ตัวแปรเชิงคุณภาพกับการวิเคราะห์การถดถอย

ชวลิต ทับสีรัก*

บทนำ

การวิเคราะห์การถดถอย เป็นเทคนิคการพยากรณ์ที่นิยมใช้กันอย่างกว้างขวางมากที่สุดในสาขาวิชาต่าง ๆ โดยใช้ตัวแปรอิสระ (Independent Variable) หรือตัวแปรพยากรณ์ (Predictor) อย่างน้อยหนึ่งตัวไปพยากรณ์หรืออธิบายความผันแปรของตัวแปรตาม (Dependent Variable) ถ้าการวิเคราะห์การถดถอยประกอบด้วยตัวแปรอิสระ 1 ตัว สำหรับพยากรณ์ตัวแปรตาม 1 ตัว ซึ่งวัดเป็นค่าในเชิงปริมาณ เรียกว่า การวิเคราะห์การถดถอยอย่างง่าย (Simple Regression) แต่ถ้าใช้ตัวแปรอิสระมากกว่า 1 ตัว พยากรณ์ตัวแปรตามที่วัดในเชิงปริมาณ 1 ตัว เรียกว่า การวิเคราะห์การถดถอยพหุ (Multiple Regression) (ศิริชัย กาญจนวาสี. 2550)

ตัวแปร (Variable) ที่ใช้ในการเขียนสมการทางสังคมศาสตร์ แบ่งออกเป็น 2 ประเภท คือ ตัวแปรเชิงปริมาณ (Quantitative Variable) เป็นตัวแปรที่สามารถวัดค่าได้ เช่น คะแนนจากการสอบนักเรียน น้ำหนัก อายุ เป็นต้น ตัวแปรอีกประเภทหนึ่ง ได้แก่ ตัวแปรเชิงคุณภาพ (Qualitative Variable) เป็นตัวแปรที่ไม่สามารถวัดค่าได้ เช่น เพศ อาชีพ วิธีสอน เป็นต้น โดยที่ตัวแปรเหล่านี้ก็มีอิทธิพลต่อตัวแปรตามเช่นกัน เช่น เพศ มีอิทธิพลต่อค่าใช้จ่ายในการซื้อเสื้อผ้า เนื่องจากเพศหญิงจะมีค่าใช้จ่ายในการซื้อเสื้อผ้าสูงกว่าเพศชาย ดังนั้นในการเขียนสมการพยากรณ์เพื่อพยากรณ์ตัวแปรตาม นอกจากจะให้ความสำคัญกับตัวแปรเชิงปริมาณแล้ว ตัวแปรเชิงคุณภาพก็มีความสำคัญไม่น้อยไปกว่ากัน แต่ในการวิเคราะห์การถดถอยโดยทั่วไปจะเน้นศึกษาเฉพาะตัวแปรเชิงปริมาณ เพราะสามารถวัดค่าออกมาได้ ทำให้ง่ายต่อการนำไปวิเคราะห์ และมองข้ามตัวแปรเชิงคุณภาพซึ่งก็มีอิทธิพลต่อตัวแปรตามเช่นกัน ทำให้ขาดความรู้ความเข้าใจเกี่ยวกับการนำตัวแปรเชิงคุณภาพไปใช้ในการวิเคราะห์การถดถอย แต่ต้องทำความเข้าใจในเบื้องต้นก่อนว่าตัวแปรเชิงคุณภาพที่ใช้ในการวิเคราะห์การถดถอย จะนำไปใช้เฉพาะตัวแปรพยากรณ์เท่านั้น ส่วนตัวแปรตามต้องเป็นตัวแปรที่วัดได้ในเชิงปริมาณ ซึ่งถ้ามีความจำเป็นในการนำตัวแปรเหล่านี้มาใช้ในสมการพยากรณ์ สามารถกระทำได้โดยจัดกระทำตัวแปร

* นิสิตปริญญาเอก สาขาวิจัยและประเมินผลการศึกษา มหาวิทยาลัยมหาสารคาม รุ่นที่ 5

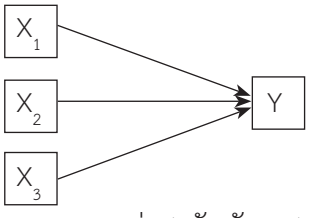
ดังกล่าวให้เป็นตัวแปรจัดประเภท (Categorical Variable) หรือตัวแปรหุ่น (Dummy Variable) บทความเรื่องนี้จะอธิบายถึงวิธีการจัดกระทำตัวแปรเชิงคุณภาพให้เป็นตัวแปรหุ่นเพื่อนำไปวิเคราะห์การถดถอย (ศิริชัย กาญจนวาสี. 2533)

กรอบแนวคิดในการวิเคราะห์การถดถอย อาศัยความสัมพันธ์ระหว่างตัวแปรเป็นหลัก และคำตอบที่ต้องการคือ มีตัวแปรใดบ้างที่สามารถพยากรณ์ตัวแปรตามที่เราสนใจศึกษา และตัวแปรใดพยากรณ์ได้มากน้อยกว่ากัน รวมทั้งส่งผลในทางบวกหรือทางลบ ซึ่งการวิจัยในลักษณะนี้จะต้องอาศัยการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง (Review Literature) มาเป็นอย่างดี และสรุปเป็นกรอบแนวคิดในการวิจัย และนำไปสร้างเครื่องมือตามกรอบแนวคิดที่กำหนดไว้ สามารถเขียนเป็นสมการกรอบแนวคิดดังภาพประกอบ 1 (Kerlinger, 1973)

