



ระเบียบวิธีที่ใช้ในการวิจัย

จากการศึกษารายงานการวิจัยและทฤษฎีต่าง ๆ ที่เกี่ยวข้อง ตัวแปรที่คิดว่าน่าจะมีผลต่อทัศนคติของชาวชนบทภาคตะวันออกเฉียงเหนือในการไปทำงานต่างถิ่นที่นำมาใช้ในการวิเคราะห์ครั้งนี้มีดังต่อไปนี้ คือ

๑. สถานะของบุคคลที่มีความเกี่ยวข้องกับหัวหน้าครัวเรือน ในที่นี้แบ่งออกเป็น ๔ กลุ่มคือ
 - ๑.๑ หัวหน้าครัวเรือน
 - ๑.๒ สามีนหรือภรรยา
 - ๑.๓ บุตรที่ยังไม่สมรส
 - ๑.๔ อื่น ๆ เช่น บุตรที่สมรสแล้ว เขยหรือสะใภ้ หลาน ปิตามารดา ญาติ คนอาศัยและอื่น ๆ
๒. จำนวนสมาชิกในครัวเรือน
๓. จำนวนสมาชิกในครัวเรือนที่ยังไม่ได้ประกอบอาชีพ
๔. เพศ แบ่งเป็น ๒ กลุ่ม คือ เพศชายและเพศหญิง
๕. อายุเต็มปี ใช้อายุปัจจุบันโดยคิดอายุเต็มปีของสมาชิกในครัวเรือนทุกคน เศษของเดือนให้ตัดทิ้ง
๖. สถานภาพสมรส แบ่งออกเป็น ๒ กลุ่มคือ
 - ๖.๑ โสด
 - ๖.๒ สมรส หม้าย หย่า แยกกันอยู่
๗. ระดับการศึกษา ในที่นี้วัดด้วยจำนวนปีตั้งแต่เริ่มเรียนจนถึงจบการศึกษาสูงสุด

๘. เคยย้ายมาจากถิ่นอื่นหรือไม่ แบ่งเป็น ๒ กลุ่มคือ

๘.๑ เคยย้ายมาจากถิ่นอื่น

๘.๒ ไม่เคย

๙. สถานภาพการมีงานทำ แบ่งออกเป็น ๓ กลุ่ม ดังนี้

๙.๑ มีงานทำหรือกำลังรอฤดูกาล เกษตร

๙.๒ กำลังว่างงาน

๙.๓ ไม่ได้ทำงาน

๑๐. สถานภาพการทำงาน ของผู้มีงานทำหรือรอฤดูกาล เกษตร

๑๐.๑ นายจ้าง

๑๐.๒ ลูกจ้างรัฐบาล

๑๐.๓ ลูกจ้างเอกชน

๑๐.๔ ทำงานส่วนตัว

๑๐.๕ ช่วยธุรกิจในครัวเรือนโดยไม่ได้รับค่าจ้าง

๑๑. อาชีพ ในที่นี้แบ่งกลุ่มอาชีพออกเป็น ๔ กลุ่ม โดยพิจารณาจากข้อมูลที่ได้จาก

โครงการสำรวจแรงงานชนบทในประเทศไทย พ.ศ. ๒๕๒๐ และจัดกลุ่มอาชีพโดยจำแนกตามการจัดประเภทมาตรฐานอาชีพ (ประเทศไทย) ซึ่งกรมแรงงาน กระทรวงมหาดไทย ได้ร่วมงานกับองค์การแรงงานระหว่างประเทศ และนักการศึกษา นักภาษาศาสตร์ นักธุรกิจ นักอุตสาหกรรม ตลอดจนนายช่างและผู้ประกอบอาชีพสาขาต่าง ๆ จัดทำเอกสารนี้ขึ้นมา ซึ่งในการวิจัยเรื่องนี้จะแบ่งเพียง ๔ กลุ่ม ดังนี้

๑๑.๑ ผู้ทำงานกลีกรรรม ผู้เลี้ยงสัตว์และผู้ทำงานป่าไม้ ชาวประมงและพราน

๑๑.๒ ผู้ปฏิบัติงานที่เกี่ยวกับการผลิต และผู้ปฏิบัติงานที่เกี่ยวข้องกันพนักงานใช้
อุปกรณ์ในการขนส่ง และคนงานหรือกรรมกร

๑๑.๓ ผู้ปฏิบัติงานที่ใช้วิชาชีพ วิชาการ และผู้ปฏิบัติงานที่เกี่ยวข้องกัน

๑๑.๔ อื่น ๆ

๑๒. จำนวนเดือนที่ทำงานในรอบปี
๑๓. ประกอบอาชีพในขณะรอการ เก็บเกี่ยวหรือไม่
- ๑๓.๑ ไม่ทำ
- ๑๓.๒ ทำ
๑๔. ประกอบอาชีพเมื่อพ้นระยะเก็บเกี่ยวหรือไม่
- ๑๔.๑ ไม่ทำ
- ๑๔.๒ ทำ
๑๕. รายได้ทั้งปี เป็นรายได้ทั้งปีจากการประกอบอาชีพหรือทำงานทุกชนิดของแต่ละคนในครัวเรือน
๑๖. รายได้ของครัวเรือน เป็นรายได้ทั้งปีของสมาชิกในครัวเรือนทุกคนรวมกัน
๑๗. ความต้องการทำงานมากขึ้น
- ๑๗.๑ ต้องการ
- ๑๗.๒ ไม่ต้องการ ไม่ทราบ
๑๘. ความต้องการฝึกอาชีพ
- ๑๘.๑ ต้องการฝึกอาชีพ ช่างฝีมือ
- ๑๘.๒ ต้องการฝึกอาชีพ งานเกษตร
- ๑๘.๓ ต้องการฝึกอาชีพ อื่น ๆ
- ๑๘.๔ ไม่ต้องการฝึก
๑๙. จำนวนคนในครัวเรือนที่ไปอยู่ที่อื่น
๒๐. จากแผนการทำงานในรอบปี ๒๕๖๕ คิดไปทำงานในถิ่นอื่นหรือไม่
- ๒๐.๑ คิด
- ๒๐.๒ ไม่คิด
- ๒๐.๓ ไม่ทราบ

