิธี แต่ในที่นี้ให้เลือกใช้ Wilks' lambda ซึ่งเป็นวิธีที่คนนิยมใช้ และในโปรแกรม SPSS มักเลือกใช้ ดังภาพ



3.3 Click Continue จะกลับมาที่หน้าต่าง Discriminant Analysis ให้เลือก Classify ในส่วนของ Prior Probabilities เลือก All groups equal ในส่วนของ Display เลือก Summary table ในส่วนของ Use Covariance Matrix เลือก Within – groulps ในส่วนของ Plots เลือก Combined – groups , Separate – groups และ Territorial map ดังภาพ



3.4 Click Continue จะกลับมาที่หน้าต่าง Discriminant Analysis ให้เลือก Save ในส่วนนี้ให้เลือก Predicted group membership , Discriminant scores และ Probabilities of group membership



3.5 Click Continue แล้ว OK จะได้ผลลัพธ์ดังนี้

ตารางที่ 5.5 ตารางแสดงผลการทดสอบความเท่ากันของค่าเฉลี่ยของตัวแปรแต่ละกลุ่มในการทดสอบ

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Eng	.950	1.480	1	28	.234
Math	.503	27.624	1	28	.000
Thai	.805	6.774	1	28	.015
Science	.781	7.872	1	28	.009

จากตาราง 5.5 เป็นการนำเสนอผลการทดสอบความเท่ากันของค่าเฉลี่ยของตัวแปรแต่ละกลุ่มในการทดสอบนั้น โดยการวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA) และให้ค่าวิลค์แลมบ์ดา ซึ่งทั้งสองการทดสอบต้องอาศัยสถิติ F ทดสอบเช่นเดียวกัน โดยถ้าค่า sig. มีค่าต่ำกว่าระดับนัยสำคัญที่เราตั้งไว้ (α = 0.05) ก็แสดงว่ากลุ่ม 2 กลุ่มนั้น มีค่าเฉลี่ยในตัวแปรนั้นๆ แตกต่างกัน และจากตาราง Tests of Equality of Group Means พบว่า มีค่าเฉลี่ยของตัวแปร Math Thai และ Science ใน 2 กลุ่ม แตกต่างกัน (Sig.= .000 , Sig.= .015, Sig.= .009) และทั้งสองกลุ่มมีค่าเฉลี่ยไม่แตกต่างกันในตัวแปร Eng (ค่า sig. = .234)

ตารางที่ 5.6 ตารางแสดงค่าความแปรปรวนร่วม (Covariance) ของทั้ง 2 กลุ่มรวมกันและค่าสัมประสิทธิ์สหสัมพันธ์

Pooled Within-Groups Matrices[a]

		Eng	Math	Thai	Science
Covariance	Eng	234.281	-7.774	-23.226	37.836
	Math	-7.774	244.352	-51.710	87.500
	Thai	-23.226	-51.710	216.995	72.607
	Science	37.836	87.500	72.607	329.600
Correlation	Eng	1.000	-.032	-.103	.136
	Math	-.032	1.000	-.225	.308
	Thai	-.103	-.225	1.000	.271
	Science	.136	.308	.271	1.000

a. The covariance matrix has 28 degrees of freedom.

จากตาราง 5.6 พบว่า ตัวแปรทั้ง 4 มีความสัมพันธ์กันทั้งทางบวกและทางลบ ดังนี้ คือ Eng กับ science มีความสัมพันธ์กันในทางบวก (.136) ส่วน Eng กับ Maths มีความสัมพันธ์กันในทางลบ (-.032) และ Eng กับ Thai มีความสัมพันธ์กันในทางลบ (-.103)

ตารางที่ 5.7 ตารางแสดงตัวแปรที่อยู่ในสมการที่ดีที่สุดที่จะนำไปสร้างสมการวิเคราะห์

Variables Entered/Removed[a,b,c,d]