X 1 ตัว Y 1 ตัว

การวิเคราะห์การถดถอยอย่างง่าย

Simple Regression Analysis



X มากกว่า 1 ตัว กับ Y 1 ตัว

การวิเคราะห์การถดถอยพหุคูณ

Multiple Regression Analysis

ภาพประกอบ 1 กรอบแนวคิดการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรพยากรณ์

ตัวแปรที่วัดในมาตรานามบัญญัติ (Nominal Scale) หรือตัวแปรจัดประเภท เป็นตัวแปรเชิงคุณภาพที่แสดงการจำแนกรายการหรือบรรยายลักษณะของสิ่งนั้น เช่น ตัวแปรเพศ ศาสนา กลุ่มการทดลอง ภูมิลำเนา อาชีพ สถานภาพสมรส เป็นต้น ตัวแปรที่จำแนกประเภทได้เพียง 2 ลักษณะ เรียกว่า ตัวแปรทวิภาค (Dichotomous Variable) เช่น เพศ (ชาย, หญิง) ผลการสอบ (ผ่าน, ไม่ผ่าน) การสูบบุหรี่ (สูบ, ไม่สูบ) เป็นต้น สำหรับตัวแปรที่จำแนกประเภทได้มากกว่า 2 ลักษณะ เรียกว่า ตัวแปรพหุวิภาค (Polytomous Variable) เช่น ศาสนา อาชีพ ภูมิลำเนา เป็นต้น

ตัวแปรทวิภาค (Dichotomous Variable) เป็นตัวแปรที่เป็นไปได้เพียง 2 ลักษณะ (g=2) เช่น เพศ (ชาย, หญิง) การไปใช้สิทธิ์เลือกตั้ง (ไป, ไม่ไป) การตอบข้อสอบ (ถูก, ผิด) เป็นต้น โดยจะขอยกตัวอย่างการจัดกระทำตัวแปรเพศ (Sex) ให้เป็นตัวแปรหุ่น โดยใช้วิธีการ Dummy Coding ซึ่งจะต้องกำหนดรหัส (Codes) ของตัวแปรเพศ ให้เป็นตัวแปรใหม่ X_1 และ X_2 จากนั้นเลือกกลุ่มใดกลุ่มหนึ่งเป็นกลุ่มอ้างอิง (Reference Group) สำหรับใช้เป็นกลุ่มเปรียบเทียบและกำหนดค่าให้เป็น 0 ส่วนอีกกลุ่มหนึ่งกำหนดค่าให้เป็น 1 ดังแสดงในตาราง 1

ตาราง 1 การกำหนดค่าของตัวแปรทวิภาค เพศ

เพศ	X_1	X_2
หญิง	1	0
ชาย	0	1

จากตาราง 1 ตัวแปร X_1 แสดงถึง “หญิง (1)” กับ “ชาย (0)” ซึ่งอาจเรียกตัวแปรใหม่นี้ว่า “ความเป็นหญิง” ไม่จำเป็นที่จะต้องมีตัวแปร X_2 (ความเป็นชาย) เนื่องจาก X_2 มีลักษณะที่ตรงข้ามกับ X_1 และ $r_{X_1X_2}=-1$ นอกจากนี้ความสัมพันธ์ระหว่าง X_1 หรือ X_2 กับ Y จะมีค่าเท่ากัน เพียงแต่เครื่องหมายตรงข้าม $R_{YX_1}=-R_{YX_2}$ แสดงว่า X_2 ซ้ำซ้อนกับ X_1 อย่างสมบูรณ์ ดังนั้นตัวแปรเพศจึง Code ให้เป็นเพียงตัวแปร X_1 หรือ X_2 เพียงตัวใดตัวหนึ่งก็มีความหมายเพียงพอสำหรับการคำนวณอิทธิพลของเพศ ดังตาราง 2