นอกจากนั้นยังมีตัวแปรอื่น ๆ ที่เกี่ยวข้องกับกลุ่มคนที่มีความคิดจะไปทำงานต่างถิ่น ดังนี้คือ

๑. จะไปทำงานอะไร

- ๑.๑ ทำงานอะไรก็ได้
- ๑.๒ พนักงาน เสมียน
- ๑.๓ คนงานโรงงาน รับจ้างตามโรงงาน
- ๑.๔ กรรมกร คนงานก่อสร้าง
- ๑.๕ รับจ้างทั่วไป
- ๑.๖ รับใช้ในบ้าน
- ๑.๗ พวกช่างฝีมือต่าง ๆ
- ๑.๘ พนักงานขับรถ
- ๑.๙ รับราชการ
- ๑.๑๐ อื่น ๆ

๒. จะไปทำงานทำที่ไหน

- ๒.๑ ที่ไหนก็ได้
- ๒.๒ จังหวัดเดิม
- ๒.๓ จังหวัดอื่น ๆ ยกเว้น
- ๒.๔ กรุงเทพมหานคร
- ๒.๕ ต่างประเทศ

๓. คิดไปทำงานนานเท่าไร

- ๓.๑ ถาวร
- ๓.๒ ชั่วคราว

๔. รายได้ที่คาดว่าจะได้จากการไปทำงานต่างถิ่น คิดเฉลี่ยทั้งปี

ในการศึกษาเรื่องปัจจัยที่มีผลต่อทัศนคติของชาวชนบทภาคตะวันออกเฉียงเหนือในการไปทำงานต่างถิ่นนี้ ผู้วิจัยจะใช้ข้อมูลจากชาวชนบทที่มีอายุ ๑๑ ปีขึ้นไป ที่อาศัยอยู่ในภาคตะวันออกเฉียงเหนือ มาวิเคราะห์เป็นประชากรตัวอย่างของการศึกษา ซึ่งสามารถจัดกลุ่มตัวแปรดังนี้

ตัวแปรที่ใช้ในการวิจัย

ในการวิจัยนี้ ได้กำหนดตัวแปรที่ใช้ในการศึกษาตามปัจจัยต่าง ๆ ดังนี้

ตัวแปรอิสระ

ปัจจัยทางเศรษฐกิจ

- รายได้จากอาชีพที่ทำอยู่ทุกอาชีพ
- รายได้ของครัวเรือน

ปัจจัยทางประชากรและสังคม / อายุ

- ✓ - เพศ
- ✓ - ระดับการศึกษา
- ✓ - สถานภาพสมรส
 - จำนวนสมาชิกในครัวเรือน
 - จำนวนสมาชิกในครัวเรือนที่ยังไม่ได้ประกอบอาชีพ
 - จำนวนคนในครัวเรือนที่ย้ายไปอยู่ที่อื่น
 - สถานะของบุคคลที่มีความเกี่ยวข้องกับหัวหน้าครัวเรือน
- ✓ - สถานภาพการมีงานทำ
 - สถานภาพการทำงาน
- ✓ - อาชีพ
 - จำนวนเดือนที่ทำงานในรอบปี
 - ความต้องการฝึกอาชีพ
 - ความต้องการทำงานมากขึ้น
- ✓ - ประสบการณ์ในการย้ายถิ่น

ปัจจัยด้านอื่น ๆ ที่เกี่ยวกับความคิดในการไปทำงานต่างถิ่น

- ✓ - ระยะเวลาที่คิดว่าจะไปทำงาน
 - ตำแหน่งงานที่คิดจะไปทำงาน
 - สถานที่ที่คาดว่าจะไปทำ
 - รายได้ที่คาดว่าจะได้รับจากการไปทำงานต่างถิ่น

ตัวแปรตามทัศนคติของชาวชนบทภาคตะวันออกเฉียงเหนือในการไปทำงานต่างถิ่น วัดจาก
มีความคิดว่าจะไปทำงาน ไม่มีความคิดว่าจะไปทำ ไม่ทราบ

การวิเคราะห์ปัจจัยที่มีผลต่อทัศนคติของชาวชนบทภาคตะวันออกเฉียงเหนือในการไป
ทำงานต่างถิ่น จะใช้วิธีการวิเคราะห์ ๓ วิธีด้วยกันคือ

๑. ในการทดสอบความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม เพื่อต้องการทราบว่า
ตัวแปรอิสระแต่ละตัวจะมีความสัมพันธ์กับตัวแปรตามหรือไม่ โดยแยกศึกษาทีละปัจจัยใช้การทดสอบ
Chi-Square กำหนดให้มีระดับนัยสำคัญเท่ากับ ๐.๐๕