Step	Entered	Wilks' Lambda							
		Statistic	df1	df2	df3	Exact F			
						Statistic	df1	df2	Sig.
1	Math	.503	1	1	28.000	27.624	1	28.000	.000
2	Thai	.396	2	1	28.000	20.585	2	27.000	.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 8.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

จากตารางที่ 5.7 พบว่า ตัวแปรที่อยู่ในสมการที่ดีที่สุดที่จะนำไปสร้างสมการวิเคราะห์ต่อไป คือ ตัวแปร Math กับ Thai

ตารางที่ 5.8 ตารางแสดงค่า Eigenvalue และ Canonical Correlation

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.525[a]	100.0	100.0	.777

a. First 1 canonical discriminant functions were used in the analysis.

จากตารางที่ 5.8 พบว่าเนื่องจากข้อมูลที่กำหนดให้มี 2 กลุ่ม จึงมีเพียงสมการเดียว และมีค่า Eigenvalue เพียงค่าเดียว ซึ่งมีค่าเท่ากับ 1.525 และมีค่าความสัมพันธ์คาโนนิคอล (Canonical Correlation) เท่ากับ .777 เมื่อนำค่า Canonical Correlation มายกกำลังสอง จะเป็นค่าที่แสดงให้เห็นว่าตัวแปรในสมการจำแนกกลุ่ม สามารถอธิบายความแปรปรวนของตัวแปรตามได้ร้อยละเท่าไร ซึ่งกรณีนี้อธิบาย 0.777^2 เท่ากับ 61.0%

ตารางที่ 5.9 ตารางแสดงค่าสถิติที่ใช้ทดสอบสมการจำแนกกลุ่ม

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.396	25.007	2	.000

จากตารางที่ 5.9 เป็นการนำเสนอสถิติที่ใช้ทดสอบสมการจำแนกกลุ่มที่วิเคราะห์ได้ว่ามีนัยสำคัญทางสถิติหรือไม่ โดยพิจารณาที่ค่าความมีนัยสำคัญ (Sig.) ถ้าค่า Sig. น้อยกว่าระดับนัยสำคัญที่กำหนดไว้ แสดงว่าสมการจำแนกกลุ่มสามารถจำแนกได้อย่างมีนัยสำคัญทางสถิติ ซึ่งจากการวิเคราะห์พบว่าสมการมีนัยสำคัญ (.000 < .05) นั่นแสดงว่า สมการที่ได้จากการวิเคราะห์จำแนกกลุ่มสามารถจำแนกออกเป็น 2 กลุ่มได้

ตารางที่ 5.10 ตารางแสดงค่าสัมประสิทธิ์ (ค่าน้ำหนัก) ของตัวแปรจำแนกในสมการจำแนกกลุ่ม

Standardized Canonical Discriminant Function Coefficients

	Function
	1
Math	.941
Thai	.610

จากตารางที่ 5.10 เป็นการนำเสนอค่าน้ำหนักของตัวแปรแต่ละตัว ซึ่งจะเห็นว่าตัวแปร Math มีค่าน้ำหนัก 0.941 ซึ่งมากกว่าค่าน้ำหนัก Thai ซึ่งมีค่าน้ำหนัก 0.610 แสดงว่าตัวแปร Math มีความสำคัญในการจำแนกกลุ่มในสมการจำแนกกลุ่ม ซึ่งจากผลวิเคราะห์สามารถเขียนเป็นสมการจำแนกกลุ่มได้ดังนี้

$$Z_y = 0.941\ (Z_{\mathbf{Math}}) + 0.610\ (Z_{Thai})$$

ตารางที่ 5.11 ตารางแสดงค่า Structure Matrix

Structure Matrix

	Function
	1
Math	.804
Science[a]	.456
Thai	.398
Eng[a]	-.093

จากตารางที่ 5.11 สามารถนำไปใช้ตีความหมายสมการจำแนกกลุ่มอีกวิธีหนึ่ง โดยจะช่วยประเมินว่าตัวแปรแต่ละตัวมีผลอย่างไรต่อการจำแนก เป็นการดูความสัมพันธ์ระหว่างค่าของตัวแปรตัวกับค่าคะแนนจำแนกที่คำนวณจากสมการจำแนก จากตารางแสดงให้เห็นว่าตัวแปร Eng มีความสำคัญต่อสมการจำแนกน้อยมาก