ตาราง 2 การกำหนดค่าของตัวแปรทวิภาค เพศ โดยวิธีการ Dummy Coding

เพศ	X_1
หญิง	1
ชาย	0

สำหรับตัวแปรพหุวิภาค (Polytomous Variable) เป็นตัวแปรที่มีจำนวนรายการมากกว่า 2 ลักษณะ เช่น ศาสนา (คริสต์, อิสลาม, พุทธและฮินดู) ภูมิลำเนา กลุ่มการทดลอง ประเภทของสินค้า เป็นต้น ยกตัวอย่างตัวแปรศาสนา เมื่อนำมาใช้ในสมการพยากรณ์จะไม่สามารถกำหนดค่าของตัวแปรศาสนาให้เป็นตัวแปรเดียวเหมือนอย่างตัวแปรต่อเนื่องอื่น ๆ เช่น IQ, น้ำหนัก, รายได้, อายุ เป็นต้น เนื่องจากตัวแปรตัวเดียวคงไม่สามารถใช้เป็นตัวแทนที่เพียงพอของทุกศาสนาได้ ถ้าตัวแปรศาสนาจำแนกเป็น g รายการ เราสามารถสร้างตัวแปรอิสระที่แตกต่างกันขึ้นมาได้ g-1 ตัวแปร ตัวอย่างเช่น ตัวแปรศาสนาถูกจำแนกเป็น 4 รายการ (g=4) ได้แก่ คริสต์, อิสลาม, พุทธ และฮินดู เราสามารถสร้างตัวแปร Dummy ใหม่ได้ 3 ตัวแปร (X_1, X_2 และ X_3) ที่ใช้เป็นตัวแทนของตัวแปรศาสนาได้สารสนเทศอย่างสมบูรณ์ โดยที่ไม่ต้องใช้ตัวแปร 4 ตัว (X_1, X_2, X_3 และ X_4) ซึ่งจะทำให้เกิดความซ้ำซ้อนและไม่ได้สารสนเทศอะไรเพิ่มเติมและทำให้การคำนวณมีปัญหาและผลการวิเคราะห์ไม่สามารถให้ค่าเฉพาะที่คงที่ได้ ดังนั้นตัวแปรที่ 4 ในกรณีนี้จึงไม่จำเป็น

แนวคิดของ Dummy Coding เป็นการจัดกระทำตัวแปรมาตรานามบัญญัติ ซึ่งมีสมาชิกอยู่ g กลุ่ม ให้เป็นตัวแปรทวิวิภาค (Dichotomous) ชุดใหม่ จำนวน g-1 ตัว โดยมีการเลือกกลุ่มใดกลุ่มหนึ่งเป็นกลุ่มอ้างอิง (Reference Group) สำหรับใช้เป็นกลุ่มเปรียบเทียบ โดยกลุ่มอ้างอิงจะถูกกำหนดค่าให้เป็น 0 สำหรับกลุ่มอื่น ๆ ที่ต้องการนำมาเปรียบเทียบกำหนดค่าให้เป็น 1 ดังตาราง 3

ตาราง 3 Dummy Variable Coding ของตัวแปรมาตรานามบัญญัติที่มี g=4 กลุ่ม

กลุ่ม	ศาสนา	ภาค	กลุ่มการทดลอง	X_1	X_2	X_3
G1	คริสต์	เหนือ	กลุ่มการทดลอง 1	1	0	0
G2	อิสลาม	กลาง	กลุ่มการทดลอง 2	0	1	0
G3	พุทธ	ใต้	กลุ่มการทดลอง 3	0	0	1
(Reference) G4	ฮินดู	ตะวันออกเฉียงเหนือ	กลุ่มการทดลอง 4	0	0	0

จากตาราง 3 สมมติว่ามีตัวแปรที่วัดอยู่ในมาตรานามบัญญัติ ซึ่งสามารถจัดประเภทหรือกลุ่มได้ 4 ลักษณะ เช่น ตัวแปรศาสนา ตัวแปรภาค ตัวแปรกลุ่มการทดลอง จึงสามารถสร้างตัวแปรหุ่นได้ 3 ตัวแปร ได้แก่ ตัวแปร X_1, X_2, และ X_3 (ฉัตรศิริ ปิยะพิมลสิทธิ์. 2554 : ออนไลน์)

ผู้เขียนจะยกตัวอย่างการใช้ตัวแปรเชิงคุณภาพร่วมกับตัวแปรเชิงปริมาณ ซึ่งเป็นตัวแปรต่อเนื่อง (Continuous Variable) เพื่อใช้ในการวิเคราะห์การถดถอยและพยากรณ์ตัวแปรตาม โดยการสมมติข้อมูลขึ้นมา ซึ่งมีคำถามว่า ทัศนคติต่อการคุมกำเนิด (Y) ของประชากร ขึ้นอยู่กับตัวแปรใดบ้าง จึงได้ทำการศึกษาตัวแปรที่คาดว่าจะส่งผลต่อทัศคติต่อการคุมกำเนินของประชากร พบว่า มีตัวแปรที่เกี่ยวข้อง ได้แก่ ตัวแปรศาสนา (Religion) ซึ่งแบ่งเป็นศาสนาคริสต์ (X_1), ศาสนาอิสลาม (X_2), ศาสนาพุทธ (X_3) และศาสนาฮินดู ตัวแปรอายุ (Age) (X_4) และตัวแปรความฉลาดทางอารมณ์ (E.Q.) (X_5)