๒. ในการวิเคราะห์เพื่อต้องการทราบลำดับอิทธิพลและทิศทางของปัจจัยที่มีผลต่อทัศนคติ
ของชาวชนบทภาคตะวันออกเฉียงเหนือในการไปทำงานต่างถิ่น จะนำตัวแปรอิสระทุกตัวที่มีความ
สัมพันธ์กับตัวแปรตามมาทำการวิเคราะห์ร่วมกัน ตัวแปรที่ใช้เป็นตัวแปรเชิงปริมาณหรือถ้าเป็นตัวแปร
เชิงคุณภาพจะใช้วิธีจัดตัวแปรใหม่โดยทำให้เป็นตัวแปรดัมมี่ (Dummy Variables) ใช้วิธี
Discriminant Analysis ในการวิเคราะห์เพื่อทดสอบการจำแนกประเภทของกลุ่มชาวชนบท
ที่มีทัศนคติในการไปทำงานต่างถิ่นว่าคิดจะไปทำงานต่างถิ่นหรือไม่

๓. ในการวิเคราะห์กลุ่มที่มีความคิดว่าจะไปทำงานต่างถิ่นนั้น จะใช้สถิติเชิงพรรณนาใน
การบรรยายลักษณะทางประชากรบางประการ เช่น อายุ เพศ สถานภาพสมรสว่าจะมีอิทธิพล
อย่างไรบ้าง โดยใช้ค่าร้อยละ ค่าความถี่และค่าเฉลี่ย

๒.๑ การทดสอบปัจจัยที่มีผลต่อทัศนคติในการไปทำงานต่างถิ่น

จะนำตัวแปรอิสระทุกตัวที่ใช้ในการวิจัยนี้มาทำการทดสอบศึกษาดูว่า มีตัวแปรใดบ้างที่มี
ความสัมพันธ์ และไม่มีความสัมพันธ์กับทัศนคติในการไปทำงานต่างถิ่น จุดประสงค์ในการทดสอบ
ตัวแปรในครั้งนี้เพื่อต้องการทราบถึงความสัมพันธ์ของตัวแปรอิสระทุกตัวที่มีต่อตัวแปรตามว่าจะมีความ
สัมพันธ์กันอย่างไร เมื่อแบ่งกลุ่มของตัวแปรอิสระแต่ละตัวให้แตกต่างกัน และแบ่งทัศนคติในการไป
ทำงานต่างถิ่นออกเป็น ๓ ลักษณะที่แตกต่างกันคือ คิดจะไปทำงานในถิ่นอื่น ไม่ทราบ ไม่คิดจะไป

ในการทดสอบปัจจัยต่าง ๆ เหล่านี้ ผู้วิจัยได้ใช้วิธีการทางสถิติคือ การทดสอบ Chi-Square เนื่องจากในการทดสอบนี้ต้องการทราบว่าตัวแปรอิสระทุกตัวและตัวแปรตามเป็นอิสระต่อกันหรือมีความสัมพันธ์ต่อกันหรือไม่ โดยตั้งสมมติฐานว่าตัวแปรที่นำมาหาความสัมพันธ์กันนั้นเป็นอิสระต่อกัน หรือไม่มีความสัมพันธ์กัน

ได้จัดแบ่งข้อมูลของตัวแปรอิสระ เป็นกลุ่มขึ้นตามลักษณะของตัวแปรและจะวิเคราะห์ข้อมูลจากแบบสอบถาม ใช้โปรแกรมสำเร็จรูป SPSS (Statistical Package for Social Science) รายละเอียดของการวิเคราะห์มีดังนี้

๑. หาค่าของความถี่ และร้อยละของความถี่ ในคำตอบแต่ละข้อ

๒. สร้างตารางจากการแจกแจงสองทาง (Contingency Table) ระหว่างตัวแปรที่น่าสนใจ เช่น ความต้องการฝึกอาชีพ กับทัศนคติของชาวชนบทภาคตะวันออกเฉียงเหนือในการไปทำงานต่างถิ่น

ทดสอบสมมติฐานเกี่ยวกับความเป็นอิสระของตัวแปรแต่ละคู่ในข้อ ๒ โดยตั้งสมมติฐานเพื่อการทดสอบเป็น

H_0 : ตัวแปรทั้งสองเป็นอิสระต่อกัน

H_1 : ตัวแปรทั้งสองไม่เป็นอิสระต่อกัน

๒.๑.๑ ตัวสถิติที่ใช้ในการทดสอบคือ

$$\chi^2_c = \sum_{m=1}^r \sum_{n=1}^k \frac{(O_{mn} - E_{mn})^2}{E_{mn}}$$

O_{mn} คือความถี่ที่ได้จากการสังเกต ในแถวที่ m สดมภ์ที่ n

E_{mn} คือความถี่ที่ได้จากการคาดหวังหรือทฤษฎี ในแถวที่ m

$$\text{สดมภ์ที่ } n = \frac{(C_n R_m)}{N}$$

C_n = ผลรวมของความถี่ในสดมภ์ที่ n

R_m = ผลรวมของความถี่ในแถวที่ m

N = จำนวนความถี่รวมทั้งหมด

r คือ จำนวนแถว

k คือ จำนวนสดมภ์

d.f คือ ชั้นของความเป็นอิสระ (degree of freedom) = $(r - 1) (k - 1)$

		สดมภ์			
		1	k	
	1				R_1
	2				R_2
แถว	.				.
	.				.
	.				.
	r				R_r
		C_1	C_k	N

k สดมภ์

๒.๑.๒ กำหนดระดับนัยสำคัญ α

๒.๑.๓ ขอบเขตการปฏิเสธสมมติฐาน

$$X^2_C \geq X^2_{(1-\alpha), (r-1)(k-1)}$$

๒.๒ การวิเคราะห์จำแนกประเภท (Discriminant Analysis)