ตารางที่ 5.12 ตารางแสดงค่าสัมประสิทธิ์ของตัวแปรจำแนกในสมการจำแนกกลุ่ม

Canonical Discriminant Function Coefficients

	Function
	1
Math	.060
Thai	.041
(Constant)	-5.768

Unstandardized coefficients

จากตารางที่ 5.12 แสดงค่าสัมประสิทธิ์ของตัวแปรจำแนกในสมการจำแนกกลุ่ม ซึ่งเป็นสมการในรูปคะแนนดิบ ค่าน้ำหนักที่ได้จึงไม่อยู่ในรูปมาตรฐาน (Unstandardized Coefficients) เป็นผลจากการเลือก Unstandardized ใน Statistics Option ผลที่ได้มีทั้งค่าน้ำหนักในแต่ละตัวแปร และค่าคงที่ (Constant) จากผลการวิเคราะห์สามารถนำมาเขียนเป็นสมการในรูปคะแนนดิบได้ดังนี้

$$Y' = -5.768 + .060\ (Math) + .041\ (Thai)$$

ตารางที่ 5.13 ตารางแสดงค่ากลางของกลุ่ม

Functions at Group Centroids

Group	Function
	1
กลุ่มอ่อน	1.193
กลุ่มเก่ง	-1.193

Unstandardized canonical discriminant functions evaluated at group means

จากตารางที่ 5.13 แสดงค่าที่สามารถใช้ประเมินสมการจำแนกคาโนนิคอลด้วยค่าเฉลี่ยของกลุ่ม (Canonical Discriminant Functions Evaluated at Group Means) ผลการวิเคราะห์เป็นคะแนนดิบ (Unstandardized) ซึ่งค่ากลางหาได้จากการหาคะแนนจำแนกของแต่ละหน่วยวิเคราะห์ โดยใช้ค่าสัมประสิทธิ์ของตัวแปรคูณกับค่าของตัวแปรของแต่ละหน่วยวิเคราะห์ เมื่อได้คะแนนของแต่ละหน่วยจากสมการแล้วก็หาค่าเฉลี่ยหรือค่ากลางของแต่ละกลุ่มได้ โดยเอาผลรวมของค่าคะแนนจำแนกของแต่ละหน่วยในกลุ่มนั้นหารด้วยจานวนหน่วยในกลุ่มนั้น โดยจากตารางพบว่า กลุ่ม 1 มีค่ากลางของกลุ่ม (Group Centroids) เท่ากับ 1.193 ส่วนกลุ่ม 2 มีค่า -1.193 ซึ่งตรงข้ามกัน แสดงว่า สมการ

สมการดังกล่าวสามารถจำแนกได้ดี ในกรณีที่มีหน่วยวิเคราะห์ใหม่ก็สามารถคำนวณหาคะแนนจำแนกแล้วนำมาเปรียบเทียบกับค่ากลางของแต่ละกลุ่ม ถ้ามีแนวโน้มเข้าใกล้ค่ากลางกลุ่มใดก็มีโอกาสในการเป็นสมาชิกกลุ่มนั้นๆ

ดังกล่าวสามารถจำแนกได้ดี ในกรณีที่มีหน่วยวิเคราะห์ใหม่ก็สามารถคำนวณหาคะแนนจำแนกแล้วนำมาเปรียบเทียบกับค่ากลางของแต่ละกลุ่ม ถ้ามีแนวโน้มเข้าใกล้ค่ากลางกลุ่มใดก็มีโอกาสในการเป็นสมาชิกกลุ่มนั้นๆ