ตาราง 4 ตัวอย่างข้อมูลแสดงการใช้ Dummy Coding g สำหรับ g=4 กลุ่ม

คนที่	กลุ่ม	Y	X_1	X_2	X_3	X_4	X_5	คนที่	กลุ่ม	Y	X_1	X_2	X_3	X_4	X_5
1	G2	61	0	1	0	25	75	21	G1	56	1	0	0	25	67
2	G4	78	0	0	0	26	79	22	G3	135	0	0	1	48	83
3	G1	47	1	0	0	20	70	23	G1	64	1	0	0	30	76
4	G2	65	0	1	0	23	74	24	G1	130	1	0	0	46	81
5	G2	45	0	1	0	21	69	25	G3	74	0	0	1	23	75
6	G4	106	0	0	0	40	80	26	G4	58	0	0	0	22	74
7	G1	120	1	0	0	45	85	27	G1	116	1	0	0	45	87
8	G2	49	0	1	0	24	82	28	G4	60	0	0	0	25	69
9	G4	45	0	0	0	26	72	29	G3	84	0	0	1	29	70
10	G4	62	0	0	0	25	71	30	G1	68	1	0	0	36	77
11	G2	79	0	1	0	30	78	31	G1	90	1	0	0	38	75
12	G4	54	0	0	0	23	68	32	G1	112	1	0	0	42	86
13	G1	140	1	0	0	50	88	33	G3	94	0	0	1	37	79
14	G2	52	0	1	0	25	71	34	G1	80	1	0	0	30	80
15	G1	88	1	0	0	29	76	35	G3	110	0	0	1	38	78
16	G2	70	0	1	0	39	80	36	G1	102	1	0	0	37	75
17	G2	56	0	1	0	30	80								
18	G3	124	0	0	1	46	79								
19	G4	98	0	0	0	49	74								
20	G2	69	0	1	0	40	79								

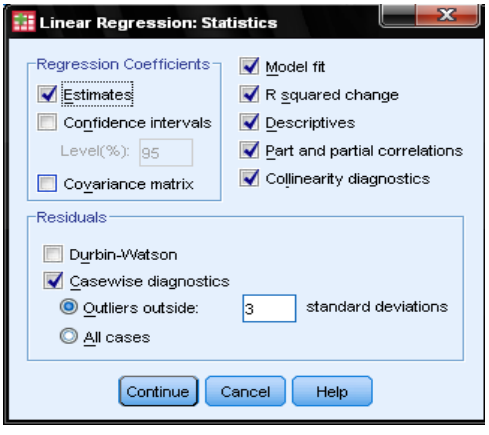
จากตาราง 4 มีกลุ่มตัวอย่าง 36 คน โดยที่แต่ละคนนับถือศาสนาที่แตกต่างกัน มีอายุต่างกัน ใช้เครื่องมือวัดความฉลาดทางอารมณ์และวัดคะแนนทัศนคติต่อการคุมกำเนิดในแต่ละคน จากนั้นนำข้อมูลไปวิเคราะห์การถดถอยพหุคูณ ซึ่งมีขั้นตอนการวิเคราะห์ด้วยโปรแกรม SPSS ดังนี้ (ทรงศักดิ์ ภูสีอ่อน. 2554)

1. Click Analyze ⟶ Regression ⟶ Linear

2. เลือกตัวแปรตามเข้าไปในช่อง Dependent ซึ่งในตัวอย่างคือ ตัวแปรคะแนนทัศนคติ (Y) และเลือกตัวแปรพยากรณ์ เข้าไปในช่อง Independent ซึ่งได้แก่ ศาสนาคริสต์ (X_1), ศาสนาอิสลาม (X_2), ศาสนาพุทธ (X_3), อายุ (X_4), และความฉลาดทางอารมณ์ (X_5)

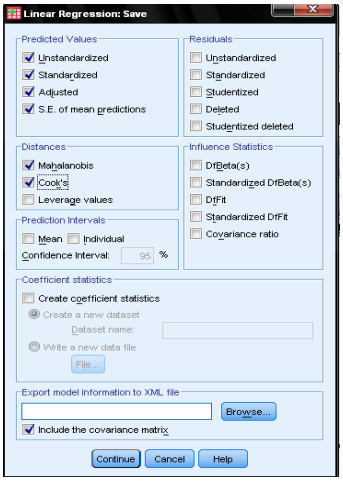
3. ในส่วนของ Method เลือกวิธี Enter

4. ในส่วนของ Statistics ให้เลือกผลการวิเคราะห์ที่มีความสำคัญ ดังภาพ



จากภาพเลือก Model fit ซึ่งจะแสดงผลการวิเคราะห์ข้อมูลเกี่ยวกับความสัมพันธ์ระหว่างตัวแปรตามกับชุดของตัวแปรพยากรณ์ หรืออาจกล่าวได้ว่า ตัวแปรพยากรณ์สามารถทำนายตัวแปรตามได้อย่างมีนัยสำคัญทางสถิติหรือไม่ Descriptive จะแสดงค่าสถิติพื้นฐาน R squared change ซึ่งเป็นค่าที่ระบุว่าความแปรปรวนของตัวแปรตามสามารถอธิบายได้จากชุดของตัวแปรพยากรณ์ได้ร้อยละเท่าใด Part and partial correlation ให้ค่าสัมประสิทธิ์สหสัมพันธ์ สัมประสิทธิ์สหสัมพันธ์บางตอนและสัมประสิทธิ์สหสัมพันธ์ส่วนย่อย โดยจะปรากฏในตาราง Model Summary

5. ในส่วนของ Save ให้เลือกผลการวิเคราะห์ที่มีความสำคัญ ดังภาพ



จากภาพ หากเราต้องการให้โปรแกรมคำนวณค่าต่าง ๆ ที่ต้องการก็สามารถเลือกได้ เช่น ค่าของตัวแปรตามที่ได้จากการทำนายในรูปคะแนนดิบ (Unstandardized) คะแนนมาตรฐาน (Standardized) คะแนนที่ปรับแก้ (Adjusted) หรือค่าความคลาดเคลื่อนมาตรฐานของการพยากรณ์ (S.E. of mean Predictions) ซึ่งผลการวิเคราะห์เหล่านี้จะไม่ปรากฏใน out put แต่จะสร้างตัวแปรดังกล่าวใน Data Editor ดังนั้นจึงขึ้นอยู่กับว่าผู้วิจัยว่าต้องการนำผลการวิเคราะห์ไปใช้หรือไม่