การวิเคราะห์โดยวิธีนี้เพื่อต้องการศึกษาให้ทราบถึงลำดับอิทธิพลและทิศทางของปัจจัยที่มีผลต่อทัศนคติของชาวชนบทภาคตะวันออกเฉียงเหนือในการไปทำงานต่างถิ่น และการวิเคราะห์จำแนกประเภทนี้จุดประสงค์เพื่อคัดเลือกตัวแปรกลุ่มหนึ่งหรือชุดหนึ่ง ซึ่งผู้วิจัยคิดว่าตัวแปรเหล่านี้มีความสัมพันธ์กับทัศนคติในการไปทำงานต่างถิ่น จนถึงขั้นที่จะสามารถแยกประชากรออกเป็น ๒ กลุ่ม คือ กลุ่มที่คิดว่าจะไปทำงานในถิ่นอื่นและไม่คิดไป ซึ่งประชากรแต่ละกลุ่มนี้จะมีตัวแปรที่สัมพันธ์กับกลุ่มเป็นตัวแปรที่ใช้แบ่งแยกกลุ่มออกจากกลุ่มอื่น ๆ ได้

สมการที่ได้เรียกว่า สมการจำแนกประเภท จำนวนของสมการที่ได้จะน้อยกว่าจำนวนกลุ่มของประชากรที่ต้องการจำแนก ๑ กลุ่มเสมอ หรือถ้ามีจำนวนกลุ่มมากกว่าตัวแปร จำนวนสมการจะเท่ากับจำนวนตัวแปรที่ใช้จำแนกประเภท (discriminant variables) ซึ่งสมการแต่ละสมการที่ได้จะพยายามแบ่งแยกกลุ่มต่าง ๆ ที่ต่างกันให้มากที่สุด

ประโยชน์ของสมการจำแนกประเภทมี ๒ ประการคือ การวิเคราะห์และการจำแนกประเภทในด้านการวิเคราะห์จะมีสถิติต่าง ๆ เกี่ยวกับตัวแปรแต่ละตัวในสมการที่ใช้ในการทดสอบความสำคัญของตัวแปรซึ่งช่วยในการตีความหมายข้อมูล ค่าสัมประสิทธิ์ของตัวแปรต่าง ๆ ที่ได้ ในสมการบ่งชี้ถึงน้ำหนักหรือความสัมพันธ์ของตัวแปรนั้นในสมการนั้นต่อการ เป็นสมาชิกของกลุ่ม

ในด้านการจำแนกประเภท เมื่อได้สมการมาแล้วก็สามารถที่จะระบุว่ากรณีใดจะเป็นสมาชิกของกลุ่มใดได้โดยอาศัยตัวแปรและค่าสัมประสิทธิ์ต่าง ๆ ที่ปรากฏอยู่ในสมการจำแนกเป็นเครื่องมือช่วยระบุจำแนกประเภท

ความสามารถของสมการในการจำแนกประเภทอาจทดสอบได้จากการจำแนกประเภทของกรณีที่เราว่าเป็นสมาชิกของกลุ่มใด และใช้สมการประมาณเปรียบเทียบว่าตรงกับที่เป็นจริงหรือไม่

สมการจำแนกประเภทสำหรับการจำแนกประชากรออกเป็นสองกลุ่ม

$$Y_{it} = \beta_1 X_{i1t} + \beta_2 X_{i2t} + \dots + \beta_p X_{ipt} \quad , \quad i = 1, 2$$

$$j = 1, 2, \dots, p$$

$$t = 1, 2, \dots, n_i$$

$$\dots \dots \dots (1)$$

X_{ijt} แทนค่าของตัวแปรตัวที่ j ในคำสั่งเกิด t ของกลุ่มที่ i

การหาสมการที่เหมาะสมวิธีหนึ่งคือการทำให้ค่า λ มากที่สุดโดย

$$\lambda = \frac{\text{between-group variation}}{\text{within-group variation}} \quad \dots \dots \dots (1 a)$$

กรณีที่ต้องการจำแนกประชากรออกเป็น ๒ กลุ่ม จะได้ว่า

$$\mu_{Y_1} = \beta_1 \mu_{X_{11}} + \beta_2 \mu_{X_{12}} + \dots + \beta_p \mu_{X_{1p}} \quad \dots \dots \dots (2)$$

และ

$$\mu_{Y_2} = \beta_1 \mu_{X_{21}} + \beta_2 \mu_{X_{22}} + \dots + \beta_p \mu_{X_{2p}} \quad \dots \dots \dots (3)$$

$\mu_{X_{ij}}$ แทนค่าเฉลี่ยของตัวแปรตัวที่ j ในกลุ่มที่ i

จากสมการที่ 1, 2, 3 หากค่าเบี่ยงเบนจากค่าเฉลี่ย μ_{Y_i} ของ Y_t แต่ละตัวได้ดังนี้

$$Y_{it} - \mu_{Y_i} = \beta_1 (X_{i1t} - \mu_{X_{i1}}) + \beta_2 (X_{i2t} - \mu_{X_{i2}}) + \dots + \beta_p (X_{ipt} - \mu_{X_{ip}}) \dots (4)$$

ผลรวมกำลังสองของ $Y_{it} - \mu_{Y_i}$ เป็นค่าวัดของความแปรปรวนภายในกลุ่ม (within-group variation)

จะประมาณค่า μ_{Y_i} ด้วย $\bar{Y}_i = \sum_{t=1}^{n_i} Y_{it}/n_i$, $i = 1, 2$

ดังนั้นจะประมาณค่าเฉลี่ยประชากรของตัวแปร X ด้วย $\bar{X}_{ij} = \sum_{t=1}^{n_i} X_{ijt}/n_i$