ตารางที่ 5.14 ตารางแสดงค่าสัมประสิทธิ์ (ค่าน้ำหนัก) และค่าคงที่ของสมการจำแนกโดยแยกเป็นกลุ่ม ตามวิธีของ Fisher's

Classification Function Coefficients

	Group	
	กลุ่มอ่อน	กลุ่มเก่ง
Math	.372	.228
Thai	.389	.291
(Constant)	-26.552	-12.791

Fisher's linear discriminant functions

จากตารางที่ 5.14 แสดงผลการวิเคราะห์ค่าสัมประสิทธิ์ (ค่าน้ำหนัก) และค่าคงที่ของสมการจำแนกโดยแยกเป็นกลุ่ม ตามวิธีของ Fisher's (Fisher's linear discrimination function) จำนวนสมการจะมีเท่ากับจำนวนกลุ่ม (ในกรณีนี้มี 2 สมการ) จากผลการวิเคราะห์จะได้สมการดังนี้

สมการของกลุ่ม 1 y'_1 = -26.552 + .372Math + .389Thai

สมการของกลุ่ม 2 y'_2 = -12.791 + .228 Math + .291Thai

จากผลการวิเคราะห์สามารถนำไปใช้ประโยชน์ได้ โดยการพยากรณ์หน่วยวิเคราะห์ว่าควรจัดให้อยู่กลุ่มใด โดยการแทนค่าตัวแปรอิสระ (Math กับ Thai) ของหน่วยวิเคราะห์ นั้น ๆ ลงในทั้ง 2 สมการ ถ้าสมการใดมีค่ามากกว่าก็จัดอยู่ในกลุ่มนั้น

ตารางที่ 5.15 ตารางแสดงการตรวจสอบและพิจารณาความน่าเชื่อถือของสมการจำแนกกลุ่ม

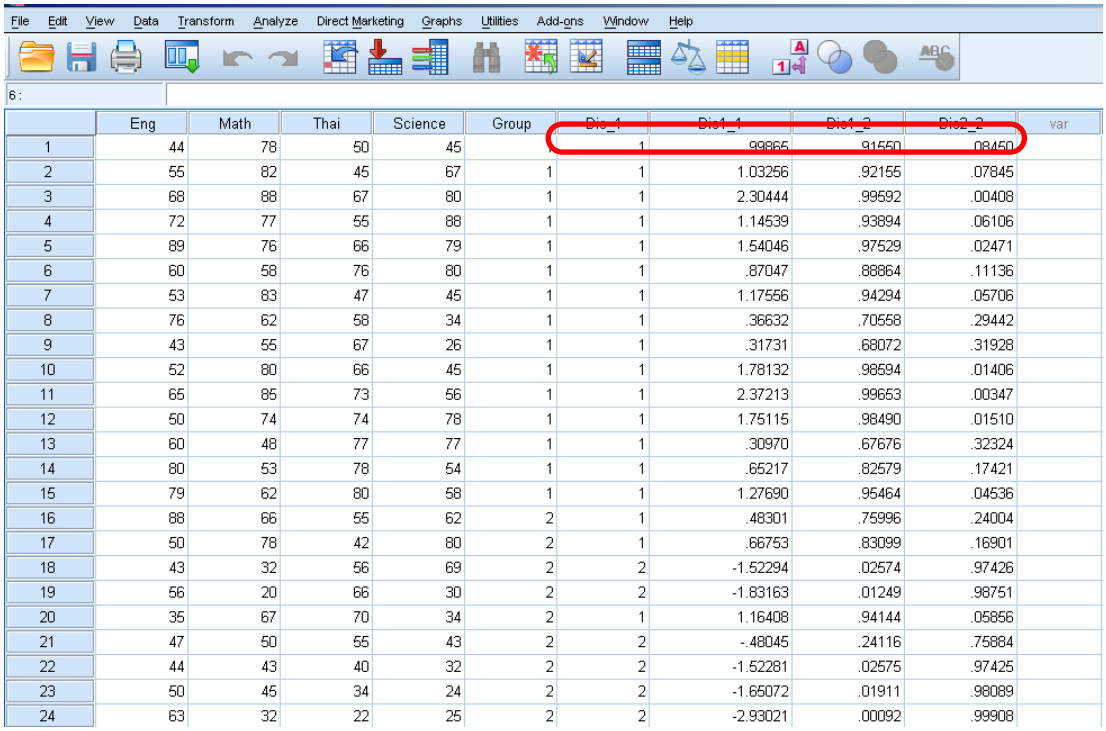
Classification Results[a]