6. Click Continue และ Click OK จะได้ผลการวิเคราะห์ตามที่ต้องการ ซึ่งมีผลการวิเคราะห์ดังนี้

REGRESSION

ตัวแปร	Mean	Std.Deviation	N
ทัศนคติ	81.69	27.880	36
ศาสนาคริสต์	.36	.487	36
ศาสนาอิสลาม	.25	.439	36
ศาสนาพุทธ	.17	.378	36
อายุ	32.97	9.263	36
ความฉลาดทางอารมณ์	76.72	5.460	36

ผลการวิเคราะห์ในส่วนนี้จะให้ค่าสถิติพื้นฐานของตัวแปรตามและตัวแปรพยากรณ์ซึ่งเป็นตัวแปรในมาตรานามบัญญัติที่จัดกระทำให้เป็นตัวแปรหุ่น คือ ศาสนาคริสต์ อิสลามและพุทธและเป็นตัวแปรต่อเนื่อง คือ ทัศนคติ อายุและความฉลาดทางอารมณ์

ผลการวิเคราะห์ส่วนนี้จะให้ค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย ระหว่างตัวแปรตามกับตัวแปรพยากรณ์ ซึ่งเป็นตัวแปรหุ่นและตัวแปรต่อเนื่อง และระหว่างตัวแปรพยากรณ์ด้วยกันเอง จะเห็นว่า ตัวแปรพยากรณ์ทุกตัวสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญที่ระดับ .05 ซึ่งมีทั้งสัมพันธ์ทางบวกและทางลบ และเมื่อพิจารณาความสัมพันธ์ระหว่างตัวแปรพยากรณ์ พบว่า มีทั้งในส่วนที่สัมพันธ์กันและไม่สัมพันธ์กัน แต่ไม่มีตัวแปรพยากรณ์คู่ใดที่มีความสัมพันธ์กันสูงเกินไป (เกิน .80) ซึ่งจะเป็นการฝ่าฝืนข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอย

Correlations

		ทัศนคติ	ศาสนาคริสต์	ศาสนาอิสลาม	ศาสนาพุทธ	อายุ	ความฉลาดทางอารมณ์
Pearson Correlation	ทัศนคติ	1.000	.318	-.442	.355	.880	.701
	ศาสนาคริสต์	.318	1.000	-.434	-.336	.281	.275
	ศาสนาอิสลาม	-.442	-.434	1.000	-.258	-.279	-.030
	ศาสนาพุทธ	.355	-.336	-.258	1.000	.189	.051
	อายุ	.880	.281	-.279	.189	1.000	.711
	ความฉลาดทางอารมณ์	.701	.275	-.030	.051	.711	1.000
Sig. (1-tailed)	ทัศนคติ	.	.030	.004	.017	.000	.000
	ศาสนาคริสต์	.030	.	.004	.022	.049	.052
	ศาสนาอิสลาม	.004	.004	.	.064	.050	.432
	ศาสนาพุทธ	.017	.022	.064	.	.135	.384
	อายุ	.000	.049	.050	.135	.	.000
	ความฉลาดทางอารมณ์	.000	.052	.432	.384	.000	.
N	ทัศนคติ	36	36	36	36	36	36
	ศาสนาคริสต์	36	36	36	36	36	36
	ศาสนาอิสลาม	36	36	36	36	36	36
	ศาสนาพุทธ	36	36	36	36	36	36
	อายุ	36	36	36	36	36	36
	ความฉลาดทางอารมณ์	36	36	36	36	36	36

ผลการวิเคราะห์ในส่วนนี้จะให้ค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ (R) คือความสัมพันธ์ระหว่างตัวแปรตามกับชุดตัวแปรพยากรณ์ ซึ่งมีค่า .934 ซึ่งมีความสัมพันธ์กันสูง ส่วนค่า R_2 เป็นค่าที่ระบุความแปรปรวนของตัวแปรตามที่สามารถอธิบายได้จากชุดของตัวแปรพยากรณ์ทั้งหมด ซึ่งหาได้จากการนำค่า R_2 คูณด้วย 100 ซึ่งจากตารางจะได้ค่า R_2 เท่ากับ 87.2 หมายความว่า ตัวแปรพยากรณ์สามารถร่วมกันอธิบายความแปรปรวนของตัวแปรตามได้ร้อยละ 87.2

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.934[a]	.872	.851	10.756	.872	41.029	5	30	.000

a. Predictors: (Constant), ความฉลาดทางอารมณ์, ศาสนาอิสลาม, ศาสนาพุทธ, ศาสนาคริสต์, อายุ

b. Dependent Variable: ทัศนคติ

ผลการวิเคราะห์ในส่วนนี้จะทดสอบเกี่ยวกับความสัมพันธ์ระหว่างตัวแปรตามกับชุดของตัวแปรพยากรณ์ โดยใช้สถิติ F ทดสอบ ซึ่งจากตารางพบว่ามีนัยสำคัญทางสถิติ (ค่า Sig น้อยกว่า .05) หมายความว่า ตัวแปรตามสามารถพยากรณ์ด้วยชุดของตัวแปรพยากรณ์ได้อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นั่นเอง