ซึ่ง $i = 1, 2$ และ $j = 1, 2, \dots, p$ ดังนั้นสำหรับกลุ่มตัวอย่าง ความแปรปรวนภายในกลุ่ม (within-group variation) จากสมการ (๔) จะได้ว่า

$$\sum_{i=1}^2 \sum_{t=1}^{n_i} (Y_{it} - \bar{Y}_i)^2 = \sum_{i=1}^2 \sum_{t=1}^{n_i} [\beta_1 (X_{i1t} - \bar{X}_{i1}) + \beta_2 (X_{i2t} - \bar{X}_{i2}) + \dots + \beta_p (X_{ipt} - \bar{X}_{ip})]^2 \dots \dots \dots (5)$$

เขียนสมการใหม่โดยทำในรูป matrix X' ดังนี้

$$X' = \begin{bmatrix} X_{111} & X_{112} & \dots & X_{11n_1} & | & X_{211} & X_{212} & \dots & X_{21n_2} \\ X_{121} & X_{122} & \dots & X_{12n_1} & | & X_{221} & X_{222} & \dots & X_{22n_2} \\ \vdots & & & \vdots & | & \vdots & \vdots & & \vdots \\ X_{1p1} & X_{1p2} & \dots & X_{1pn_1} & | & X_{2p1} & X_{2p2} & \dots & X_{2pn_2} \end{bmatrix}$$

แบ่ง matrix X' ออกเป็น ๒ ส่วนโดยใช้เส้นขีดตามแนวตั้งดังนี้

$$X' = [X'_1 \quad \vdots \quad X'_2]$$

X'_1 ประกอบด้วยจำนวนค่าสังเกต n_1 ซึ่งแต่ละค่ามีตัวแปร p ตัว (กลุ่ม I)

X'_2 ประกอบด้วยจำนวนค่าสังเกต n_2 ซึ่งแต่ละค่ามีตัวแปร p ตัว (กลุ่ม II)

ทั้ง X'_1 และ X'_2 เป็น matrix ของค่าสังเกตซึ่งแต่ละ matrix จะได้เวกเตอร์ค่าเฉลี่ย ดังนี้

$$\bar{X}'_1 = [\bar{X}_{11} \quad \bar{X}_{12} \quad \dots \quad \bar{X}_{1p}]$$

$$\bar{X}'_2 = [\bar{X}_{21} \quad \bar{X}_{22} \quad \dots \quad \bar{X}_{2p}]$$

โดยมี covariation matrices $x_1'x_1$ และ $x_2'x_2$ ให้ $\beta' = [\beta_1 \beta_2 \dots \beta_p]$
ดังนั้นสมการ (๔) สามารถเขียนใหม่ได้ดังนี้

$$\begin{aligned} \sum_{i=1}^2 \sum_{t=1}^{n_i} (y_{it} - \bar{y}_i)^2 &= \beta'x_1'x_1\beta + \beta'x_2'x_2\beta \\ &= \beta'(x_1'x_1 + x_2'x_2)\beta \end{aligned}$$

และ covariance matrix Σ ประมาณด้วย S_* คือ pooled covariance matrix

$$\text{โดย } S_* = \frac{1}{n_1+n_2-2} (x_1'x_1 + x_2'x_2) = \frac{\left[\sum_{t=1}^{n_1} (x_{1t}-\bar{x}_1)(x_{1t}-\bar{x}_1)' + \sum_{t=1}^{n_2} (x_{2t}-\bar{x}_2)(x_{2t}-\bar{x}_2)' \right]}{n_1 + n_2 - 2}$$

ดังนั้น within-group variation $\sum_{i=1}^2 \sum_{t=1}^{n_i} (y_{it} - \bar{y}_i)^2 = \beta'[(n_1+n_2-2)S_*]\beta$

และ between-group variation สำหรับกลุ่มตัวอย่าง คือ $(\bar{y}_1 - \bar{y}_2)^2$ จากสมการ

(2) และ (3) ค่าเฉลี่ยของประชากรแทนด้วยค่าเฉลี่ยของตัวอย่าง จะได้

$$\begin{aligned} (\bar{y}_1 - \bar{y}_2)^2 &= (\beta'\bar{x}_1 - \beta'\bar{x}_2)^2 \\ &= \beta'(\bar{x}_1 - \bar{x}_2)(\bar{x}_1 - \bar{x}_2)'\beta \end{aligned}$$

จากสมการ (1a) ให้ $L = \frac{\text{between-group variation}}{\text{within-group variation}} = \frac{1}{n_1+n_2-2} \frac{\beta'(\bar{x}_1 - \bar{x}_2)(\bar{x}_1 - \bar{x}_2)'\beta}{\beta'S_*\beta}$

ดังนั้นการทำให้ค่า L มากที่สุดคือการทำให้อัตราส่วน L มากที่สุด โดย

$$L = \frac{\beta'(\bar{x}_1 - \bar{x}_2)(\bar{x}_1 - \bar{x}_2)'\beta}{\beta'S_*\beta} \dots \dots \dots (6)$$

$\frac{\partial L}{\partial \beta} = 0$ จะได้ $c(\bar{x}_1 - \bar{x}_2) - S_*\beta = 0$

$$\hat{\beta} = cS_*^{-1}(\bar{x}_1 - \bar{x}_2)$$

กำหนดให้ $c = 1$ ดังนั้น $\hat{\beta} = S_*^{-1} (\bar{X}_1 - \bar{X}_2)$

แทนค่า $\hat{\beta}$ ในสมการ (6) $\hat{d} = (\bar{X}_1 - \bar{X}_2)' S_*^{-1} (\bar{X}_1 - \bar{X}_2) = D^2$