Group			Predicted Group Membership		Total
			กลุ่มอ่อน	กลุ่มเก่ง	
Original	Count	กลุ่มอ่อน	15	0	15
		กลุ่มเก่ง	3	12	15
	%	กลุ่มอ่อน	100.0	.0	100.0
		กลุ่มเก่ง	20.0	80.0	100.0

a. 90.0% of original grouped cases correctly classified.

จากตาราง 5.15 แสดงผลการวิเคราะห์ประสิทธิภาพของสมการจำแนกว่าสามารถจำแนกกลุ่มได้ถูกต้องมากน้อยเพียงใด โดยเปรียบเทียบกลุ่มที่แบ่งไว้เดิม (Original) กับการแบ่งกลุ่มที่ได้จากการทำนายจากสมการ (Predicted Group Membership) พบว่า ในกลุ่มที่ 1 เดิมมี 15 cases จากการทำนายโดยใช้สมการจำแนกกลุ่มพบว่า ทำนายได้ถูกต้องทั้ง 15 cases คิดเป็นร้อยละ 100.0 ส่วนในกลุ่ม 2 เดิมมี 15 cases แต่ทำนายโดยใช้สมการจำแนกกลุ่มได้ถูกต้อง 12 cases คิดเป็นร้อยละ 80.0 (12 ใน 15) เมื่อคิดรวมทั้งหมด (ทั้ง 30 cases) พบว่าสมการจำแนกกลุ่มได้ถูกต้องร้อยละ 90.0 (15+12 = 27 ใน 30)

สำหรับผลจากการเลือก Save จะไม่ปรากฏใน output แต่จะนำเสนอใน SPSS for windows Data Editor (File ข้อมูลในส่วนของ “Data View”) โดยสร้างเป็นตัวแปรต่อจากข้อมูลเดิมจำนวน 4 ตัวแปร ดังภาพ



File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

6 :

	Eng	Math	Thai	Science	Group	Dis_1	Dis1_1	Dis1_2	Dis2_2	var
1	44	78	50	45	1	1	.99865	.91550	.08450	
2	55	82	45	67	1	1	1.03256	.92155	.07845	
3	68	88	67	80	1	1	2.30444	.99592	.00408	
4	72	77	55	88	1	1	1.14539	.93894	.06106	
5	89	76	66	79	1	1	1.54046	.97529	.02471	
6	60	58	76	80	1	1	.87047	.88864	.11136	
7	53	83	47	45	1	1	1.17556	.94294	.05706	
8	76	62	58	34	1	1	.36632	.70558	.29442	
9	43	55	67	26	1	1	.31731	.68072	.31928	
10	52	80	66	45	1	1	1.78132	.98594	.01406	
11	65	85	73	56	1	1	2.37213	.99653	.00347	
12	50	74	74	78	1	1	1.75115	.98490	.01510	
13	60	48	77	77	1	1	.30970	.67676	.32324	
14	80	53	78	54	1	1	.65217	.82579	.17421	
15	79	62	80	58	1	1	1.27690	.95464	.04536	
16	88	66	55	62	2	1	.48301	.75996	.24004	
17	50	78	42	80	2	1	.66753	.83099	.16901	
18	43	32	56	69	2	2	-1.52294	.02574	.97426	
19	56	20	66	30	2	2	-1.83163	.01249	.98751	
20	35	67	70	34	2	1	1.16408	.94144	.05856	
21	47	50	55	43	2	2	-.48045	.24116	.75884	
22	44	43	40	32	2	2	-1.52281	.02575	.97425	
23	50	45	34	24	2	2	-1.65072	.01911	.98089	
24	63	32	22	25	2	2	-2.93021	.00092	.99908	