วัตถุประสงค์หลักของการวิเคราะห์การถดถอย คือ การสร้างสมการพยากรณ์ ซึ่งการแปลความหมายของสมการพยากรณ์ที่มีตัวแปรหุ่นร่วมด้วยในส่วนนี้จะแตกต่างจากการแปลความหมายของตัวแปรในเชิงปริมาณหรือตัวแปรต่อเนื่องทั่ว ๆ ไป ดังนี้

สมการพยากรณ์คะแนนทัศนคติต่อการคุมกำเนิด (Y) ในรูปคะแนนมาตรฐาน เมื่อตัวแปรพยากรณ์จัดอยู่ในมาตรานามบัญญัติ (ตัวแปรเชิงคุณภาพ) ซึ่งจัดกระทำให้เป็นตัวแปรหุ่นและมีตัวแปรต่อเนื่องร่วมด้วย

$$\hat{Z} = B_1Z_{X1} + B_2Z_{X2} + B_3Z_{X4} + B_5Z_{X5}$$

เมื่อ $\hat{Z}$ = ค่าของตัวแปรตามที่ได้จากการพยากรณ์ในรูปคะแนนมาตรฐาน

B = ค่าสัมประสิทธิ์การถดถอยส่วนย่อยมาตรฐาน

จากตาราง Coefficients เขียนสมการพยากรณ์ที่อยู่ในรูปคะแนนมาตรฐาน ได้ดังนี้

$$\hat{Z}_Y = .075Z_{X1} - .181Z_{X2} + .208Z_{X3} + .601Z_{X4} + .238Z_{X5}$$

จากสมการจะเห็นว่า ตัวแปรที่สามารถใช้พยากรณ์ทัศนคติต่อการคุมกำเนิดของประชากรได้ดีที่สุด คือ อายุ ซึ่งมีค่าสัมประสิทธิ์การถดถอยส่วนย่อยเท่ากับ .601 รองลงมา คือ ความฉลาดทางอารมณ์

สมการพยากรณ์คะแนนทัศนคติต่อการคุมกำเนิด (Y) ในรูปคะแนนดิบ เมื่อตัวแปรพยากรณ์จัดอยู่ในมาตรานามบัญญัติ (ตัวแปรเชิงคุณภาพ) ซึ่งจัดกระทำให้เป็นตัวแปรหุ่นและมีตัวแปรต่อเนื่องร่วมด้วย

$$\hat{Y} = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5$$

เมื่อ $\hat{Y}$ = ค่าของตัวแปรตามที่ได้จากการพยากรณ์ในรูปคะแนนดิบ

X = ค่าสังเกตของตัวแปรพยากรณ์

a = ค่าคงที่หรือจุดตัดแกน Y

b = ค่าสัมประสิทธิ์การถดถอยหรือค่าเฉลี่ยการเปลี่ยนแปลงใน Y เมื่อ X เปลี่ยนแปลงไป 1 หน่วย

จากตาราง Coefficients เขียนสมการพยากรณ์ที่อยู่ในรูปคะแนนดิบ ได้ดังนี้

$$\hat{Y} = -72.22 + 4.28X_1 - 11.47X_2 + 15.31X_3 + 1.80X_4 + 1.21X_5$$

จากสมการพยากรณ์ แทน ศาสนาคริสต์ (X_1) = 1, ศาสนาอิสลาม (X_2) = 0 ศาสนาพุทธ (X_3) = 0 ส่วนค่าของตัวแปรอายุ (X_4) และตัวแปรความฉลาดทางอารมณ์ (X_5) ไม่ต้องแทนค่าโดยให้ถือว่าเป็นค่าคงที่

จะได้ $\hat{Y} = -72.22 + 4.28(1) - 11.47(0) + 15.31(0) + 1.80X_4 + 1.21X_5$

$\hat{Y}_1 = -67.94 + 1.80X_4 + 1.21X_5$

จากสมการพยากรณ์ แทน ศาสนาคริสต์ (X_1) = 0, ศาสนาอิสลาม (X_2) = 1 ศาสนาพุทธ (X_3) = 0 ส่วนค่าของตัวแปรอายุ (X_4) และตัวแปรความฉลาดทางอารมณ์ (X_5) ไม่ต้องแทนค่าโดยให้ถือว่าเป็นค่าคงที่

จะได้ $\hat{Y} = -72.22 + 4.28(1) - 11.47(0) + 15.31(0) + 1.80X_4 + 1.21X_5$

$\hat{Y}_2 = -83.69 + 1.80X_4 + 1.21X_5$

จากสมการพยากรณ์ แทน ศาสนาคริสต์ (X_1) = 0, ศาสนาอิสลาม (X_2) = 0 ศาสนาพุทธ (X_3) = 1 ส่วนค่าของตัวแปรอายุ (X_4) และตัวแปรความฉลาดทางอารมณ์ (X_5) ไม่ต้องแทนค่าโดยให้ถือว่ามีค่าคงที่