ซึ่งเรียกว่า Mahalanobis' D^2 หรือ generalized distance

กำหนดให้ Y^* เป็นเส้นแบ่งระหว่าง μ_{Y_1} และ μ_{Y_2} สำหรับตัวอย่างนั้นจะแสดง
ค่า $Y^* = \frac{1}{2} (\bar{Y}_1 + \bar{Y}_2)$ โดย $\bar{Y}_1 = (\bar{X}_1 - \bar{X}_2)' S_*^{-1} \bar{X}_1$ และ $\bar{Y}_2 = (\bar{X}_1 - \bar{X}_2)' S_*^{-1} \bar{X}_2$

$$\begin{aligned} \text{ดังนั้นจาก } Y^* &= \frac{1}{2} \hat{\beta}' (\bar{X}_1 + \bar{X}_2) \\ &= \frac{1}{2} (\bar{X}_1 - \bar{X}_2)' S_*^{-1} (\bar{X}_1 + \bar{X}_2) \end{aligned}$$

กฎการจำแนกประเภท (Discriminant Rule)

จัดค่าสังเกตใหม่ อยู่ในกลุ่ม I ถ้า $Y_{it} \geq Y^*$

จัดค่าสังเกตใหม่ อยู่ในกลุ่ม II ถ้า $Y_{it} < Y^*$

หรือกำหนดเวกเตอร์ Y' โดย

$$Y' = [Y_{11} \ Y_{12} \ \dots \ Y_{1n_1} \ Y_{21} \ Y_{22} \ \dots \ Y_{2n_2}]$$

จากสมการ (1) จะเขียนรูปแบบใหม่สำหรับกลุ่มตัวอย่างได้ดังนี้

$$Y = \hat{\beta} X'$$

ดังนั้น จะจัดให้อยู่ในกลุ่ม I ถ้า $\hat{\beta} X' - \frac{1}{2} \hat{\beta}' (\bar{X}_1 + \bar{X}_2) \geq 0$

และจัดให้อยู่ในกลุ่ม II ถ้า $\hat{\beta} X' - \frac{1}{2} \hat{\beta}' (\bar{X}_1 + \bar{X}_2) < 0$

วิธีการคัดเลือกตัวแปรเข้าไปในสมการจำแนกประเภท (Selection of Variables)

จากวิธีการที่กล่าวมาแล้วข้างต้นเป็นหลักการโดยทั่ว ๆ ไปของวิธีการสร้างสมการจำแนกประเภท
ในการศึกษาวิจัยมีตัวแปรจำนวนมาก การคัดเลือกตัวแปรให้เหลือจำนวนน้อยที่สุดแต่มีความสามารถ
ในการจำแนกมากที่สุดนั้น เราสามารถทำได้โดยใช้การคัดเลือกตัวแปรอิสระทีละตัว โดยที่จะหา
ตัวแปรที่ดีที่สุดตัวแรก และตัวแปรที่ดีที่สุดตัวต่อ ๆ ไป ที่จะช่วยการจำแนกให้ดีขึ้นตามลำดับ

ซึ่งในการเลือกตัวแปรอิสระเพื่อศึกษาในสมการวิเคราะห์จำแนกประเภท ผู้วิจัยได้คำนึงถึงปัญหาของความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) ด้วย ในการคัดเลือกตัวแปรทำนาย (Predict) ได้ใช้วิธีการคัดเลือกแบบขั้นตอน (Stepwise) และทดสอบค่าสถิติ F-test ที่ระดับนัยสำคัญทางสถิติ ๐.๐๕

ขั้นตอนการเลือกตัวแปร ได้พิจารณาตามค่าอัตราส่วน F เชิงส่วน (partial multivariate F ratio) และตัวแปรจะได้รับการคัดเลือกเข้าสมการ จำแนกต่อเมื่อค่าของอัตราส่วน F เชิงส่วนมีค่ามากกว่าที่กำหนดไว้ (ค่า F เชิงส่วนนี้วัดความสามารถในการแบ่งแยกที่เกิดจากการนำเอาตัวแปรตัวหนึ่งเข้ามาเมื่อได้พิจารณาถึงความสามารถในการแบ่งแยกที่เกิดจากตัวแปรตัวอื่น ๆ ที่ได้รับการคัดเลือกแล้ว) การทดสอบค่า F เชิงส่วนจะกระทำขึ้นก่อนที่จะนำตัวแปรเข้าสู่ขั้นตอนของสมการวิเคราะห์จำแนก

การวิเคราะห์แบบขั้นตอน ในการคัดเลือกตัวแปร เข้าสมการวิเคราะห์จำแนกประเภท การทดสอบสมมติฐานกระทำดังนี้ คือ

๑. คำนวณหาค่า generalized distance D_p^2

$$\text{โดยที่ } D_p^2 = (\bar{X}_1 - \bar{X}_2)' S_*^{-1} (\bar{X}_1 - \bar{X}_2)$$