1.ตัวแปร Dis_1 เป็นตัวแปรที่ระบุถึงการเป็นสมาชิกกลุ่มของหน่วยวิเคราะห์ที่ได้จากการทำนายของสมการจำแนกกลุ่ม

2. ตัวแปร Dis1_1 เป็นตัวแปรที่บอกถึงคะแนนจำแนกของหน่วยวิเคราะห์ โดยทำนายจากความน้ำหนักและค่าที่วัดได้จากตัวแปร (Math กับ Thai) ในสมการจำแนก (ตัวอย่างนี้เป็นคะแนนจากสมการในรูปคะแนนดิบ)

3. ตัวแปร Dis1_2 เป็นตัวแปรที่บอกถึงโอกาสในการเป็นสมาชิกกลุ่ม 1 ของแต่ละหน่วยวิเคราะห์

4. ตัวแปร Dis2_2 เป็นตัวแปรที่บอกถึงโอกาสในการเป็นสมาชิกกลุ่ม 2 ของแต่ละหน่วยวิเคราะห์

เช่น ใน Case ที่ 1 มีคะแนนจำแนกเท่ากับ 0.99865 มีค่าเข้าใกล้ Group Centroids ของกลุ่มที่ 1 มากกว่ากลุ่มที่ 2 โดยมีโอกาสหรือความน่าจะเป็นการเป็นสมาชิกของกลุ่ม 1 เท่ากับ 0.91550 และมีโอกาสในการเป็นสมาชิกกลุ่ม 2 เพียง 0.08450 จึงจัดให้ Case ที่ 1 เป็นสมาชิกของกลุ่มที่ 1 (กลุ่มที่ได้จากการทำนายโดยการวิเคราะห์จำแนกกลุ่ม) ซึ่งผลการวิเคราะห์มีความสอดคล้องกับกลุ่มเดิม

6. บทสรุป

การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis : DA) เป็นเทคนิคการวิเคราะห์ตัวแปรพหุ (Multivariate Analysis) ที่มีวัตถุประสงค์หลักเพื่อสร้างสมการจำแนกกลุ่มที่ดีที่สุดจากข้อมูล และนำสมการจำแนกกลุ่มมาใช้พยากรณ์หน่วยวิเคราะห์ใหม่ว่าสมควรจัดให้อยู่ในกลุ่มใดลักษณะของตัวแปรที่นำมาวิเคราะห์เหมือนกับการวิเคราะห์ถดถอยโลจิสติคคือ ตัวแปรตาม (ตัวแปรที่ถูกจำแนกหรือถูกทำนาย) เป็นตัวแปรที่แบ่งเป็นกลุ่มๆ ตั้งแต่ 2 กลุ่มขึ้นไป ส่วนตัวแปรอิสระ (ตัวแปรที่ใช้ในการจำแนกหรือตัวแปรที่ใช้ในการทำนาย) เป็นตัวแปรต่อเนื่องที่มีระดับการวัดอยู่ในมาตราอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) แต่จะมีส่วนที่แตกต่างกันคือ ในเทคนิควิเคราะห์ถดถอยโลจิสติคตัวแปรอิสระ ถ้าไม่ได้อยู่ในมาตราอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) ไม่ควรแปลงเป็นตัวแปรหุ่น (Dummy Coding) เพราะจะทำให้เกิดปัญหาในการวิเคราะห์ แต่ในเทคนิคการวิเคราะห์จำแนกกลุ่ม สามารถแปลงตัวแปรอิสระที่ไม่ได้อยู่ในมาตราการวัดอันตรภาค (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) ให้เป็นตัวแปรหุ่น (Dummy Coding) ได้ และการวิเคราะห์จำแนกกลุ่ม ตัวแปรอิสระต่องเปนอิสระและถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติหลายตัวแปร รวมทั้งกลุ่มแต่ละกลุ่มต้องมีความแปรปรวนและความแปรปรวนร่วมเท่ากัน (Equal variance/covariance matrices) แต่ในการวิเคราะห์ถดถอยโลจิสติคไม่มีข้อกำหนดนี้ ดังนั้นก่อนที่จะทำการวิเคราะห์โดยใช้เทคนิคการวิเคราะห์จำแนกกลุ่มจึงต้องมีการตรวจสอบข้อตกลงเบื้องต้นเสียก่อน

แบบฝึกหัดบทที่ 5
เรื่อง การวิเคราะห์จำแนกกลุ่ม

---

1. บริษัทจำหน่ายรถยนต์แห่งหนึ่งแบ่งกลุ่มพนักงานขายออกเป็น 2 กลุ่ม คือ กลุ่มที่ควรได้รับรางวัล และกลุ่มที่ไม่ควรได้รับรางวัล ซึ่งคาดว่าปัจจัยหรือตัวแปรที่ทำให้พนักงานถูกจัดอยู่ในกลุ่มที่ต่างกันคือ ยอดขาย (คัน : เดือน) อายุการทำงาน และคะแนนความพึงพอใจของลูกค้า (คะแนนเต็ม 20 คะแนน) เก็บรวบรวมข้อมูลจากตัวอย่างพนักงานจำนวน 50 คน ได้ข้อมูลดังนี้

พนักงานคนที่	ยอดขาย (คัน : เดือน)	อายุการทำงาน (ปี)	คะแนนความพึงพอใจของลูกค้า	กลุ่ม (1= กลุ่มที่ควรได้รับรางวัล , 2= กลุ่มที่ไม่ควรได้รับรางวัล)
1	10	7	16	1
2	15	8	17	1
3	18	6	16	1
4	12	8	15	1
5	10	7	19	1
6	15	4	18	1
7	10	5	16	1
8	18	9	15	1
9	20	10	19	1
10	14	12	19	1
11	13	10	18	1
12	16	5	16	1
13	17	9	19	1
14	10	8	18	1
15	12	10	18	1
16	14	12	18	1
17	15	13	19	1
18	15	10	19	1
19	13	7	17	1
20	12	8	17	1
21	10	6	17	1
22	11	7	16	1
23	10	8	16	1

พนักงานคนที่	ยอดขาย (คัน : เดือน)	อายุการทำงาน (ปี)	คะแนนความพึงพอใจของลูกค้า	กลุ่ม (1= กลุ่มที่ควรได้รับรางวัล , 2= กลุ่มที่ไม่ควรได้รับรางวัล)
24	7	4	12	2
25	5	7	12	2
26	4	5	14	2
27	7	3	13	2
28	6	4	12	2
29	4	5	13	2
30	5	3	13	2
31	3	8	12	2
32	7	10	14	2
33	8	7	16	2
34	4	6	13	2
35	5	8	12	2
36	6	8	13	2
37	3	4	14	2
38	8	4	13	2
39	9	5	14	2
40	4	6	12	2
41	5	6	14	2
42	7	4	13	2
43	6	3	14	2
44	6	3	12	2
45	7	2	13	2
46	7	3	12	2
47	8	4	13	2
48	4	5	14	2
49	5	4	14	2
50	5	3	14	2

1.1 จงตรวจสอบข้อตกลงเบื้องต้นของเทคนิคการวิเคราะห์ Discriminant Analysis พร้อมทั้งอธิบายผลลัพธ์ที่ได้มาโดยละเอียด

1.2 จงวิเคราะห์ข้อมูลโดยใช้เทคนิค Discriminant Analysis เพื่อหาสมการจำแนกที่เหมาะสมสำหรับนำไปใช้ในการจัดกลุ่มพนักงาน พร้อมทั้งอธิบายผลลัพธ์ที่ได้มาโดยละเอียด

2. สถานบำบัดผู้ป่วยยาเสพติดแห่งหนึ่ง แบ่งกลุ่มของผู้เข้ารับการบำบัดเป็น 2 กลุ่มคือ กลับมาบำบัดซ้ำ กับ ไม่กลับมาบำบัดซ้ำ โดยคาดว่าปัจจัยที่ทำให้ผู้เข้ารับการบำบัดอยู่ในกลุ่มที่ต่างกัน ได้แก่ รายได้ต่อเดือน จำนวนปีที่เคยเสพ อายุ เก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างจำนวน 30 คน ได้ข้อมูลดังนี้

คนที่	รายได้ต่อเดือน	จำนวนปีที่เคยเสพ	อายุ	กลุ่ม (1= กลับมาบำบัดซ้ำ, 2= ไม่กลับมาบำบัดซ้ำ)
1	5,000	3	18	1
2	4,000	3	20	1
3	3,500	3	18	1
4	7,500	4	20	1
5	8,000	5	22	1
6	9,000	5	25	1
7	8,500	5	27	1
8	5,000	3	18	1
9	6,000	2	20	1
10	6,000	5	22	1
11	9,000	5	27	1
12	9,500	5	29	1
13	8,700	4	28	1
14	4,000	4	15	2
15	8,000	1	14	2
16	9,000	3	16	2
17	8,500	2	27	2
18	5,000	4	29	2
19	5,000	5	28	2
20	4,000	3	30	2
21	3,500	4	32	2
22	7,500	5	28	2
23	8,000	4	19	2
24	9,000	4	29	2
25	8,500	5	30	2
26	10,000	2	32	2
27	9,000	4	28	2

คนที่	รายได้ต่อเดือน	จำนวนปีที่เคยเสพ	อายุ	กลุ่ม (1= กลับมาบำบัดซ้ำ, 2= ไม่กลับมาบำบัดซ้ำ)
28	6,000	3	29	2
29	6,000	4	31	2
30	9,000	1	19	2

2.1 จงตรวจสอบข้อตกลงเบื้องต้นของเทคนิคการวิเคราะห์ Discriminant Analysis พร้อมทั้งอธิบายผลลัพธ์ที่ได้มาโดยละเอียด

2.2 จงวิเคราะห์ข้อมูลโดยใช้เทคนิค Discriminant Analysis เพื่อหาสมการจำแนกที่เหมาะสมสำหรับนำไปใช้ในการจัดกลุ่มผู้เข้ารับการบำบัด พร้อมทั้งอธิบายผลลัพธ์ที่ได้มาโดยละเอียด

เอกสารอ้างอิงบทที่ 5

กัลยา วานิชย์บัญชา. (2551). **การวิเคราะห์สถิติขั้นสูงด้วย SPSS for Windows**. พิมพ์ครั้งที่ 6. กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

ทรงศักดิ์ ภูสีอ่อน. (2554). **การประยุกต์ใช้ SPSS วิเคราะห์ข้อมูลงานวิจัย**. พิมพ์ครั้งที่ 5. มหาสารคาม : มหาวิทยาลัยมหาสารคาม.

สุชาติ ประสิทธิ์รัฐสินธุ์. 2540. **เทคนิคการวิเคราะห์ตัวแปรหลายตัวสำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์.** พิมพ์ครั้งที่ 4. กรุงเทพฯ : เลี่ยงเชียง.

การวิเคราะห์การจำแนกกลุ่ม Discriminant Analysis. (2557). ค้นเมื่อ 6 พฤษภาคม 2557, จาก www.saruthipong.com/port/document/299-705/299-705-6.pdf.

การวิเคราะห์การจำแนกประเภท. (2557). ค้นเมื่อ 6 พฤษภาคม 2557, จาก https://docs.google.com/document/d/.../edit?hl=th.