จะได้ $\hat{Y} = -72.22 + 4.28(1) - 11.47(0) + 15.31(0) + 1.80X_4 + 1.21X_5$

$\hat{Y}_3 = -56.91 + 1.80X_4 + 1.21X_5$

จากสมการพยากรณ์ แทน ศาสนาคริสต์ (X_1) = 0, ศาสนาอิสลาม (X_2) = 0 ศาสนาพุทธ (X_3) = 0 ส่วนค่าของตัวแปรอายุ (X_4) และตัวแปรความฉลาดทางอารมณ์ (X_5) ไม่ต้องแทนค่าโดยให้ถือว่าเป็นค่าคงที่

จะได้ $\hat{Y} = -72.22 + 4.28(1) - 11.47(0) + 15.31(0) + 1.80X_4 + 1.21X_5$

$\hat{Y}_4 = -72.22 + 1.80X_4 + 1.21X_5$

การแปลผล

เนื่องจากตัวแปรพยากรณ์ในที่นี้มีทั้งที่เป็นตัวแปรหุ่น (Dummy Variable) คือตัวแปรศาสนา (X_1, X_2, X_3) ซึ่งแบ่งออกเป็น 4 ลักษณะ การเปลี่ยนแปลง 1 หน่วยของตัวแปรศาสนาเป็นการเปลี่ยนแปลงจาก 0 (ศาสนาฮินดู) เป็น 1 (ศาสนาคริสต์, ศาสนาอิสลามและศาสนาพุทธ) หรือเปลี่ยนจากระดับหนึ่งไปสู่อีกระดับหนึ่ง เช่น $X_1 = 1$ เป็นการเปลี่ยนแปลงจาก ศาสนาฮินดูเป็นศาสนาคริสต์ ส่วนตัวแปรพยากรณ์ที่เป็นตัวแปรต่อเนื่องอื่น ๆ ให้แทนค่าไปตามค่าของตัวแปรนั้น ๆ แต่ในตัวอย่างนี้จะให้ค่าของตัวแปรอายุและตัวแปรความฉลาดทางอารมณ์มีค่าคงที่ (อาจจะแทนค่าตัวแปรอายุและตัวแปรความฉลาดทางอารมณ์ลงในสมการก็ได้ แต่ตัวแปรแต่ละตัวต้องมีค่าเท่ากันทุก ๆ สมการพยากรณ์) เพื่อเป็นการเปรียบเทียบลักษณะของตัวแปรหุ่น จากผลการแทนค่า จะได้ค่าต่าง ๆ ดังนี้ (บุญเรียง ขจรศิลป์, 2538, 2533)

$\hat{Y}_1 = -67.94 + 1.80X_4 + 1.21X_5$ หมายถึง ค่าเฉลี่ยของคะแนนทัศนคติต่อการคุมกำเนิด (Y) สำหรับความเป็นศาสนาคริสต์ มีค่าสูงกว่าศาสนาฮินดู ซึ่งเป็นกลุ่มอ้างอิง ($\hat{Y}_4 = -72.22 + 1.80X_4 + 1.21X_5$) อยู่ $\hat{Y}_1 - \hat{Y}_4$ เท่ากับ 4.28 โดยค่าของตัวแปรอายุและ ตัวแปรความฉลาดทางอารมณ์มีค่าคงที่ นั่นแสดงว่า ถ้ากลุ่มประชากรที่ต้องการอ้างอิงเป็นกลุ่ม ที่นับถือศาสนาคริสต์และศาสนาฮินดู และมีค่าเฉลี่ยของอายุและความฉลาดทางอารมณ์เท่ากันทั้งสองกลุ่ม ประชากรที่นับถือศาสนาคริสต์จะมีค่าเฉลี่ยของคะแนนทัศนคติต่อการคุมกำเนิดสูงกว่าประชากรที่นับถือศาสนาฮินดูอยู่ 4.28 คะแนน

$\hat{Y}_2 = -83.69 + 1.80X_4 + 1.21X_5$ หมายถึง ค่าเฉลี่ยของคะแนนทัศนคติต่อการคุมกำเนิด (Y) สำหรับความเป็นศาสนาอิสลาม มีค่าต่ำกว่าศาสนาฮินดู ซึ่งเป็นกลุ่มอ้างอิง ($\hat{Y}_4 = -72.22 + 1.80X_4 + 1.21X_5$) อยู่ $\hat{Y}_2 - \hat{Y}_4$ เท่ากับ 11.47 โดยค่าของตัวแปรอายุและตัวแปรความฉลาดทางอารมณ์มีค่าคงที่ นั่นแสดงว่า ถ้ากลุ่มประชากรที่ต้องการอ้างอิงเป็นกลุ่มที่นับถือศาสนาอิสลามและศาสนาฮินดูและมีค่าเฉลี่ยของอายุและความฉลาดทางอารมณ์เท่ากันทั้งสองกลุ่ม ประชากรที่นับถือศาสนาอิสลามจะมีค่าเฉลี่ยของคะแนนคะแนนทัศนคติต่อการคุมกำเนิดต่ำกว่าประชากรที่นับถือศาสนาฮินดูอยู่ 11.47 คะแนน

$\hat{Y}_3 = -56.91 + 1.80X_4 + 1.21X_5$ หมายถึง ค่าเฉลี่ยของคะแนนทัศนคติต่อการคุมกำเนิด (Y) สำหรับความเป็นศาสนาพุทธมีค่าสูงกว่าศาสนาฮินดู ซึ่งเป็นกลุ่มอ้างอิง ($\hat{Y}_4 = -72.22 + 1.80X_4 + 1.21X_5$) อยู่ $\hat{Y}_3 - \hat{Y}_4$ เท่ากับ 15.31 โดยค่าของตัวแปรอายุและตัวแปรความฉลาดทางอารมณ์มีค่าคงที่ นั่นแสดงว่า ถ้ากลุ่มประชากรที่ต้องการอ้างอิงเป็นกลุ่มที่นับถือศาสนาพุทธและศาสนาฮินดูและมีค่าเฉลี่ยของอายุและความฉลาดทางอารมณ์เท่ากันทั้งสองกลุ่ม ประชากรที่นับถือศาสนาพุทธจะมีค่าเฉลี่ยของคะแนนทัศนคติต่อการคุมกำเนิดสูงกว่าประชากรที่นับถือศาสนาฮินดูอยู่ 15.31 คะแนน

$\hat{Y}_4 = -72.22 + 1.80X_4 + 1.21X_5$ หมายถึง ค่าเฉลี่ยของคะแนนทัศนคติต่อการคุมกำเนิด (Y) สำหรับความเป็นศาสนาฮินดู ซึ่งมีค่าเท่ากับค่าคงที่ของสมการพยากรณ์ โดยค่าของตัวแปรอายุและตัวแปรความฉลาดทางอารมณ์มีค่าคงที่

สรุป

ในการวิเคราะห์การถดถอย ในบางครั้งตัวแปรที่สำคัญที่มีส่วนในการพยากรณ์ตัวแปรตามที่ได้จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง รวมทั้งทฤษฎีต่าง ๆ นั้น นอกจากจะเป็นตัวแปรเชิงปริมาณซึ่งเป็นตัวแปรต่อเนื่องแล้ว ตัวแปรเชิงคุณภาพก็มีบทบาทในการพยากรณ์และมีส่วนสำคัญเช่นเดียวกัน ซึ่งตัวแปรเหล่านี้ไม่สามารถวัดค่าออกมาได้ว่ามากหรือน้อย ดังนั้นก่อนที่เราจะนำตัวแปรเชิงคุณภาพไปใช้ในการวิเคราะห์การถดถอย จำเป็นต้องกำหนดค่าที่แตกต่างกันให้แต่ละกลุ่มที่แตกต่างกันของตัวแปรเชิงคุณภาพที่เป็นตัวแปรแบบแบ่งประเภทโดยใช้แนวคิดของ Dummy Coding เป็นการจัดกระทำตัวแปรมาตรานามบัญญัติ ซึ่งมีสมาชิกอยู่ g กลุ่ม ให้เป็นตัวแปรทวิวิภาค (Dichotomous) ชุดใหม่ จำนวน g-1 ตัว โดยมีการเลือกกลุ่มใดกลุ่มหนึ่งเป็นกลุ่มอ้างอิง (Reference Group) นอกจากนี้การแปลความหมายจากสมการที่ได้นั้นแตกต่างจากสมการที่มีเฉพาะตัวแปรเชิงปริมาณทั่วไป ซึ่งศึกษาได้จากรายละเอียดที่ผ่านมา

บรรณานุกรม

ฉัตรศิริ ปิยะพิมลสิทธิ์. **ตัวแปรดัมมี่ (ออนไลน์)**. 30 กันยายน 2554. แหล่งที่มา http://www.watpon.com/www.

ทรงศักดิ์ ภูสีอ่อน. **การประยุกต์ใช้ SPSS วิเคราะห์ข้อมูลงานวิจัย**. พิมพ์ครั้งที่ 4. มหาสารคาม : โรงพิมพ์มหาวิทยาลัยมหาสารคาม, 2554.

บุญเรียง ขจรศิลป์. **การวิเคราะห์ข้อมูลในการวิจัยโดยใช้โปรแกรมสำเร็จรูป SPSSx**. กรุงเทพฯ : ฟิสิกส์เซ็นเตอร์การพิมพ์, 2533.

_____. **การวิเคราะห์ข้อมูลในการวิจัยโดยใช้โปรแกรมสำเร็จรูป SPSSx/PC+**. พิมพ์ครั้งที่ 2. กรุงเทพฯ : โครงการงานบริการจัดพิมพ์เอกสารวิชาการและตำรา, 2538.

ศิริชัย กาญจนวาสี. **การวิเคราะห์พหุระดับ**. พิมพ์ครั้งที่ 4. กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2550.

_____. “ตัวแปรสำหรับการวิจัย : การคัดเลือก การวัดและการควบคุม” **วารสาร วิธีวิทยาการวิจัย**. ปีที่ 5 ฉบับที่ 3 กันยายน–ธันวาคม 2533.

Kerlinger, Fred N. and Pedhazur, Elazar J. **Multiple Regression in Behavioral Research**. New York : Holt, Rinehart and Winaton, Inc., 1973.