๒. คำนวณหาค่า generalized distance D_{p+q}^2

p, q เป็นจำนวนตัวแปร

๓. เปรียบเทียบค่าของ generalized distance ระหว่าง D_p^2 และ D_{p+q}^2

๔. สมมติฐานของการทดสอบ

H_0 : จำนวนตัวแปร q ไม่ทำให้อำนาจในการจำแนกดีขึ้น

H_1 : จำนวนตัวแปร q ทำให้อำนาจในการจำแนกดีขึ้น

๕. คำนวณหาค่า $D_{p+q}^2 - D_p^2$

๖. กำหนดระดับนัยสำคัญ α

๗. ขอบเขตของการปฏิเสธสมมติฐาน

$$F = \frac{n_1+n_2-p-q-1}{q} \frac{m(D_1^2 - D_2^2)}{1+mD_2^2} > F_{\alpha; q, n_1+n_2-p-q-1}$$

$$\text{โดย } m = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-2)}$$

นอกจากนี้ค่า Λ Criterion ซึ่งคิดค้นโดย Wilks (Rao 1973 : 570)

$$\text{หาได้จาก } \Lambda = \frac{|W|}{|W+B|}$$

$$\text{โดย } B_{ij} = \sum_{r=1}^k \frac{T_{ir} T_{jr}}{N_r} - \frac{T_i T_j}{N}$$

$$W_{ij} = \sum_{r=1}^k S_{ij}^{(r)} = S_{ij} - B_{ij}$$

B_{ij} เป็น Sum of products between populations

W_{ij} เป็น Sum of products within populations

k เป็น จำนวนกลุ่ม

N เป็น จำนวนค่าสังเกตทั้งหมด

$$m = t - \frac{p+q+1}{2}, \quad t = N - 1, \quad q = k - 1$$

$$\lambda = \frac{pq - 2}{4}, \quad s = \frac{\sqrt{(p^2 q^2 - 4)}}{\sqrt{p^2 + q^2 - 5}}, \quad r = \frac{pq}{2}$$

และ $-m \log_e \Lambda \sim \chi^2$ มี d.f. เป็น $p(k-1)$

นอกจากนี้ค่า Λ สามารถแปลงให้อยู่ในรูป F โดย

$$F = \frac{ms-2\lambda}{2r} \frac{1-\Lambda^{1/s}}{\Lambda^{1/s}} \quad \text{มี d.f. เป็น } 2r \text{ และ } (ms-2\lambda)$$

จากวิธีการดังกล่าวข้างต้นจะทำให้ได้สมการจำแนกประเภทและค่าสถิติบางตัวที่สำคัญ ซึ่งสามารถใช้ในการตัดสินความสำคัญของสมการวิเคราะห์จำแนกประเภทได้ สถิติตัวนั้นคือ Wilks' Lambda ซึ่ง เป็นมาตรวัดความแตกต่างระหว่างกลุ่มและภายในกลุ่มด้วยค่าของ Wilks' Lambda ค่า หมายถึงความว่าสมการที่ได้สามารถแบ่งแยกกลุ่มได้ดี เราสามารถทดสอบว่าสมการนี้มีอำนาจในการ จำแนกกลุ่มได้อย่างมีนัยสำคัญทางสถิติหรือไม่ โดยแปลงค่าของ Lambda เป็น Chi-square ได้ ดังที่กล่าวมาแล้ว ถ้าค่า χ^2 มากกว่าค่าที่กำหนดจากตารางที่ระดับนัยสำคัญหนึ่งแล้ว สมการนั้น สามารถใช้จำแนกกลุ่มได้อย่างมีนัยสำคัญทางสถิติ

นอกจากนั้นการทดสอบความสำคัญของสมการจำแนกประเภท สามารถทดสอบได้จากค่า eigenvalue โดยสามารถแปลงเป็นค่า Bartlett's V statistic ในการทดสอบสมการ ได้ดังนี้

$$V = - [N-1-(p+k)/2] \log_e \Lambda$$

$$V = [N-1-(p+k)/2] \log_e \frac{1}{\Lambda}$$

$$\text{และ } \frac{1}{\Lambda} = (1+\lambda_1)(1+\lambda_2)\dots\dots(1+\lambda_r)$$

ในกรณีเป็นประชากร ๒ กลุ่ม จะหา nonzero eigenvalue λ_1 ได้เพียงตัวเดียว

ดังนั้น $V = [N-1-(p+k)/2] \log_e (1+\lambda_1)$ โดย λ เป็น nonzero eigenvalue ของ $W^{-1}B$ และ V จะมีการกระจายเป็น chi-square ด้วย degree of freedom $p(k-1)$ โดย p คือ จำนวนตัวแปร, k คือจำนวนกลุ่ม

นอกจากนั้นสมการจำแนกประเภทที่ได้มานั้น มีความสามารถในการแบ่งแยกกลุ่มได้ดีหรือไม่นั้น เราสามารถดูได้จากค่าร้อยละของกรณีที่ทำนายความเป็นสมาชิกกลุ่มได้ถูกต้อง โดยการที่เปรียบเทียบดูว่าสิ่งที่เราศึกษาอยู่นั้นเป็นสมาชิกของประชากรกลุ่มหนึ่ง เมื่อใช้สมการจำแนก ประเภทแล้วสามารถจำแนกได้ถูกต้องกับที่เป็นจริงเท่าไร

ซึ่งการวิเคราะห์จำแนกประเภทในการวิจัยนี้ใช้โปรแกรมสำเร็จรูป SPSS (Statistical Package for Social Science) ในการคัดเลือกตัวแปรและหาค่าสถิติต่าง ๆ

วิธีการวิเคราะห์จำแนกประเภท (Discriminant Analysis) มีสูตรสมการ ดังนี้

$$D_i = d_{i1}Z_1 + d_{i2}Z_2 + \dots + d_{ip}Z_p$$

เมื่อ D = คะแนนจำแนกประเภทในสมการ (Discriminant score)

d = ความสัมพันธ์โดยน้ำหนัก (Weighting coefficients)

Z = ค่าคะแนนมาตรฐานของตัวแปรที่ได้จากการจำแนกประเภท

p = ตัวแปรที่เข้าสมการจำแนกประเภท

ตัวแปรที่จะนำเข้าไปในสมการจำแนกประเภทนั้นบางตัวมีลักษณะเป็นตัวแปรเชิงคุณภาพ เช่น สถานะของบุคคลที่มีความเกี่ยวข้องกับหัวหน้าครัวเรือน เพศ สถานภาพสมรส สถานภาพการมีงานทำ อาชีพ ความต้องการทำงานมากขึ้น ความต้องการฝึกอาชีพ จึงจำเป็นต้องกำหนดตัวแปรตัวนี้มีขึ้นเพื่อช่วยในการวิเคราะห์

ตารางที่ ๒.๑ คะแนนของตัวแปรตัวมีของสถานะของบุคคลที่มีความเกี่ยวข้องกับหัวหน้าครัวเรือน

สถานะของบุคคลที่มีความเกี่ยวข้องกับหัวหน้าครัวเรือน	ตัวแปรตัวมี		
	D1	D2	D3
หัวหน้าครัวเรือน	๑	๐	๐
สามีหรือภรรยา	๐	๑	๐
บุตรที่ยังไม่สมรส	๐	๐	๑

ตารางที่ ๒.๒ คะแนนของตัวแปรคัมมี่ของอาชีพ

กลุ่มอาชีพ	ตัวแปรคัมมี่				
	D6	D7	D8	D9	D10
อาชีพกลุ่มที่ ๑	๑	๐	๐	๐	๐
อาชีพกลุ่มที่ ๒	๐	๑	๐	๐	๐
อาชีพกลุ่มที่ ๓	๐	๐	๑	๐	๐
ว่างงาน	๐	๐	๐	๑	๐
ไม่ได้ทำงาน	๐	๐	๐	๐	๑

นอกจากนั้นกรณีในตัวแปร เป็นตัวแปรเชิงคุณภาพและจัดแบ่ง เป็น ๒ กลุ่ม จะกำหนดตัวแปรคัมมี่ ดังนี้

- เพศ กำหนดตัวแปรคัมมี่ เป็น D_4
 $D_4 = 1$ ถ้า เป็นชาย
 $D_4 = 0$ ถ้า เป็นหญิง
- สถานภาพสมรส กำหนดตัวแปรคัมมี่ เป็น D_5
 $D_5 = 1$ ถ้า เป็นโสด
 $D_5 = 0$ ถ้า สมรส หม้าย หย่า แยกกันอยู่
- ความต้องการทำงานมากขึ้น กำหนดตัวแปรคัมมี่ เป็น D_{11}
 $D_{11} = 1$ ถ้า ต้องการทำงานมากขึ้น
 $D_{11} = 0$ ถ้า ไม่ต้องการทำงานมากขึ้นหรือไม่ทราบ
- ความต้องการฝึกอาชีพ กำหนดตัวแปรคัมมี่ เป็น D_{12}
 $D_{12} = 1$ ถ้า ต้องการฝึกอาชีพ
 $D_{12} = 0$ ถ้า ไม่ต้องการฝึกอาชีพหรือไม่ทราบ

การใช้สมการวิเคราะห์จำแนกนี้ ผู้วิจัยได้แบ่งกลุ่มศึกษาออกเป็น ๒ กลุ่มคือ กลุ่มที่มีความคิดว่าจะไปทำงานในถิ่นอื่น กับกลุ่มที่ไม่คิดจะไป โดยไม่ได้นำเอากลุ่มที่ไม่ทราบว่าจะไปหรือไม่ไปเข้ามารวมในการศึกษา ทั้งนี้เพราะต้องการทราบถึงปัจจัยที่จำแนกประเภทระหว่างกลุ่มคนชนบทที่มีความคิดว่าจะไปทำงานในถิ่นอื่นจริง ๆ และที่ยังไม่มีความคิดว่าจะไปนั้นมีปัจจัยอะไรบ้างเป็นตัวกำหนดให้มีความคิดว่าจะไปทำงานในถิ่นอื่น

การวิเคราะห์โดยสมการจำแนกประเภทจะได้สมการ-ในการทำนาย ๑ สมการ เมื่อแบ่งกลุ่มศึกษาเป็น ๒ กลุ่ม

การตีความหมายของสมการหรือสัมประสิทธิ์ของตัวแปรต่าง ๆ ในสมการ สมการการวิเคราะห์จำแนกประเภทที่จะได้ คือ สมการปรับฐานแล้ว (Standardized) ซึ่งตัวแปรแต่ละตัวจะมีค่าเฉลี่ยเป็นศูนย์ และค่าเบี่ยงเบนมาตรฐานมีค่าเป็น ๑ ในกรณีเช่นนี้ค่าสัมประสิทธิ์ของตัวแปรอิสระแต่ละตัวจะบ่งชี้ถึงความสำคัญของตัวแปรนั้น

แต่ในขณะที่เกี่ยวกับสมการจำแนกที่ได้จะให้สมการที่ไม่ได้ปรับฐาน (Unstandardized) พร้อมทั้งค่าสัมประสิทธิ์ของคะแนนดิบมาด้วย เพื่อให้สะดวกกับการนำไปใช้ประโยชน์ ซึ่งค่าของ discriminant score จะเหมือนกับการใช้สมการปรับฐานแล้ว เมื่อเราปรับค่าด้วยค่าคงที่ (constant)

๒.๓ ในการศึกษาถึงลักษณะทางประชากรบางประการของชาวชนบทที่มีความคิดว่าจะไปทำงานต่างถิ่น

กลุ่มที่มีความคิดจะไปทำงานต่างถิ่นนั้น คนในวัยหนุ่มสาวมีจำนวนมากกว่าคนในวัยอื่น ๆ และเพศชายมีจำนวนมากกว่าเพศหญิง คนที่เป็นโสดก็มีมากกว่าคนที่สมรสแล้ว จะเสนอผลการวิเคราะห์ เป็นคำร้อยละและวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา