บทที่ 2

วรรณคดีที่เกี่ยวข้อง

ผู้วิจัยนำเสนอผลการศึกษาค้นคว้าเอกสารและงานวิจัยเกี่ยวข้องในบทนี้โดยแบ่งออกเป็น 3 ตอนดังนี้ ตอนแรก การวิเคราะห์การถดถอยโลจิสติก ในตอนนี้ผู้วิจัยนำเสนอสังกัปของการวิเคราะห์การถดถอยโลจิสติก ข้อตกลงเบื้องต้น โมเดลการวิเคราะห์การถดถอยโลจิสติก การตรวจสอบความตรงของโมเดล และการคัดเลือกตัวแปรพยากรณ์ ตอนที่ 2 การวิเคราะห์อิทธิพล ในตอนนี้ผู้วิจัยนำเสนอสังกัปเบื้องต้นของการวิเคราะห์อิทธิพล ความสัมพันธ์ และความเป็นสาเหตุระหว่างตัวแปร การพัฒนาโมเดลแสดงอิทธิพล โมเดลลิสเรลสำหรับการวิเคราะห์อิทธิพล และหลักการวิเคราะห์โมเดลความสัมพันธ์เชิงสาเหตุในการวิเคราะห์อิทธิพล และ ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการศึกษาวิจัยเรื่องตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนและงานวิจัยเกี่ยวกับการประยุกต์ใช้การวิเคราะห์การถดถอยโลจิสติกร่วมกับการวิเคราะห์อิทธิพล ในตอนนี้ผู้วิจัยนำเสนอผลการสังเคราะห์ผลงานวิจัยที่เกี่ยวข้องกับการวิจัยครั้งนี้เพื่อประมวลตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียน และพัฒนาโมเดลความสัมพันธ์เชิงสาเหตุของความคาดหวังในการศึกษาต่อของนักเรียนเป็นสมมุติฐานวิจัย รวมทั้งเสนองานวิจัยที่เกี่ยวข้องกับการประยุกต์ใช้การวิเคราะห์การถดถอยโลจิสติกร่วมกับการวิเคราะห์อิทธิพล รายละเอียดเนื้อหาสาระแต่ละตอนมีดังต่อไปนี้

ตอนที่ 1 การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression Analysis)

การวิเคราะห์การถดถอยโลจิสติก เป็นเทคนิคการวิเคราะห์ข้อมูลทางสถิติวิธีหนึ่งที่ใช้สำหรับการพยากรณ์ค่าความน่าจะเป็นของการเกิดเหตุการณ์ซึ่งเป็นตัวแปรผลจากชุดของตัวแปรอิสระ โมเดลการถดถอยโลจิสติกเป็นที่รู้จักกันอย่างดีในหมู่นักสถิติ การประยุกต์โมเดลนี้ในการวิจัย พบในการวิจัยทางการแพทย์และสาธารณสุขก่อน และได้แพร่หลายเข้าสู่การวิจัยทางพฤติกรรมศาสตร์และสังคมศาสตร์ในรอบทศวรรษที่ผ่านมา (นงลักษณ์ วิรัชชัย, 2538; Hosmer and Lemmeshow, 1989)

สังกัปของการวิเคราะห์การถดถอยโลจิสติก

การวิเคราะห์การถดถอยโลจิสติก เป็นเทคนิคการวิเคราะห์ที่มีหลักการเหมือนกับการวิเคราะห์การถดถอยพหุคูณเชิงเส้นตรง (multiple linear regression analysis) แต่มีความเหมาะสมที่จะใช้กับข้อมูลลักษณะแตกต่างกัน การวิเคราะห์การถดถอยพหุคูณเชิงเส้นตรงใช้วิเคราะห์หารูปแบบความสัมพันธ์เชิงเส้นตรง ระหว่างตัวแปรอิสระหลายตัวกับตัวแปรตามตัวเดียว โดยที่ตัวแปรทั้งหมดมีระดับการวัดแบบมาตรอันตรภาคหรือมาตรอัตราส่วน (interval or ratio scale) ในกรณีที่ตัวแปรอิสระมีระดับการวัดแบบมาตรนามบัญญัติหรือมาตรเรียงอันดับ (nominal or ordinal scale) อาจแปลงให้เป็นตัวแปรดัมมี่ (dummy variable) ก่อน ส่วนการวิเคราะห์การถดถอยโลจิสติกใช้วิเคราะห์หารูปแบบความสัมพันธ์เชิงเส้นโค้งรูปตัวเอสระหว่างตัวแปรอิสระหลายตัวกับตัวแปรตามที่มีการกระจายแบบทวินาม (binomial) โดยที่ตัวแปรอิสระมีระดับการวัดได้ตั้งแต่ระดับนามบัญญัติ (nominal) ไปจนถึง ระดับอัตราส่วน (ratio) ในกรณีที่ตัวแปรอิสระมีระดับการวัดอยู่ในระดับอันตรภาคหรืออัตราส่วนอาจจัดรวมกลุ่มค่าของตัวแปรเพื่อให้ระดับการวัดต่ำลงมาได้ตามที่ผู้วิจัยต้องการ แต่วิธีนี้ไม่เป็นที่นิยมใช้เพราะเป็นวิธีที่ทำให้สูญเปล่าสารสนเทศของตัวแปร (นงลักษณ์ วิรัชชัย, 2538; Hosmer and Lemmeshow, 1989; Wittink, 1988)

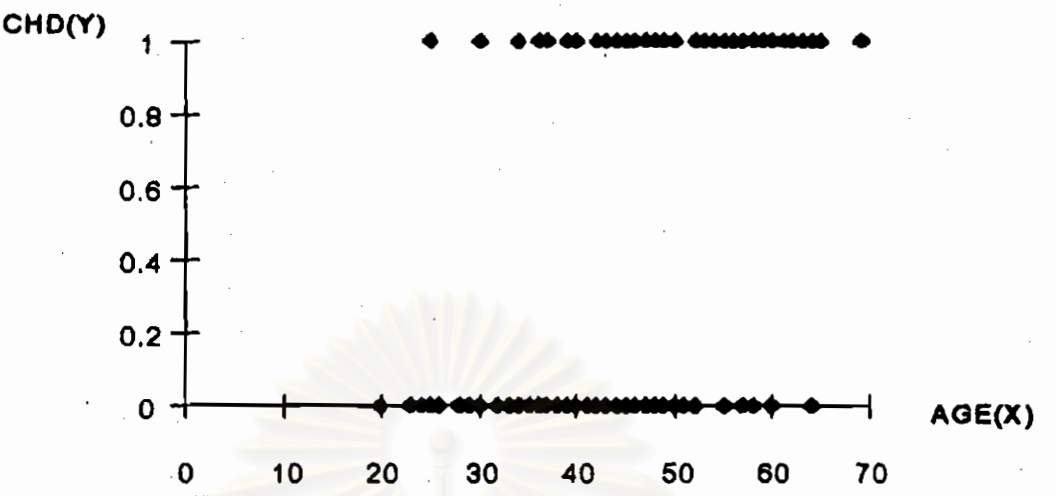
เพื่อให้เห็นถึงความแตกต่างและความเหมือนระหว่างการวิเคราะห์การถดถอยโลจิสติกและการวิเคราะห์การถดถอยพหุคูณ ขอยกตัวอย่างงานวิจัยที่ศึกษาเกี่ยวกับความสัมพันธ์ระหว่างอายุกับการเป็นโรคหลอดเลือดหัวใจ (ศิริเดช สุชีวะ, 2539; Norusis, 1992) ดังต่อไปนี้

ข้อมูลจากตารางที่ 1 แสดงค่าตัวแปรจำนวน 3 ตัว ได้แก่ กลุ่มอายุ (AGRP) อายุ (AGE) และการเป็นโรคหลอดเลือดหัวใจ (CHD) มี 2 ค่า คือ 0 และ 1 หมายถึงการไม่เป็นโรคหลอดเลือดหัวใจและการเป็นโรคหลอดเลือดหัวใจตามลำดับ เมื่อหาความสัมพันธ์ระหว่างอายุกับการเป็นโรคหลอดเลือดหัวใจจะได้แผนภาพการกระจาย (scatterplot) ดังภาพที่ 1

ตารางที่ 1 อายุและการเป็นโรคหลอดเลือดหัวใจของกลุ่มตัวอย่างจำนวน 100 คน

ID	AGRP	AGE	CHD	ID	AGRP	AGE	CHD	ID	AGRP	AGE	CHD
1	1	20	0	35	3	38	0	68	6	51	0
2	1	23	0	36	3	39	0	69	6	52	0
3	1	24	0	37	3	39	1	70	6	52	1
4	1	25	0	38	4	40	0	71	6	53	1
5	1	25	1	39	4	40	1	72	6	53	1
6	1	26	0	40	4	41	0	73	6	54	1
7	1	26	0	41	4	41	0	74	7	55	0
8	1	28	0	42	4	42	0	75	7	55	1
9	1	28	0	43	4	42	0	76	7	55	1
10	1	29	0	44	4	42	0	77	7	58	1
11	2	30	0	45	4	42	1	78	7	56	1
12	2	30	0	46	4	43	0	79	7	56	1
13	2	30	0	47	4	43	0	80	7	57	0
14	2	30	0	48	4	43	1	81	7	57	0
15	2	30	0	49	4	44	0	82	7	57	1
16	2	30	1	50	4	44	0	83	7	57	1
17	2	32	0	51	4	44	1	84	7	57	1
18	2	32	0	52	4	44	1	85	7	57	1
19	2	33	0	53	5	45	0	86	7	58	0
20	2	33	0	54	5	45	1	87	7	58	1
21	2	34	0	55	5	46	0	88	7	58	1
22	2	34	0	56	5	46	1	89	7	59	1
23	2	34	1	57	5	47	0	90	7	59	1
24	2	34	0	58	5	47	0	91	8	60	0
25	2	34	0	59	5	47	1	92	8	60	1
26	3	35	0	60	5	48	0	93	8	61	1
27	3	35	0	61	5	48	1	94	8	62	1
28	3	36	0	62	5	48	1	95	8	62	1
29	3	36	1	63	5	49	0	96	8	63	1
30	3	36	0	64	5	49	0	97	8	64	0
31	3	37	0	65	5	49	1	98	8	64	1
32	3	37	1	66	6	50	0	99	8	65	1
33	3	37	0	67	6	50	1	100	8	69	1
34	3	38	0								

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



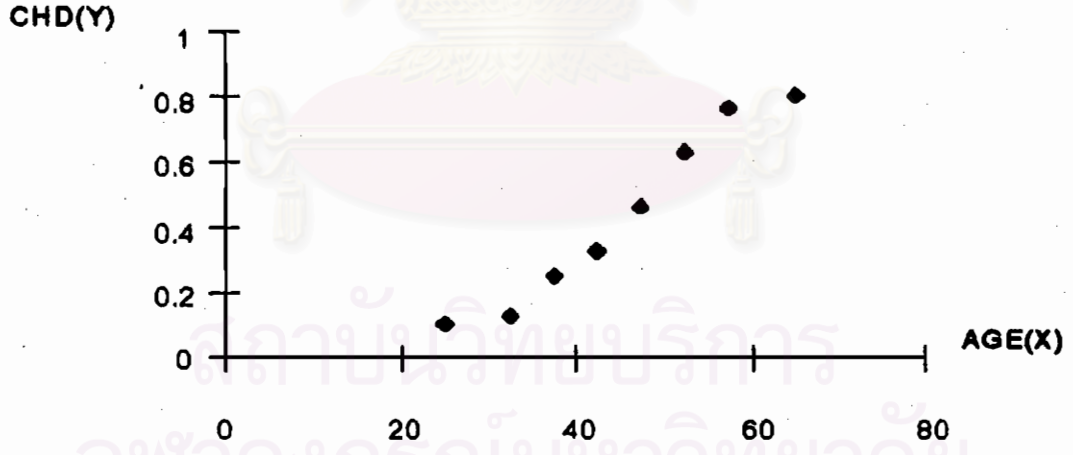
ภาพที่ 1 แผนภาพแสดงความสัมพันธ์ของการเป็นโรคหลอดเลือดหัวใจกับอายุ

จากแผนภาพการกระจายจะเห็นได้ว่าจุดทุกจุดที่ตกบนเส้นขนานแสดงการเป็นหรือไม่เป็นโรคหลอดเลือดหัวใจ (CHD = 1, CHD = 0) ซึ่งไม่สามารถมองเห็นลักษณะของความสัมพันธ์ระหว่างการเป็นโรคหลอดเลือดหัวใจกับอายุได้อย่างชัดเจน เนื่องจากความแปรปรวนของโรคหลอดเลือดหัวใจในทุกอายุมีมาก วิธีการที่จะลดความแปรปรวนดังกล่าวแต่ยังคงรูปแบบความสัมพันธ์ไว้ได้ คือ การสร้างอันตรภาคชั้นของตัวแปรอิสระและคำนวณค่าเฉลี่ยของตัวแปรตามในแต่ละอันตรภาคดังแสดงในตารางที่ 2 โดยใช้ตัวแปรกลุ่มอายุมาจัดอันตรภาคของอายุและใช้สัดส่วนของการเป็นโรคหลอดเลือดหัวใจแทนค่าเฉลี่ยในแต่ละกลุ่มอายุ จะทำให้เห็นความสัมพันธ์ระหว่างสัดส่วนของการเป็นโรคหลอดเลือดหัวใจกับช่วงอายุได้ชัดเจนมากยิ่งขึ้น ซึ่งแผนภาพดังกล่าวคล้ายกับแผนภาพการกระจายที่ได้จากการวิเคราะห์การถดถอยพหุคูณ ดังภาพที่ 2

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2 สัดส่วนของการเป็นโรคหลอดเลือดหัวใจตามกลุ่มอายุ

กลุ่มอายุ	จำนวน	โรคหลอดเลือดหัวใจ		สัดส่วนของการเป็นโรคหลอดเลือดหัวใจ
		ไม่เป็น	เป็น	
20 - 29	10	9	1	0.10
30 - 34	15	13	2	0.13
35 - 39	12	9	3	0.25
40 - 44	15	10	5	0.33
45 - 49	13	7	6	0.46
50 - 54	8	3	5	0.63
55 - 59	17	4	13	0.76
60 - 69	10	2	8	0.80
รวม	100	57	43	0.43



ภาพที่ 2 สัดส่วนของการเป็นโรคหลอดเลือดหัวใจกับช่วงอายุ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์การถดถอยพหุคูณและการวิเคราะห์การถดถอยโลจิสติกมีข้อแตกต่างกัน 3 ประการ ประการแรกเป็นความแตกต่างของลักษณะความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ในการวิเคราะห์การถดถอยพหุคูณนั้น ลักษณะความสัมพันธ์ระหว่างตัวแปรเป็นแบบเส้นตรง ค่าเฉลี่ยของตัวแปรตามจะถูกกำหนดค่าโดยตัวแปรอิสระที่เรียกว่า ค่าเฉลี่ยแบบมีเงื่อนไข (conditional mean) ใช้สัญลักษณ์ $E(Y|x)$ เมื่อ Y เป็นตัวแปรตาม และ x เป็นตัวแปรอิสระ ในการวิเคราะห์ถดถอยกรณีตัวแปรอิสระตัวเดียว ค่าเฉลี่ยนี้แสดงได้ดังสมการ $E(Y|x) = \beta_0 + \beta_1 x$ สมการนี้แสดงให้เห็นว่าค่า $E(Y|x)$ ที่เป็นไปได้บนค่าของ x มีพิสัยอยู่ระหว่าง $-\infty$ ถึง $+\infty$ ส่วนในการวิเคราะห์การถดถอยโลจิสติกซึ่งมีตัวแปรตามเป็นตัวแปรทวิวิภาคนั้น ตัวแปรตามมี 2 ค่า เช่น กำหนดให้ Y เป็นตัวแปรตาม มีค่าเป็น 1 และ 0 ค่าเฉลี่ยแบบมีเงื่อนไขจะมีค่าอยู่ระหว่าง 0 ถึง 1 ($0 \le E(Y|x) \le 1$) เมื่อนำมาเขียนกราฟระหว่างตัวแปรอิสระกับสัดส่วนของตัวแปรตาม จะได้กราฟเป็นรูปตัวเอส (S-shape) คล้ายกับรูปของการแจกแจงแบบสะสมของตัวแปรสุ่ม (cumulative distribution of a random variable) (ศิริเดช สุชีวะ, 2539; Hosmer and Lemmeshow, 1989; Wittink, 1988)

เพื่อให้ได้ฟังก์ชันที่ง่ายขึ้น จึงกำหนดให้ $\pi(x) = E(Y|x)$ แทนค่าเฉลี่ยแบบมีเงื่อนไขของ Y ที่กำหนดโดย x เมื่อใช้การแจกแจงแบบโลจิสติก โมเดลการถดถอยโลจิสติกเขียนได้ดังสมการ

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

การแปลงค่า $\pi(x)$ เป็นการแปลงแบบโลจิท (logit transformation) ซึ่งเขียนในรูปของลอการิทึมของอัตราส่วนแต้มต่อ (log odds ratio) ได้ดังนี้

$$g(X) = \ln\left[\frac{\pi(x)}{1 - \pi(x)}\right]$$ ในที่นี้ต่อไปจะมีค่า = Z

$$= \beta_0 + \beta_1 x$$

โดยที่ $\frac{\pi(x)}{1 - \pi(x)}$ = อัตราส่วนแต้มต่อ คำนวณจากอัตราส่วนของค่าความน่าจะเป็นที่จะเกิดเหตุการณ์ที่ Y มีค่าเป็น 1 กับค่าความน่าจะเป็นที่จะเกิดเหตุการณ์ที่ Y มีค่าเป็น 0

การแปลงแบบโลจิททำให้ g(X) หรือ Z มีคุณสมบัติตรงตามการวิเคราะห์การถดถอยเชิงเส้นตรงซึ่งอาจจะมีค่าต่อเนื่องได้ตั้งแต่ $-\infty$ ถึง $+\infty$ ขึ้นอยู่กับพิสัยของ X ดังนั้นการแปลความหมายผลการวิเคราะห์การถดถอยโลจิสติกจึงต้องระบุผลจากการเปลี่ยนแปลงของตัวแปรต้นที่มีต่ออัตราส่วนแต้มต่อในลักษณะเดียวกับการแปลความหมายในการวิเคราะห์ถดถอยพหุคูณ และนักวิจัยต้องนำค่าดอการิทึมของอัตราส่วนแต้มต่อมาแปลความหมายในรูปค่าความน่าจะเป็นได้โดยการคำนวณตามสูตร

$$\pi(X) = \frac{e^{g(x)}}{1 + e^{g(x)}} = \frac{e^{z}}{1 + e^{z}}$$

ความแตกต่างประการที่สองคือ ความแตกต่างของลักษณะการแจกแจงความคลาดเคลื่อน ในการวิเคราะห์การถดถอยพหุคูณนั้น ค่าของตัวแปรตามแสดงได้ดังสมการ $Y = E(Y|x) + \varepsilon$ โดยที่ ε เป็นค่าความคลาดเคลื่อน มีข้อตกลงเบื้องต้นว่าความคลาดเคลื่อนของตัวแปรตามที่ได้จากการพยากรณ์มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวนคงที่ แต่ในการวิเคราะห์การถดถอยโลจิสติกซึ่งตัวแปรตามเป็นตัวแปรทวิวิภาค (dichotomous) จะแสดงค่าของตัวแปรตามได้ดังสมการ $Y = \pi(x) + \varepsilon$ โดยที่ความคาดเคลื่อนมีค่าที่เป็นไปได้ 2 ค่า คือ ถ้า Y เท่ากับ 1 แล้ว ความคลาดเคลื่อนจะเท่ากับ $1 - \pi(x)$ ด้วยความน่าจะเป็น $\pi(x)$ และถ้า Y เท่ากับ 0 แล้ว ความคลาดเคลื่อนจะเท่ากับ $-\pi(x)$ ด้วยความน่าจะเป็น $1 - \pi(x)$ ดังนั้นความคลาดเคลื่อนจึงมีการแจกแจงแบบทวินาม มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเท่ากับ $\pi(x)\ [\ 1 - \pi(x)\ ]$

ความแตกต่างประการที่สาม คือ ความแตกต่างในเทคนิคการประมาณค่าพารามิเตอร์ กล่าวคือ ในการวิเคราะห์การถดถอยโลจิสติก จะประมาณค่าพารามิเตอร์ในโมเดลด้วยวิธีไลค์ลิฮู้ดสูงสุด (maximum likelihood) อันเป็นการคำนวณทวนซ้ำ (iterative algorithm) เพื่อให้ได้ค่าประมาณของพารามิเตอร์ที่ใกล้เคียงกับข้อมูลเชิงประจักษ์มากที่สุด แต่ในการประมาณค่าพารามิเตอร์ของการวิเคราะห์การถดถอยพหุคูณจะประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุด (ordinary least squares) โดยคัดเลือกสัมประสิทธิ์การถดถอยที่ทำให้ผลรวมกำลังสองของความคลาดเคลื่อนในการพยากรณ์มีค่าน้อยที่สุด โดยที่ความคลาดเคลื่อนคำนวณจากผลต่างระหว่างคะแนนที่วัดได้กับคะแนนที่เป็นค่าพยากรณ์

ข้อตกลงเบื้องต้น

การวิเคราะห์การถดถอยโลจิสติกมีข้อตกลงเบื้องต้น 2 ประการคือ ประการที่ 1 ตัวแปรตามเป็นตัวแปรทวิวิภาค (dichotomous or binary) ที่มีการแจกแจงแบบทวินาม ประการที่ 2 ตัวแปรอิสระเป็นตัวแปรที่มีการแจกแจงแบบปกติพหุนามซึ่งมีเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมเท่ากัน (multivariate normally distributed with equal covariance matrices) หรือเป็นตัวแปรทวิวิภาคและเป็นอิสระต่อกัน (multivariate independent dichotomous variables) หรือเป็นตัวแปรที่มีค่าไม่ต่อเนื่อง ส่งผลเป็นลักษณะความสัมพันธ์ที่เป็นเส้นตรงเมื่อค่าของตัวแปรนั้นถูกแปลงรูปให้อยู่ในรูปของเลขยกกำลังไม่ว่าจะเป็นกำลังสองหรือมากกว่า (Aldrich and Nelson, 1987; Hosmer and Lemmshow, 1989; Wittink, 1989)

โมเดลการวิเคราะห์การถดถอยโลจิสติก

กำหนดให้ $X_1, X_2, \ldots\ldots X_p$ เป็นตัวแปรอิสระ และ Y เป็นตัวแปรทวิภาค มีค่าเป็น 0 และ 1 โมเดลการวิเคราะห์ถดถอยโลจิสติกสามารถเขียนได้ดังสมการ

$$\text{prob (event)} = \pi(X) = \frac{e^Z}{1 + e^Z}$$

$$= \frac{1}{1 + e^{-Z}}$$

เมื่อ $\pi(X)$ = ความน่าจะเป็นของการเกิดเหตุการณ์ Y

Z = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_p X_p$

$$= \ln\left[\frac{\pi(X)}{1 - \pi(X)}\right]$$

β = ค่าสัมประสิทธิ์ที่ประมาณได้จากข้อมูล

e = ค่าคงที่ มีค่าประมาณ 2.7183

ความน่าจะเป็นของการไม่เกิดเหตุการณ์ Y จะสามารถประมาณได้จากสมการ

$$\text{Prob (event)} = \pi(X) = 1 - \text{Prob (no event)} = 1 - [1-\pi(X)]$$

หรือ $\text{Prob (no event)} = 1-\pi(X) = 1 - \text{Prob (event)}$

อัตราส่วนแต้มต่อของการเกิดเหตุการณ์ $Y = \dfrac{\pi(X)}{1 - \pi(X)}$

ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระแต่ละตัวที่ได้จากโมเดลการวิเคราะห์การถดถอยโลจิสติกสามารถแปลงค่าเป็นค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐาน (standard coefficients) ได้เช่นเดียวกันกับการวิเคราะห์การถดถอยพหุคูณ ซึ่งค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐานนี้เป็นค่าที่แสดงถึงโอกาสที่จะเกิดเหตุการณ์อย่างใดอย่างหนึ่งเมื่อตัวแปรอิสระมีการเปลี่ยนแปลงหนึ่งหน่วยของค่าเบี่ยงเบนมาตรฐาน

สำหรับการคำนวณค่าความสามารถในการอธิบายสัดส่วนความแปรปรวนของตัวแปรตามในโมเดลการวิเคราะห์การถดถอยโลจิสติก สามารถคำนวณได้จากสมการดังนี้ (McFadden, 1974)

$$\text{McFadden's } R^2 = 1 - \frac{LLf}{LLc}$$

เมื่อ LLf คือค่า ลอการิทึมของค่าไล้คลิฮูดที่ได้จากโมเดลที่มีความกลมกลืนกับข้อมูลเชิงประจักษ์ (Log Likelihood from the fitted model)

LLc คือค่า ลอการิทึมของค่าไล้คลิฮูดที่ได้จากโมเดลที่มีเฉพาะค่าคงที่ (Log Likelihood from a model containing only the constant)

การแปลความหมายของค่าสัมประสิทธิ์การพยากรณ์ McFadden's R^2 จะแปลความหมายในลักษณะของสัดส่วนความน่าจะเป็นในการเลือกทางเลือกทางใดทางหนึ่งสามารถอธิบายได้ด้วยชุดของตัวแปรอิสระในสมการนั้นๆ เหมือนกับค่าสัมประสิทธิ์การพยากรณ์ (multiple's R^2) ในการวิเคราะห์การถดถอยพหุคูณนั่นเอง

ในการวิเคราะห์การถดถอยโลจิสติก จะประมาณค่าพารามิเตอร์ในโมเดลโดยวิธีไลค์ลิฮู้ดสูงสุด (maximum likelihood) อันเป็นการคำนวณทวนซ้ำ (iterative algorithm) เพื่อให้ได้ค่าประมาณของพารามิเตอร์ที่ใกล้เคียงกับข้อมูลเชิงประจักษ์มากที่สุด ส่วนการแปลความหมายของสัมประสิทธิ์การถดถอยโลจิสติกต้องแปลความหมายในรูปของอัตราส่วนของแต้มต่อ (odds ratio) ซึ่งเป็นการเปลี่ยนแปลงของแต้มต่อ (odds) ตามการเปลี่ยนแปลงของตัวแปรอิสระ (odds หมายถึง อัตราส่วนของโอกาสที่จะเกิดกับโอกาสที่จะไม่เกิดของเหตุการณ์) ดังนั้นจึงเขียนโมเดลโลจิสติกในรูปของแต้มต่อได้ดังสมการ

$$\text{odds} = \frac{\text{prob (event)}}{\text{prob (no event)}} = \frac{\pi(x)}{1 - \pi(x)} = e^{\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p} = e^{\beta_0} e^{\beta_1 x_1} \ldots e^{\beta_p x_p}$$

ค่า e ยกกำลังด้วย β_i เป็นค่าแต้มต่อที่เปลี่ยนแปลง เมื่อตัวแปรอิสระตัวที่ i มีค่าเปลี่ยนไป 1 หน่วย ถ้า β_i มีค่าเป็นบวก เทอมนี้จะมีค่ามากกว่า 1 หมายความว่าค่าของแต้มต่อเพิ่มขึ้น ถ้า β_i มีค่าเป็นลบ เทอมนี้จะมีค่าน้อยกว่า 1 หมายความว่าค่าของแต้มต่อจะลดลง และถ้า β_i มีค่าเป็นศูนย์ เทอมนี้จะมีค่าเท่ากับ 1 หมายความว่าค่าของแต้มต่อไม่เปลี่ยนแปลง ดังนั้นในการแปลความหมายของสัมประสิทธิ์การถดถอยของการวิเคราะห์การถดถอยโลจิสติกจึงแตกต่างจากการแปลความหมายของสัมประสิทธิ์การถดถอยของการวิเคราะห์การถดถอยพหุคูณ กล่าวคือ ในการวิเคราะห์การถดถอยพหุคูณนั้นจะแปลความหมายของสัมประสิทธิ์การถดถอยได้ว่า เป็นขนาดการเปลี่ยนแปลงของตัวแปรตามเมื่อตัวแปรอิสระเปลี่ยนไป 1 หน่วย แต่การแปลความหมายของสัมประสิทธิ์การถดถอยในการวิเคราะห์การถดถอยโลจิสติกจะแปลความหมายว่า เป็นขนาดการเปลี่ยนแปลงของตัวแปรตามในรูปของอัตราส่วนแต้มต่อ เมื่อตัวแปรอิสระเปลี่ยนไป 1 หน่วย (ศิริเดช สุชีวะ, 2539; Hosmer and Lemmeshow, 1989)

สำหรับการทดสอบสมมุติฐานของโมเดลการวิเคราะห์การถดถอยโลจิสติก เพื่อทดสอบสมมุติฐานว่าค่าสัมประสิทธิ์การถดถอยไม่เท่ากับศูนย์นั้น จะใช้ค่าสถิติวาลด์ (wald statistic) ซึ่งมีการแจกแจงแบบไค-สแควร์ โดยที่ค่าสถิติวาลด์เป็นค่าที่คำนวณมาจากอัตราส่วนระหว่างค่าสัมประสิทธิ์การถดถอยกับค่าความคลาดเคลื่อนมาตรฐานของสัมประสิทธิ์การถดถอยยกกำลังสองนั่นเอง

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การตรวจสอบความตรง (Goodness of Fit) ของโมเดล

การตรวจสอบความตรงของโมเดลโลจิสติกเป็นการทดสอบความมีนัยสำคัญทางสถิติของตัวแปรอิสระทุกตัวที่จะเป็นตัวพยากรณ์ตัวแปรตามในโมเดล ซึ่งค่าสถิติที่ใช้ตรวจสอบความตรงของโมเดลคือค่าสถิติที่มีการกระจายแบบไคสแควร์ (Chi-square distribution) อาทิ Log likelihood ratio statistic, Hosmer-Lemeshow statistic, Q statistic เป็นต้น

วิธีการตรวจสอบความกลมกลืนระหว่างข้อมูลเชิงประจักษ์กับโมเดลการวิเคราะห์การถดถอยโลจิสติกทำได้หลายวิธี ดังต่อไปนี้ (ศิริเดช สุชีวะ, 2539; Hosmer and Lemmeshow, 1989; Menard, 1995)

1.ตารางจัดประเภท (Classification Table)

การเปรียบเทียบผลการพยากรณ์จากโมเดลการถดถอยโลจิสติกกับข้อมูลเชิงประจักษ์จากตารางจัดประเภท จะทำให้ผู้วิจัยทราบว่า ผลการพยากรณ์ที่ได้จากโมเดลมีความถูกต้องเป็นร้อยละเท่าใด ดังตัวอย่างงานวิจัยเกี่ยวกับการพยากรณ์การเป็นมะเร็งของ Brown (Brown, 1980 อ้างถึงใน Norusis, 1992) ที่ศึกษาเกี่ยวกับตัวแปรที่สามารถพยากรณ์การเป็นมะเร็ง (NODES) จำนวน 5 ตัวแปร คือ อายุ (AGE) ระยะของโรค (STAGE) ระดับการผิดปกติของเซล (GRADE) ผลการเอกซเรย์ (XRAY) และ กรดที่มีค่าสูงเมื่อพบว่ามีเซลที่เจริญเติบโตผิดปกติขึ้นในร่างกาย (ACID) ผลการวิจัยให้ตารางจัดประเภทดังตารางที่ 3

ตารางที่ 3 ตารางจัดประเภทในการพยากรณ์โอกาสการเป็นมะเร็ง

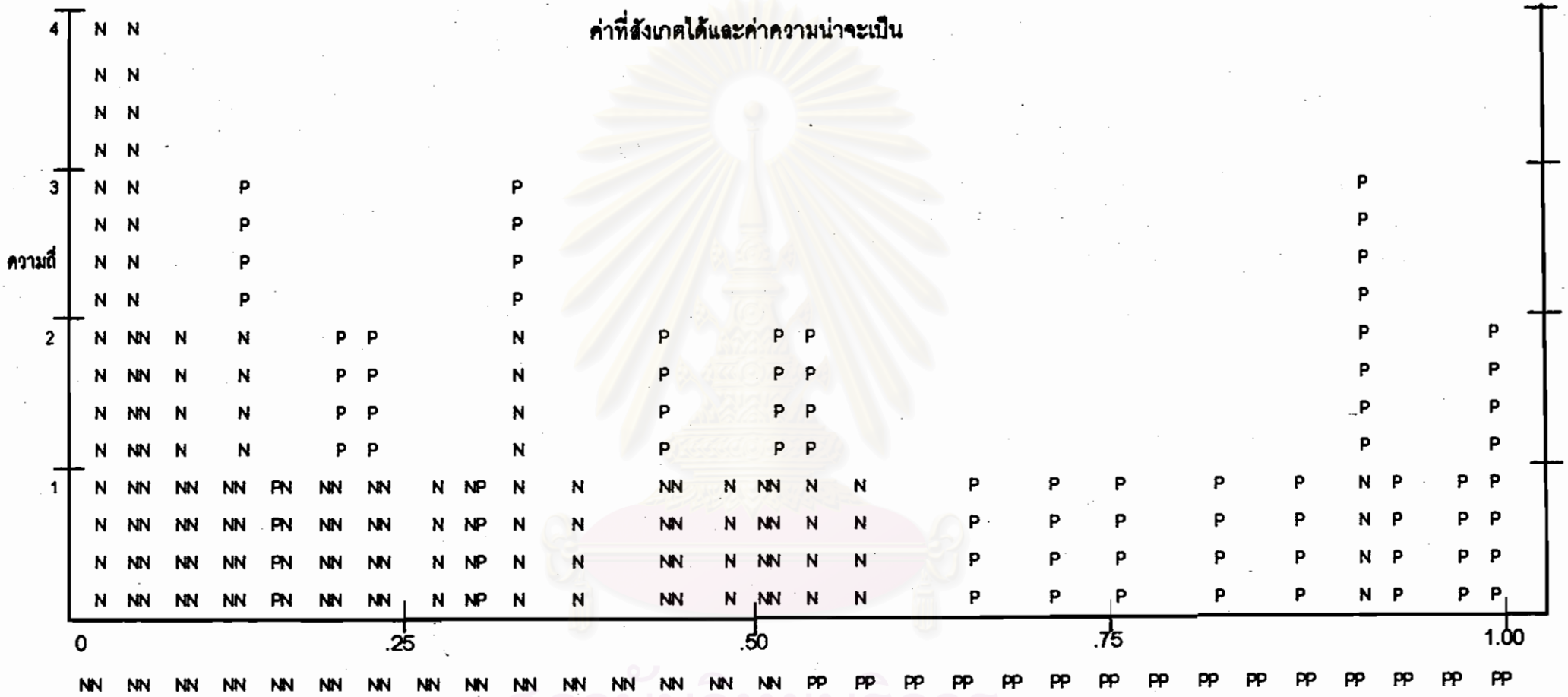
	ค่าพยากรณ์		
ค่าที่สังเกตได้	ไม่เป็นมะเร็ง	เป็นมะเร็ง	ร้อยละ
ไม่เป็นมะเร็ง	28	5	84.85%
เป็นมะเร็ง	7	13	65.00%
		ภาพรวม	77.36%

จากตารางจัดประเภท แสดงให้เห็นว่าโมเดลสามารถพยากรณ์ผู้ป่วยที่ไม่เป็นมะเร็งจำนวน 28 คนได้ถูกต้องว่าไม่เป็นมะเร็ง และสามารถพยากรณ์ผู้ที่เป็นมะเร็งจำนวน 13 คนได้อย่างถูกต้อง ส่วนในอีก 2 เซลที่เหลือเป็นจำนวนผู้ที่โมเดลพยากรณ์ผิด มีจำนวน 12 คน แสดงว่า โมเดลนี้สามารถพยากรณ์ผู้ที่ไม่เป็นมะเร็งได้ถูกต้องร้อยละ 84.85 พยากรณ์ผู้ที่เป็นมะเร็งได้ถูกต้องร้อยละ 65 และเมื่อพิจารณาภาพรวมพบว่า โมเดลสามารถพยากรณ์ได้ถูกต้องร้อยละ 77.63 จากผู้ป่วยทั้งหมด 53 คน แต่ตารางจัดประเภทไม่ได้แสดงให้เห็นถึงการแจกแจงความน่าจะเป็นของค่าประมาณที่ได้ จึงเป็นที่มาของการทดสอบแบบวิธีที่ 2

2.ฮีสโตแกรมค่าประมาณความน่าจะเป็น (Histogram of Estimated Probabilities)

การตรวจสอบความตรงของโมเดลด้วยวิธีนี้ ผลการพยากรณ์ที่ได้จากโมเดลจะอยู่ในรูปฮีสโตแกรมค่าประมาณความน่าจะเป็น การแจกแจงโอกาสความน่าจะเป็นแบ่งออกเป็น 2 ส่วน โดยใช้ค่าความน่าจะเป็นเท่ากับ 0.5 ในการแบ่ง จากตัวอย่างงานวิจัยของ Brown ผลการวิเคราะห์ข้อมูลจะได้ฮีสโตแกรมค่าประมาณความน่าจะเป็นดังภาพที่ 3 อธิบายได้ว่า สัญญลักษณ์ P แทนคนที่เป็นมะเร็ง และ N แทนคนที่ไม่เป็นมะเร็ง โดย P และ N แต่ละตัว แทนคนจำนวน 0.25 คน หากโมเดลนี้สามารถพยากรณ์คนทั้งสองกลุ่มได้อย่างถูกต้อง คนที่เป็นมะเร็งทั้งหมดจะอยู่ทางด้านขวาของค่า 0.5 ส่วนคนที่ไม่เป็นมะเร็งทั้งหมดจะอยู่ทางด้านซ้ายของค่า 0.5 ซึ่งสามารถพิจารณาจากรูปได้ว่า คนที่ไม่เป็นมะเร็งมีจำนวน 1 คนเท่านั้นที่มีค่าประมาณความน่าจะเป็นของการเป็นมะเร็งสูง มีค่าประมาณ 0.88 และคนที่เป็นมะเร็งมี จำนวน 4 คน ที่มีค่าประมาณความน่าจะเป็นโอกาสการเป็นมะเร็งต่ำกว่า 0.25

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ความน่าจะเป็นในการพยากรณ์ของสมาชิกในแต่ละกลุ่ม สัญญลักษณ์ : N = ไม่เป็นโรคมะเร็ง P = เป็นโรคมะเร็ง

สัญญลักษณ์แต่ละตัว แทนจำนวน .25 คน

ภาพที่ 3 ฮีสโตแกรมของค่าประมาณความน่าจะเป็นในการเป็นโรคมะเร็ง

3.ความกลมกลืนของโมเดล (Goodness of Fit of the Model)

การตรวจสอบความตรงของโมเดลด้วยวิธีนี้ ใช้การเปรียบเทียบค่าพารามิเตอร์ที่ประมาณได้จากโมเดลการถดถอยโลจิสติกกับค่าสถิติที่คำนวณได้จากข้อมูลเชิงประจักษ์ว่าใกล้เคียงกันมากน้อยเพียงใด วิธีการประมาณค่าทำได้โดยการสมมุติค่าพารามิเตอร์ขึ้นมา 1 ชุดแล้วหาค่าประมาณพารามิเตอร์ ที่เรียกว่า ไลค์ลิฮู้ด (likelihood) หรือความเป็นไปได้ของการที่จะได้ค่าสังเกตของกลุ่มตัวอย่างที่สมมุติค่าพารามิเตอร์ไว้ แต่โดยส่วนใหญ่แล้วไลค์ลิฮู้ดจะมีค่าน้อยกว่า 1 จึงใช้ค่า -2 คูณกับค่าลอการิทึม (logarithm) ของไลค์ลิฮู้ด (-2LL) เพื่อใช้ในการตรวจสอบความกลมกลืนของโมเดลกับข้อมูลเชิงประจักษ์ ซึ่งโมเดลที่ดีจะมีค่าไลค์ลิฮู้ดสูงเพราะว่า ถ้าหากค่าไลค์ลิฮู้ดที่ได้มีค่าเข้าใกล้ 1 หรือมีค่าเป็น 1 เมื่อใช้ -2 คูณกับค่าลอการิทึมของค่าไลค์ลิฮู้ดดังกล่าวจะทำให้มีค่าเข้าใกล้ 0 หรือเท่ากับ 0 แสดงว่าโมเดลกลมกลืนกับข้อมูลเชิงประจักษ์อย่างสมบูรณ์ ในการตรวจสอบความตรงของโมเดลจะใช้ค่า -2LL โดยกำหนดสมมุติฐานหลักว่า โมเดลมีความกลมกลืนกับข้อมูลเชิงประจักษ์ ซึ่งค่า -2LL เป็นค่าที่มีการแจกแจงแบบไค-สแควร์ มีระดับองศาอิสระเท่ากับจำนวนกลุ่มตัวอย่างลบด้วยจำนวนพารามิเตอร์ที่จะประมาณ หากผลการทดสอบสมมุติฐานพบว่า ไม่สามารถปฏิเสธสมมุติฐานหลักได้ แสดงว่าโมเดลมีความตรงและกลมกลืนกับข้อมูลเชิงประจักษ์

สถิติอีกชนิดหนึ่งที่ใช้ในการตรวจสอบความตรงของโมเดลคือ ค่าสถิติความกลมกลืน (goodness of fit statistic) ซึ่งจะเปรียบเทียบค่าความน่าจะเป็นที่สังเกตได้กับค่าความน่าจะเป็นที่พยากรณ์ได้จากโมเดล สถิติชนิดนี้มีการแจกแจงแบบไค-สแควร์ มีการกำหนดสมมุติฐานหลัก และให้ผลการทดสอบด้วยค่าสถิติความกลมกลืนเหมือนกับค่า-2ลอการิทึมไลค์ลิฮู้ด (-2LL) ทุกประการ

การคัดเลือกตัวแปรพยากรณ์

ในปัจจุบัน การวิเคราะห์การถดถอยโลจิสติกสามารถทำได้โดยใช้โปรแกรมคอมพิวเตอร์หลายแบบในการวิเคราะห์ข้อมูล เช่น SAS GLIM BMDP และ SPSS ทำให้นักวิจัยได้รับความสะดวกสบายมากยิ่งขึ้นในการวิเคราะห์ข้อมูล นักวิจัยที่ใช้การวิเคราะห์การถดถอยโลจิสติกสามารถคัดเลือกตัวแปรพยากรณ์ที่ดีได้เช่นเดียวกับการวิเคราะห์การถดถอยพหุคูณ ในคำสั่งการวิเคราะห์การถดถอยโลจิสติกอาจใช้คำสั่งรอง ENTER เพื่อให้ตัวแปรทุกตัวเข้าสู่สมการหรืออาจใช้คำสั่งรอง FSTEP สำหรับคัดเลือกแบบขั้นตอนก้าวหน้า (forward stepwise) หรือ BSTEP สำหรับ

คัดเลือกแบบขั้นตอนถอยหลัง (backward stepwise) ก็ได้ ในที่นี้จะกล่าวถึงการคัดเลือกตัวแปรพยากรณ์ที่สำคัญ 2 วิธี คือ (ศิริเดช สุชีวะ, 2539; Norusis, 1990)

1.การคัดเลือกตัวแปรพยากรณ์แบบขั้นตอนก้าวหน้า (forward stepwise) หลักการในการดำเนินการคัดเลือกตัวแปรพยากรณ์แบบขั้นตอนก้าวหน้าของการวิเคราะห์การถดถอยโลจิสติกการวิเคราะห์การถดถอยพหุคูณเป็นหลักการเดียวกัน กล่าวคือ การคัดเลือกตัวแปรพยากรณ์แบบนี้ เป็นเทคนิคการคัดเลือกตัวแปรพยากรณ์เข้าไปสร้างในโมเดลทีละขั้นตอนโดยเพิ่มตัวแปรพยากรณ์เข้าไปครั้งละหนึ่งตัว ในขณะที่ในโมเดลมีเพียงค่าคงที่เท่านั้น การคัดเลือกตัวแปรในแต่ละขั้นจะพิจารณาระดับนัยสำคัญทางสถิติของตัวแปรที่น้อยที่สุดก่อน เมื่อตัวแปรทุกตัวเข้ามาอยู่ในโมเดล ตัวแปรเหล่านั้นจะถูกทดสอบระดับนัยสำคัญเพื่อคัดเลือกออกจากโมเดลด้วยค่าสถิติวาลด์ ตัวแปรตัวใดมีระดับนัยสำคัญของค่าสถิติวาลด์มากที่สุด จะถูกคัดเลือกออกจากโมเดลเป็นอันดับแรก โดยกำหนดเกณฑ์ในการคัดเลือกตัวแปรเท่ากับ 0.1 เมื่อไม่มีตัวแปรอยู่ในโมเดลที่จะต้องคัดเลือกออกแล้ว คอมพิวเตอร์จะคัดเลือกตัวแปรเข้ามาในโมเดลและทดสอบระดับนัยสำคัญพร้อมทั้งคัดเลือกตัวแปรออกจากโมเดลอีกครั้ง หมุนเวียนไปจนกว่าไม่มีตัวแปรที่จะต้องคัดเลือกเข้าหรือคัดเลือกออกอีกแล้ว แต่การทดสอบระดับนัยสำคัญด้วยการทดสอบอัตราส่วนไล้ค์ลิฮู้ด (likelihood ratio, LR) จะมีความเหมาะสมมากกว่าการใช้ค่าสถิติวาลด์ เนื่องจากอัตราส่วนไล้ค์ลิฮู้ดมีวิธีการคำนวณที่สลับซับซ้อนมากกว่าวิธีการคำนวณด้วยค่าสถิติวาลด์ โดยที่วิธีการอัตราส่วนไล้ค์ลิฮู้ดจะทดสอบการเปลี่ยนแปลงค่าลอการิทึมไลค์ลิฮู้ด กล่าวคือ เมื่อตัวแปรถูกคัดเลือกออกจากโมเดล อัตราส่วนไล้ค์ลิฮู้ดจะทดสอบสมมติฐานหลักที่กำหนดไว้ว่า ค่าสัมประสิทธิ์การถดถอยที่ถูกคัดเลือกออกมีค่าเท่ากับ 0 คำนวณได้จากอัตราส่วนไล้ค์ลิฮู้ดของโมเดลสมบูรณ์แบบหารด้วยอัตราส่วนไล้ค์ลิฮู้ดของโมเดลที่มีการคัดเลือกตัวแปรออก ถ้าหากสมมุติฐานหลักเป็นจริง ค่า -2 ลอการิทึมไล้ค์ลิฮู้ด จะมีการแจกแจงแบบไค-สแควร์ ที่ระดับของศาอิสระเท่ากับความแตกต่างของจำนวนเทอมในโมเดลเต็มรูปกับโมเดลลดรูป เมื่อใช้ค่าอัตราส่วนไล้ค์ลิฮู้ดกับค่าจุดตัดในการคัดเลือกตัวแปรด้วยวิธีเปรียบเทียบระดับนัยสำคัญ จะมีขั้นตอนการคัดเลือกตัวแปรออกเหมือนกับที่กล่าวมา

2.การคัดเลือกตัวแปรพยากรณ์แบบขั้นตอนถอยหลัง (backward stepwise) การคัดเลือกตัวแปรพยากรณ์แบบขั้นตอนถอยหลัง เริ่มจากการนำตัวแปรพยากรณ์ทุกตัวเข้าในโมเดลเสียก่อน ต่อจากนั้น ตัวแปรที่อยู่ในโมเดลจะถูกคัดเลือกเข้าและออกทีละขั้นตอน โดยใช้ค่าสถิติในการกำหนดว่าตัวแปรควรจะถูกคัดเลือกเข้าสู่โมเดลหรือไม่ ส่วนการคัดตัวแปรออกจากโมเดลจะหลักการเหมือนกับการคัดเลือกตัวแปรพยากรณ์แบบขั้นตอนก้าวหน้า

จากที่กล่าวมา จะเห็นได้ว่าการวิเคราะห์การถดถอยโลจิสติกเป็นระเบียบวิธีทางสถิติอีกวิธีหนึ่งซึ่งเป็นทางเลือกใหม่ที่น่าสนใจสำหรับนักวิจัยที่ต้องการสร้างสรรค์งานวิจัยที่มีคุณภาพ เพราะนักวิจัยสามารถตอบคำถามวิจัยได้ลึกซึ้งเพิ่มมากขึ้นกว่าการวิเคราะห์การถดถอยพหุคูณ และวิธีการวิเคราะห์มีข้อตกลงเบื้องต้นน้อยกว่าการวิเคราะห์การถดถอยพหุคูณด้วย จุดเด่นอีกประการหนึ่งของการใช้การวิเคราะห์การถดถอยโลจิสติกในการวิจัยก็คือ การที่นักวิจัยสามารถตรวจสอบความกลมกลืนระหว่างโมเดลการวิจัยกับข้อมูลเชิงประจักษ์ได้ และยังสามารถแปลความหมายผลการวิเคราะห์ในรูปของความน่าจะเป็นของการเกิดเหตุการณ์ได้อีกด้วย

ตอนที่ 2 การวิเคราะห์อิทธิพล (Path Analysis)

การวิเคราะห์อิทธิพลเป็นวิธีการศึกษาความสัมพันธ์เชิงสาเหตุ (causal relationship) ระหว่างตัวแปรที่มีพื้นฐานการวิเคราะห์จากการวิเคราะห์การถดถอยพหุคูณ หลักการสำคัญคือการอธิบายความแปรปรวนของตัวแปรตามที่เกิดจากตัวแปรอิสระ โดยการศึกษาขนาดของอิทธิพลจากตัวแปรอิสระที่มีต่อตัวแปรตาม การวิเคราะห์อิทธิพลเริ่มต้นจากแนวคิดและผลงานของ Sewell Wright นักชีวมิติ ซึ่งได้พัฒนาเทคนิคและนำเข้ามาใช้ในวงการวิจัย จนเป็นที่แพร่หลายและได้รับการพัฒนาให้ดีมากยิ่งขึ้นในปัจจุบัน (นงลักษณ์ วิรัชชัย, 2538; Pedhazur, 1982; Bollen, 1989)

สังกัปเบื้องต้นของการวิเคราะห์อิทธิพล

การวิเคราะห์อิทธิพล เป็นวิธีการวิเคราะห์ข้อมูลทางสถิติเพื่อทดสอบสมมุติฐานเกี่ยวกับความสัมพันธ์เชิงสาเหตุในการวิจัยทั้งการวิจัยที่ไม่ใช่การทดลอง และการวิจัยเชิงทดลอง การพัฒนาในระยะแรกมุ่งใช้กับการวิจัยที่ไม่ใช่การทดลอง เนื่องจากในการวิจัยทางพฤติกรรมศาสตร์และสังคมศาสตร์นั้น การวิจัยเชิงทดลองที่แท้จริงซึ่งสามารถตอบคำถามวิจัยเกี่ยวกับความสัมพันธ์เชิงสาเหตุได้นั้นมีข้อจำกัดหลายประการ อาทิ ปัญหาในการจัดกระทำตัวแปรทางด้านพฤติกรรม ปัญหาในกระบวนการสุ่ม รวมทั้งปัญหาทางด้านจริยธรรมเกี่ยวกับการทดลอง ซึ่งผู้วิจัยต้องคำนึงถึงผลต่างๆที่จะตามมาต่อกลุ่มตัวอย่างที่ใช้ในการทดลองในระยะยาวด้วย เช่น การทำร้ายร่างกาย การมีตัวแปรจัดกระทำ ตัวแปรคั่นกลาง หรือตัวแปรที่เป็นตัวแปรกดหลายตัว จนไม่สามารถกำหนดแผนแบบการทดลองเพื่อตอบคำถามการวิจัยได้ ทำให้นักวิจัยพัฒนาการศึกษาความสัมพันธ์เชิงสาเหตุในลักษณะของการวิจัยที่ไม่ใช่การทดลองโดยการรวบรวมสารสนเทศเชิงคุณภาพจากทฤษฎี ข้อตกลงเบื้องต้น สมมุติฐานในการวิเคราะห์เชิงตรรกและสาร

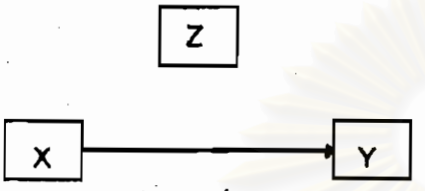
สนเทศเชิงปริมาณจากสหสัมพันธ์ชนิดต่างๆที่ได้จากการวิเคราะห์ข้อมูลเชิงปริมาณเพื่อให้ได้สารสนเทศที่มีความสมบูรณ์และเพียงพอที่จะศึกษาความสัมพันธ์เชิงสาเหตุได้ (นงลักษณ์ วิรัชชัย, 2538; ศิริชัย กาญจนวาสี, 2532)

การหาความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรในการวิจัยนั้น ทฤษฎี (theory) ที่เกี่ยวข้องกับเรื่องที่สนใจศึกษาจัดเป็นแนวคิดที่สำคัญในการเสนอคำอธิบายลำดับขั้นของการเกิดและลักษณะความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรแต่ละตัว ทฤษฎีจึงมีบทบาทสำคัญที่ช่วยเป็นพื้นฐานในการเชื่อมโยงความสัมพันธ์เชิงสาเหตุระหว่างตัวแปร เพื่อสร้างโมเดลเชิงสาเหตุ (causal models) ซึ่งแสดงรูปแบบของความเกี่ยวข้องสัมพันธ์เชิงสาเหตุระหว่างตัวแปร หรือให้แนวคิดในการปรับเปลี่ยนโครงสร้างความสัมพันธ์ระหว่างตัวแปรในโมเดล โมเดลเชิงสาเหตุนี้เป็นตัวกำหนดประเภทข้อมูลที่ต้องการ เครื่องมือในการเก็บรวบรวมข้อมูล และวิธีการวิเคราะห์ข้อมูล เพื่อตรวจสอบความน่าเชื่อถือของโมเดลเชิงสาเหตุที่พัฒนาขึ้น นักวิจัยต้องอาศัยวิธีการวิเคราะห์ด้วยเหตุและผล (logical analysis) หรืออาจทำการเก็บรวบรวมข้อมูลเชิงประจักษ์แล้วนำมาวิเคราะห์ทางสถิติ เพื่อตรวจสอบความสอดคล้องของโมเดล นักวิจัยอาจยอมรับหรือปฏิเสธโมเดลที่นักวิจัยพัฒนาขึ้นก็ได้ ถ้าโมเดลไม่สอดคล้องกับข้อมูล นักวิจัยอาจจะพิจารณาถึงทฤษฎีที่นำมาใช้เป็นพื้นฐานในการสร้างโมเดลว่ามีความเหมาะสมเพียงใด หรืออาจจะพิจารณาในเรื่องเกี่ยวกับความเหมาะสมของการออกแบบและการดำเนินการวิจัยเพื่อเสนอแนะในการวิจัยต่อไป แต่ถ้าพบว่าโมเดลมีความสอดคล้องกับข้อมูล ผลการวิเคราะห์นี้มิใช่หลักฐานของการพิสูจน์ทฤษฎีหรือโมเดล เพียงแต่แสดงว่ายังไม่มีหลักฐานเพียงพอที่จะนำมาปฏิเสธทฤษฎีหรือโมเดล กล่าวอีกนัยหนึ่งได้ว่า หลักฐานที่ได้สนับสนุนความเป็นไปได้ของทฤษฎีหรือโมเดลเชิงสาเหตุ โดยมีวิธีการวิเคราะห์ทางสถิติที่จะช่วยยืนยันหรือปฏิเสธโมเดลเชิงสาเหตุที่ผู้วิจัยพัฒนาขึ้นหลายวิธี วิธีที่นิยมกันและเป็นที่ยอมรับกันทั่วไปได้แก่ การวิเคราะห์อิทธิพลความสัมพันธ์เชิงสาเหตุระหว่างตัวแปร (path analysis) อันเป็นวิธีการที่ได้รับการพัฒนาขึ้นเป็นวิธีการสร้างโมเดลสมการโครงสร้าง (structural equation modeling) หรือ การวิเคราะห์โมเดลความสัมพันธ์โครงสร้างเชิงเส้น (linear structural relationship model = LISREL) (ศิริชัย กาญจนวาสี, 2536; Alwin and Robert, 1975; Bollen, 1989; Specht, 1975)

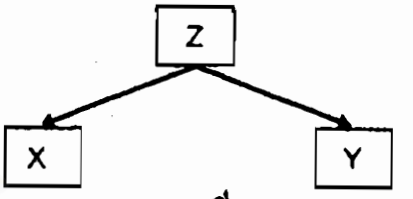
ความสัมพันธ์และความเป็นสาเหตุ (Correlation and Causation)

ในการวิจัยเพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรนั้น ในอดีตนักวิจัยใช้วิธีการศึกษาความสัมพันธ์ระหว่างตัวแปรจากการพิจารณาเปรียบเทียบค่าสหสัมพันธ์และสหสัมพันธ์พาร์เชียล

หรือสหสัมพันธ์บางส่วน (partial correlation) ประกอบกัน ผลของการเปรียบเทียบที่ต่างกัน แสดงให้เห็นความสัมพันธ์ระหว่างตัวแปรที่แตกต่างกัน นักวิจัยแบ่งลักษณะความสัมพันธ์ของตัวแปรได้เป็น 8 แบบ (Pedhazer, 1982 อ้างถึงใน นงลักษณ์ วิรัชชัย, 2538) แสดงด้วยภาพ เมื่อต้องการศึกษาความสัมพันธ์ระหว่างตัวแปร X และตัวแปร Y เมื่อมีตัวแปร Z เป็นตัวแปรทดสอบ (test variable) ดังต่อไปนี้

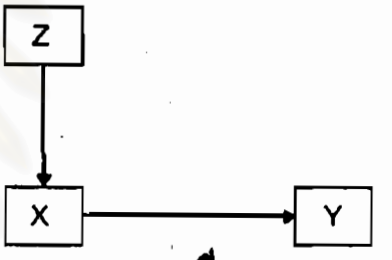


ภาพที่ 4



ภาพที่ 5

ภาพที่ 6



ภาพที่ 7

ภาพที่ 8

ภาพที่ 4 แสดงความสัมพันธ์ที่แท้จริง (true correlation)

ภาพที่ 5 แสดงความสัมพันธ์ลวง (spurious correlation)

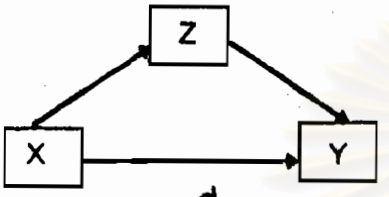
ภาพที่ 6 แสดงความสัมพันธ์เกิดจากอิทธิพลส่งผ่านตัวแปรคั่นกลาง (intervening variable)

ภาพที่ 7 แสดงความสัมพันธ์เกิดจากอิทธิพลของตัวแปรที่เป็นตัวกด (suppressor variable)

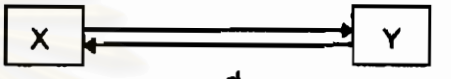
ภาพที่ 8 แสดงการไม่มีความสัมพันธ์ (no correlation)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

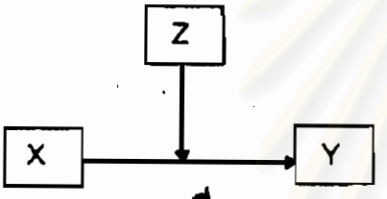
รูปแบบความสัมพันธ์ทั้ง 5 แบบที่กล่าวข้างต้นสามารถตรวจสอบได้โดยการหาค่าสัมประสิทธิ์สหสัมพันธ์แบบง่ายเปรียบเทียบกับค่าสัมประสิทธิ์สหสัมพันธ์พาร์เชียล แต่ยังมีรูปแบบความสัมพันธ์อีก 3 รูปแบบที่การวิเคราะห์สหสัมพันธ์ไม่สามารถตรวจสอบได้ ดังนี้



ภาพที่ 9



ภาพที่ 10



ภาพที่ 11

ภาพที่ 9 แสดงความสัมพันธ์มีอิทธิพลทางตรงและทางอ้อม (direct and indirect effect)
ภาพที่ 10 แสดงความสัมพันธ์เชิงสาเหตุย้อนกลับ (reciprocal causal relationship)
ภาพที่ 11 แสดงความสัมพันธ์มีเงื่อนไข (conditional relationship)

จากลักษณะรูปแบบความสัมพันธ์ระหว่างตัวแปร จะเห็นได้ว่ามีความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรแตกต่างกันหลายแบบลักษณะ ความสัมพันธ์เหล่านั้นมีอยู่จริงตามสภาพธรรมชาติ ดังนั้นในการวิจัยนักวิจัยต้องอธิบาย พยากรณ์ และควบคุมปรากฏการณ์เหล่านั้น โดยพยายามศึกษา และกำหนดลักษณะความสัมพันธ์เชิงสาเหตุเหล่านั้นให้ชัดเจน ทั้งนี้ความเป็นสาเหตุระหว่าง X และ Y นั้น ถ้า X เป็นเหตุทำให้เกิด Y จริงจะต้องมีเงื่อนไขดังนี้ ประการแรก X ต้องเกิดก่อน Y ประการที่สอง X และ Y มีความแปรปรวนร่วมกัน หรือสัมประสิทธิ์สหสัมพันธ์ระหว่าง X กับ Y มีนัยสำคัญ ประการที่สาม ต้องแสดงให้เห็นได้ว่า การเปลี่ยนแปลงค่าของตัวแปร X ทำให้เกิดการเปลี่ยนแปลงในตัวแปร Y และประการสุดท้าย มีเหตุผลในเชิงทฤษฎีสนับสนุน เพื่ออธิบายการที่ X ก่อให้เกิด Y (นงลักษณ์ วิรัชชัย, 2538; ศิริชัย กาญจนวาสี, 2536)

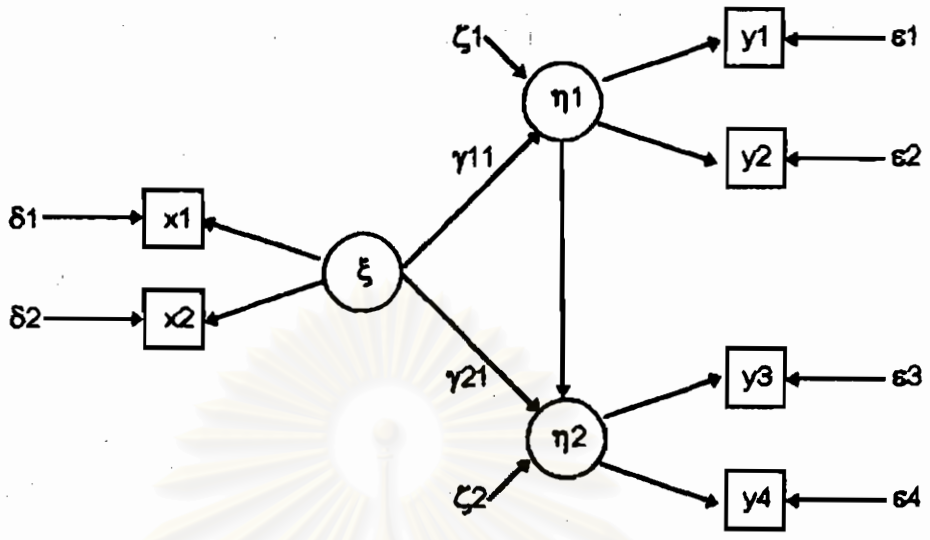
การพัฒนาโมเดลความสัมพันธ์เชิงสาเหตุ

การวิจัยทางพฤติกรรมศาสตร์และสังคมศาสตร์สามารถอธิบายปรากฏการณ์ในเชิงสาเหตุในการวิจัยที่ไม่ใช่การทดลองได้ โดยใช้เทคนิคของการวิเคราะห์โครงสร้างความสัมพันธ์ระหว่างตัวแปร ซึ่งมีชื่อเรียกต่างๆกัน เช่น การวิเคราะห์โครงสร้างความแปรปรวนร่วม (analysis of covariance structures) การวิเคราะห์สาเหตุ (causal analysis) การสร้างโมเดลสมการโครงสร้าง (structural equation modeling) การวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) หรือ โมเดลความสัมพันธ์โครงสร้างเชิงเส้น (linear structural relationship model) (Bentler, 1978; Joreskog and Sorbom, 1985 อ้างถึงใน ศิริชัย กาญจนวาสี, 2536) หลักการที่สำคัญของเทคนิคดังกล่าวคือการสร้างโมเดลเชิงสาเหตุและใช้ข้อมูลเชิงประจักษ์ตรวจสอบโมเดลที่สร้างขึ้น โมเดลเชิงสาเหตุเป็นกรอบหรือโครงสร้างความสัมพันธ์เชิงเหตุผลระหว่างตัวแปรหรือองค์ประกอบที่สามารถทดสอบได้ทางสถิติ โมเดลเชิงสาเหตุที่ดีจะต้องพัฒนาขึ้นมาจากรากฐานทางทฤษฎีที่แข็งแกร่งประกอบกับความรอบรู้ในเนื้อเรื่องของผู้สร้างโมเดล จากทฤษฎีเราสามารถโยงเข้าสู่การคัดสรรตัวแปรหรือองค์ประกอบที่เกี่ยวข้องกันเป็นระบบที่สมบูรณ์ของโมเดล ถ้าการคัดเลือกตัวแปรหรือองค์ประกอบปราศจากซึ่งความรู้ความเข้าใจเกี่ยวกับความเป็นมาและลำดับขั้นตอนของความเกี่ยวข้องเชิงสาเหตุแล้ว การคำนวณค่าสหสัมพันธ์ ไม่ว่าสหสัมพันธ์รวมหรือบางส่วน จะไม่ช่วยให้ผู้วิจัยก้าวสู่การประเมินเชิงสาเหตุของความสัมพันธ์ระหว่างตัวแปรหรือองค์ประกอบเหล่านั้น การพัฒนาโมเดลจะต้องผ่านการกลั่นกรองตรวจสอบและปรับปรุงแก้ไข จนได้โมเดลที่คิดว่าสมเหตุสมผลมากที่สุดจำนวนหนึ่ง เพื่อใช้เปรียบเทียบกัน นอกจากนี้นักวิจัยอาจใช้การสำรวจและวิเคราะห์ข้อมูลขั้นต้นมาช่วยในการปรับแก้โมเดลเหล่านั้น เมื่อปรับแก้โมเดลแล้วนักวิจัยต้องตรวจสอบความตรงของโมเดลว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์หรือไม่ ถ้าโมเดลสอดคล้องกับข้อมูลที่เก็บรวบรวมมาได้ โมเดลนั้นก็น่าเชื่อถือมากขึ้น แต่ถ้าข้อมูลขัดแย้งกับโมเดล โมเดลนั้นอาจไม่เหมาะสม ควรต้องแก้ไขปรับปรุงต่อไป ในทางอุดมคติมีความเป็นไปได้ที่อาจจะมีโมเดลหลายโมเดลสอดคล้องกับข้อมูลชุดเดียวกัน ดังนั้นจึงถือเป็นข้อแนะนำในการสร้างโมเดลเชิงสาเหตุที่มีทางเป็นไปได้หลายๆโมเดลโดยมีพื้นฐานทางทฤษฎีและข้อตกลงเบื้องต้นที่แตกต่างกัน เพื่อเป็นโมเดลเปรียบเทียบในการทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ เพื่อให้ได้โมเดลที่ดีที่สุดในการอธิบายความสัมพันธ์เชิงสาเหตุระหว่างตัวแปร (ศิริชัย กาญจนวาสี, 2536)

โมเดลลิสเรลสำหรับการวิเคราะห์อิทธิพล

การวิเคราะห์อิทธิพลด้วยวิธีดั้งเดิม เป็นวิธีการวิเคราะห์ที่อาศัยหลักการของการวิเคราะห์การถดถอยพหุคูณมีข้อตกลงเบื้องต้นส่วนหนึ่งเหมือนข้อตกลงเบื้องต้นในการวิเคราะห์การถดถอยพหุคูณ ข้อตกลงเบื้องต้นสำหรับการวิเคราะห์อิทธิพลที่สำคัญมี 5 ประการ คือ ประการที่ 1 ความสัมพันธ์ระหว่างตัวแปรเป็นเส้นตรง (linear) เป็นความสัมพันธ์เชิงบวก (additive) และเป็นความสัมพันธ์เชิงสาเหตุและผล ประการที่ 2 ส่วนเหลือ (residual) ของตัวแปรไม่มีความสัมพันธ์กับตัวแปรที่เกิดขึ้นก่อนตามที่ระบุไว้ในโมเดล ประการที่ 3 ความสัมพันธ์เชิงสาเหตุและผลระหว่างตัวแปรไม่มีการย้อนทิศทาง ประการที่ 4 การวัดค่าตัวแปรทุกตัวจะต้องมีการวัดค่าอยู่ในมาตรวัดอันตรภาค (interval scale)และประการที่ 5 การวัดค่าตัวแปรทุกค่าจะต้องไม่มีความคลาดเคลื่อน (error) ข้อตกลงเบื้องต้นทั้ง 5 ประการนี้ เป็นข้อตกลงเบื้องต้นที่เข้มงวดมากและไม่ตรงตามสภาพการณ์ที่แท้จริงเท่าใดนัก แต่การวิเคราะห์อิทธิพลด้วยลิสเรลสามารถผ่อนคลายข้อตกลงเบื้องต้นเหล่านี้ได้ อาทิ ความสัมพันธ์ระหว่างตัวแปรไม่จำเป็นต้องเป็นเส้นตรงเสมอไป ตัวแปรทุกตัวในโมเดลมีความคลาดเคลื่อนในการวัดได้ ตัวแปรทุกตัวไม่จำเป็นว่าจะต้องมีการวัดค่าอยู่ในมาตรวัดอันตรภาคเท่านั้น โปรแกรมลิสเรลสามารถวิเคราะห์ได้ทั้งกรณีที่ความสัมพันธ์ระหว่างตัวแปรไม่มีการย้อนทิศทางและมีการย้อนทิศทาง วิเคราะห์ได้ทั้งกรณีที่เทอมความคลาดเคลื่อนมีและไม่มีความสัมพันธ์กัน เป็นต้น จึงกล่าวได้ว่าในการวิเคราะห์อิทธิพลนั้น โปรแกรมลิสเรลสามารถวิเคราะห์ข้อมูลได้เหมาะสมกว่า (Bollen, 1989; Joreskog and Sorbom, 1989)

ลักษณะของโมเดลลิสเรลจะประกอบด้วยโมเดล 2 โมเดล คือ โมเดลการวัด (measurement model) และโมเดลสมการโครงสร้าง (structural equation model) โมเดลการวัดอธิบายความสัมพันธ์ระหว่างตัวแปรสังเกตได้กับตัวแปรแฝง (latent variable) ส่วนโมเดลสมการโครงสร้างอธิบายความสัมพันธ์ระหว่างตัวแปรแฝงด้วยกัน ดังแสดงในภาพที่ 12



ภาพที่ 12 ตัวแปรในโมเดลการวัด (measurement model) และ ตัวแปรในโมเดลสมการโครงสร้าง (structural equation model)

โมเดลในแผนภาพมีตัวแปรแฝงที่เป็นตัวแปรภายนอก 1 ตัว และตัวแปรแฝงที่เป็นตัวแปรภายใน 2 ตัว ตัวแปรแฝงทั้งสามตัว แต่ละตัววัดได้จากตัวแปรสังเกตได้ 2 ตัวแปร

ξ แทนเวกเตอร์ตัวแปรแฝงภายนอก
x แทนเวกเตอร์ตัวแปรภายนอกสังเกตได้
δ แทนเวกเตอร์ความคลาดเคลื่อนของ x
ζ แทนเวกเตอร์ความคลาดเคลื่อนของ η
η แทนเวกเตอร์ตัวแปรแฝงภายใน
y แทนเวกเตอร์ตัวแปรภายในสังเกตได้
ε แทนเวกเตอร์ความคลาดเคลื่อนของ y

Λy (LY) แทนเมทริกซ์สัมประสิทธิ์การถดถอยของ y บน η
Λx (LX) แทนเมทริกซ์สัมประสิทธิ์การถดถอยของ x บน ξ
β (BE) แทนเมทริกซ์สัมประสิทธิ์การถดถอยระหว่างตัวแปร η
Γ (GA) แทนเมทริกซ์สัมประสิทธิ์การถดถอยของ x ระหว่างตัวแปร ξ กับ η
Ψ (PS) แทนเมทริกซ์ความแปรปรวน - ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน ζ
Φ (PH)แทนเมทริกซ์ความแปรปรวน - ความแปรปรวนร่วมระหว่าง ξ
Θε (TE)แทนเมทริกซ์ความแปรปรวน - ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน ε
Θδ (TD) แทนเมทริกซ์ความแปรปรวน - ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน δ

ตัวแปรในโมเดลสมการโครงสร้าง (structural equation model) มีความสัมพันธ์กัน แสดงในรูปของสมการโครงสร้าง ดังนี้

$$\eta_1 = \gamma_{11}\xi_1 + \zeta_1$$

$$\eta_2 = \beta_{21}\eta_1 + \gamma_{21}\xi_1 + \zeta_2$$

ในที่นี้

$$\eta = \begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix} \quad \beta = \begin{bmatrix}0 & 0\\ \beta_{21} & 0\end{bmatrix} \quad \Gamma = \begin{bmatrix}\gamma_{11}\\ \gamma_{21}\end{bmatrix} \quad \xi = \begin{bmatrix}\xi_1\end{bmatrix} \quad \zeta = \begin{bmatrix}\zeta_1\\ \zeta_2\end{bmatrix}$$

เขียนสมการในรูปเมทริกซ์ได้ดังนี้

$$\eta = \beta\eta + \Gamma\xi + \zeta$$

ตัวแปรในโมเดลการวัด (measurement model) มีความสัมพันธ์กัน แสดงในรูปของสมการ ดังนี้

$$x_1 = \lambda_{x1}\xi + \delta_1, \qquad x_2 = \lambda_{x2}\xi_1 + \delta_2$$

$$y_1 = \lambda_{y1}\eta_1 + \varepsilon_1, \qquad y_2 = \lambda_{y2}\eta_2 + \varepsilon_2$$

$$y_3 = \lambda_{y3}\eta_3 + \varepsilon_3, \qquad y_4 = \lambda_{y4}\eta_4 + \varepsilon_4$$

ในที่นี้

$$x = \begin{bmatrix}x_1\\ x_2\end{bmatrix} \quad \Lambda x = \begin{bmatrix}\lambda_1\\ \lambda_2\end{bmatrix} \quad \xi = \begin{bmatrix}\xi_1\end{bmatrix} \quad \delta = \begin{bmatrix}\delta_1\\ \delta_2\end{bmatrix}$$

$$y = \begin{bmatrix}y_1\\ y_2\\ y_3\\ y_4\end{bmatrix} \quad \Lambda y = \begin{bmatrix}\lambda_1 & 0\\ \lambda_2 & 0\\ 0 & \lambda_3\\ 0 & \lambda_4\end{bmatrix} \quad \eta = \begin{bmatrix}\eta_2\\ \eta_1\end{bmatrix} \quad \varepsilon = \begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\end{bmatrix}$$

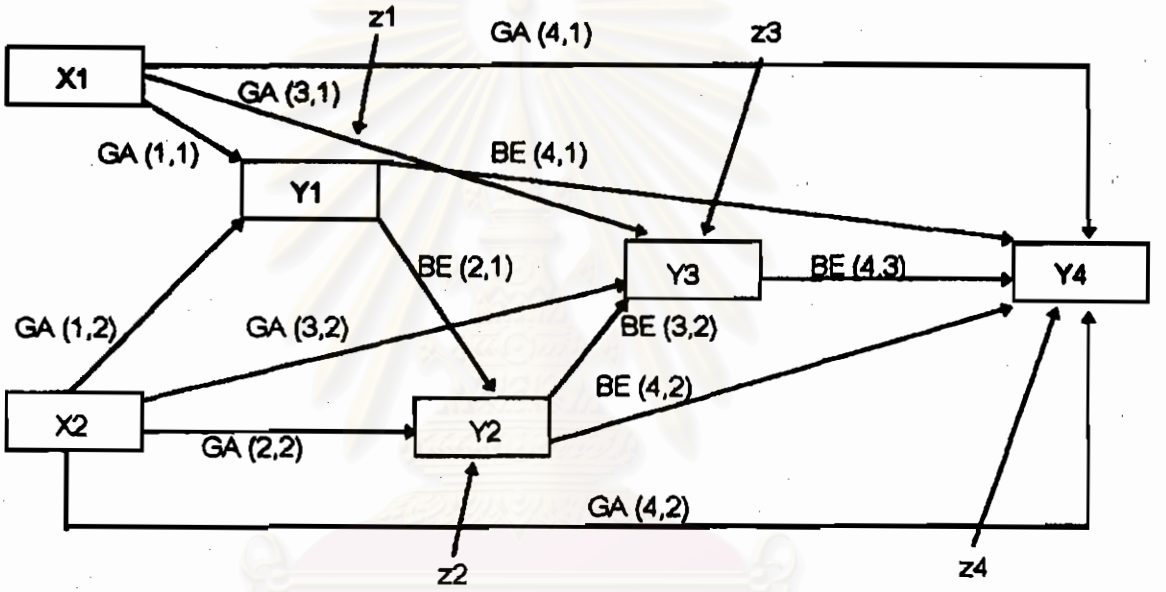
เขียนสมการในรูปเมทริกซ์ ดังนี้

$$x = \Lambda x\xi + \delta$$

$$y = \Lambda y\eta + \varepsilon$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์อิทธิพลด้วยโปรแกรมลิสเรลสามารถวิเคราะห์ได้ไม่ว่าโมเดลนั้นมีตัวแปรแฝงหรือตัวแปรสังเกตได้ วิเคราะห์ได้ทั้งโมเดลความสัมพันธ์ทางเดียวและความสัมพันธ์ย้อนกลับ แต่โมเดลลิสเรลสำหรับการวิเคราะห์อิทธิพลที่ใช้ในการศึกษาปัจจัยที่มีอิทธิพลต่อความคาดหวังในการศึกษาต่อของนักเรียนครั้งนี้ เป็นโมเดลลิสเรลสำหรับการวิเคราะห์อิทธิพลแบบไม่มีความคลาดเคลื่อนในการวัดและเป็นโมเดลความสัมพันธ์ทางเดียว ประกอบด้วยตัวแปรสังเกตได้ทุกตัว การกำหนดข้อมูลจำเพาะของโมเดลจึงกำหนดรูปแบบและสถานะของเมทริกซ์พารามิเตอร์ GA, BE, PH, และ PS เท่านั้น ดังแสดงในภาพที่ 13



ภาพที่ 13 โมเดลลิสเรลสำหรับการวิเคราะห์อิทธิพลในโมเดลความสัมพันธ์เชิงสาเหตุของความคาดหวังในการศึกษาต่อของนักเรียน

สมการโครงสร้างในรูปคะแนนมาตรฐาน

$$Y1 = GA(1,1)(X1) + GA(1,2)(X2) + z1$$

$$Y2 = GA(2,2)(X2) + BE(2,1)(Y1) + z2$$

$$Y3 = GA(3,1)(X1) + GA(3,2)(X2) + BE(3,2)(Y2) + z3$$

$$Y4 = GA(4,1)(X1) + GA(4,2)(X2) + BE(4,1)(Y1) + BE(4,2)(Y2) + BE(4,3)(Y3) + z4$$

หลักการวิเคราะห์โมเดลความสัมพันธ์เชิงสาเหตุในการวิเคราะห์อิทธิพล

หลักการวิเคราะห์โมเดลความสัมพันธ์เชิงสาเหตุในการวิเคราะห์อิทธิพลคือ การตรวจสอบความกลมกลืนระหว่างโมเดลที่เป็นสมมุติฐานการวิจัยกับข้อมูลเชิงประจักษ์ โดยอาศัยการเปรียบเทียบเมทริกซ์ความแปรปรวน - ความแปรปรวนร่วมเป็นเกณฑ์ในการเปรียบเทียบ กล่าวคือ ในการวิเคราะห์เพื่อประมาณค่าพารามิเตอร์และการตรวจสอบความกลมกลืนระหว่างโมเดลที่เป็นสมมุติฐานการวิจัยกับข้อมูลเชิงประจักษ์นั้น พารามิเตอร์จะถูกเลือกจากชุดของพารามิเตอร์ที่มีความเป็นไปได้สูงที่จะทำให้ค่าเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมที่ถูกสร้างขึ้นจากพารามิเตอร์ที่ประมาณค่าได้จากโมเดลที่เป็นสมมุติฐานการวิจัย (Σ) มีค่าใกล้เคียงกับค่าเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมที่คำนวณได้จากกลุ่มตัวอย่างอันเป็นข้อมูลเชิงประจักษ์ (S) มากที่สุด

เงื่อนไขความใกล้เคียงระหว่างเมทริกซ์ S และ Σ นั้น ใช้การสร้างฟังก์ชั่นความกลมกลืน (fit or fitting function) เป็นเกณฑ์ในการตรวจสอบ โดยรูปแบบของฟังก์ชั่นความกลมกลืนที่จะใช้คำนวณเพื่อประมาณค่าพารามิเตอร์ต้องมีคุณสมบัติ 3 ประการคือ ประการแรกฟังก์ชั่นความกลมกลืนต้องเป็นสเกลาร์ หรือเป็นเลขแสดงจำนวน มีขนาดแต่ไม่มีทิศทาง ประการที่สอง ฟังก์ชั่นความกลมกลืนต้องมีค่ามากกว่าหรือเท่ากับศูนย์ หากฟังก์ชั่นความกลมกลืนจะมีค่าเท่ากับศูนย์ได้ก็ต่อเมื่อเมทริกซ์ S และ Σ มีค่าเท่ากันเท่านั้น และประการสุดท้าย ฟังก์ชั่นความกลมกลืนต้องเป็นฟังก์ชั่นต่อเนื่อง วิธีการประมาณค่าพารามิเตอร์มีรูปแบบที่แตกต่างกันไป อาทิ วิธีกำลังสองน้อยที่สุดไม่ถ่วงน้ำหนัก (unweighted least squares) วิธีกำลังสองน้อยที่สุดวางนัยทั่วไป (generalized least squares) วิธีไลค์ลิฮู้ดสูงสุด (maximum likelihood) วิธีกำลังสองน้อยที่สุดถ่วงน้ำหนักทั่วไป (generally weighted least squares) หรือ วิธีกำลังสองน้อยที่สุดถ่วงน้ำหนักแนวทแยง (diagonally weighted least squares) เป็นต้น ดังนั้นจึงกล่าวได้ว่าหลักการวิเคราะห์โมเดลความสัมพันธ์เชิงสาเหตุก็คือ การหาค่าพารามิเตอร์ที่จะทำให้ค่าเมทริกซ์ทั้งสองใกล้เคียงกันมากที่สุดนั่นเอง (นงลักษณ์ วิรัชชัย, 2538; ศิริชัย กาญจนวาสี, 2532)

สำหรับเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมที่ใช้ในการประมาณค่าพารามิเตอร์นั้น โปรแกรมลิสเรลสามารถจัดเตรียมเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม หรือเมทริกซ์สหสัมพันธ์ได้หลายรูปแบบและมีความเหมาะสมกับลักษณะของข้อมูล เนื่องจากมีโปรแกรมพรีลิสในการเตรียมข้อมูลสำหรับการวิเคราะห์ด้วยโปรแกรมลิสเรล กล่าวคือ ในกรณีที่ข้อมูลมีระดับการวัดตั้งแต่มาตรอันตรภาคขึ้นไป นักวิจัยอาจจะใช้โปรแกรม SPSS หรือ SAS เตรียมข้อมูลสำหรับวิเคราะห์ด้วยลิสเรลได้ แต่ในกรณีที่ข้อมูลมีระดับการวัดแบบเรียงอันดับ นักวิจัยไม่สามารถใช้

i 17871141

โปรแกรม SPSS หรือ SAS ได้ ในขณะที่โปรแกรมพรีลิสสามารถเตรียมข้อมูลดังกล่าวได้ การวิเคราะห์อิทธิพลกรณีที่ตัวแปรอิสระและตัวแปรตามเป็นตัวแปรที่มีการวัดในระดับอันตรภาคจะใช้เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม หรือเมทริกซ์สหสัมพันธ์แบบเพียร์สัน (Pearson product moment correlation) ในการประมาณค่าพารามิเตอร์ ส่วนกรณีที่ตัวแปรอิสระและตัวแปรตามเป็นตัวแปรจัดประเภทหรือเป็นตัวแปรต่อเนื่องนั้น จะใช้เมทริกซ์สหสัมพันธ์แบบโพลีคอลิค (polychoric correlation) ซึ่งรวมทั้งสหสัมพันธ์เตตราคอลิค (polychoric including tetracholic correlation) และสหสัมพันธ์แบบโพลีซีเรียลซึ่งรวมทั้งสหสัมพันธ์ไบซีเรียล (polyserial including biserial correlation) ในการประมาณค่าพารามิเตอร์ (Bollen, 1989; Joreskog and Sorbom, 1988, 1989)

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับการศึกษาวิจัยเรื่องตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียน และงานวิจัยเกี่ยวกับการประยุกต์ใช้การวิเคราะห์การถดถอยโลจิสติกร่วมกับการวิเคราะห์อิทธิพล

การนำเสนองานวิจัยที่เกี่ยวข้องกับการศึกษาวิจัยครั้งนี้ ผู้วิจัยแยกนำเสนอเป็นสามตอน ตอนแรกนำเสนอการวิจัยทั้งในประเทศและต่างประเทศที่มีการศึกษาวิเคราะห์โดยใช้การวิเคราะห์การถดถอยโลจิสติก ตอนที่สองเป็นการนำเสนอเฉพาะการวิจัยในประเทศไทย ที่ศึกษาเรื่องตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียน โดยนำเสนองานวิจัยทุกเรื่อง ไม่ว่าจะใช้วิธีการวิเคราะห์แบบใด วิธีการนำเสนอเป็นการประมวลผลการวิจัยมาสังเคราะห์ สรุป จัดเป็นหมวดหมู่ตัวแปร เพื่อประมวลตัวแปรสำคัญที่เป็นสาเหตุของการศึกษาต่อและไม่ศึกษาต่อ ของนักเรียน มาสร้างโมเดลต่อไป และตอนที่สาม เป็นการนำเสนองานวิจัยเกี่ยวกับการประยุกต์ใช้การวิเคราะห์การถดถอยโลจิสติกร่วมกับการวิเคราะห์อิทธิพลเพื่อศึกษาและเปรียบเทียบความแตกต่างระหว่างงานวิจัยเรื่องดังกล่าวกับการวิจัยในครั้งนี้

ตอนที่ 3.1 การวิจัยในต่างประเทศที่มีการศึกษาเรื่องตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนโดยใช้การวิเคราะห์การถดถอยโลจิสติก

เนื่องจากการวิจัยที่ใช้การวิเคราะห์การถดถอยโลจิสติกในการศึกษาเรื่องตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนในประเทศไทยนั้น พบว่า ยังไม่มีผู้ใดได้ศึกษาเอาไว้เลย ผู้วิจัยจึงนำเสนอผลการวิจัยดังกล่าวที่มีเฉพาะในต่างประเทศเท่านั้น ดังต่อไปนี้

Valerie and David (1992) ได้ศึกษาเกี่ยวกับปัจจัยที่มีอิทธิพลต่อการออกจากโรงเรียนในขณะที่กำลังศึกษาอยู่หรือการย้ายสถานที่เรียนของนักเรียนระดับไฮสคูลในโรงเรียนรัฐบาลและโรงเรียนคาทอลิค มลรัฐมิชิแกน ประเทศสหรัฐอเมริกา ตัวแปรตามได้แก่ สถานภาพของนักเรียน แบ่งเป็น 3 กลุ่ม คือ การออกจากโรงเรียนในขณะที่กำลังศึกษาอยู่ การย้ายสถานที่เรียน และ การคงอยู่ของนักเรียน ส่วนตัวแปรพยากรณ์แบ่งออกเป็น 3 ด้าน ได้แก่ ด้านภูมิหลังของนักเรียน ประกอบด้วยตัวแปรเพศ (ชายหรือหญิง) ตัวแปรสีผิว (ผิวดำหรืออื่นๆ) ตัวแปรเชื้อชาติ (อีสพานิค หรืออื่นๆ) และ ตัวแปรอายุของผู้เรียน (อายุเกิน 16 ปี หรือตั้งแต่ 16 ปีลงไป) ด้านครอบครัว ประกอบด้วยตัวแปรสภาพครอบครัว (บิดามารดาอยู่ด้วยกันหรือหย่าร้าง) ตัวแปรจำนวนบุตรในครอบครัว และตัวแปรสถานภาพในสังคม (เกินค่าเฉลี่ยเกณฑ์มาตรฐานหรือต่ำกว่า) และด้านโรงเรียน ประกอบด้วยตัวแปรความสนใจในการเรียน (สนใจหรือไม่สนใจ) ตัวแปรการปฏิบัติตนต่อการเสี่ยง (เกินค่าเฉลี่ยเกณฑ์มาตรฐานหรือต่ำกว่า) และการขาดเรียนของนักเรียน (มากกว่า 10 ครั้งหรือน้อยกว่า) ใช้การวิเคราะห์แบบมัลติโนเมี่ยล (multinomial logit model) โดยให้กลุ่มนักเรียนที่คงอยู่เป็นฐานในการวิเคราะห์ข้อมูล ได้ข้อค้นพบดังนี้

ในโรงเรียนรัฐบาลพบว่า ตัวแปรความเป็นเพศหญิงมีอิทธิพลต่อการออกจากโรงเรียนในขณะที่ศึกษาอยู่มากกว่าการย้ายที่เรียน ตัวแปรผิวสีดำมีอิทธิพลต่อการย้ายที่เรียนมากกว่าการออกจากโรงเรียน ตัวแปรเชื้อชาติอีสพานิคมีอิทธิพลต่อการย้ายที่เรียนเพียงอย่างเดียว ตัวแปรอายุเกิน 16 ปี ตัวแปรสถานภาพในสังคม และ ตัวแปรความสนใจในการเรียน มีอิทธิพลต่อการออกจากโรงเรียนเท่านั้น ตัวแปรสภาพครอบครัว ตัวแปรจำนวนบุตรในครอบครัว ตัวแปรการปฏิบัติตนต่อการเสี่ยง และการขาดเรียนของนักเรียนเกิน 10 ครั้งมีอิทธิพลต่อการออกจากโรงเรียนในขณะที่กำลังศึกษาอยู่หรือการย้ายสถานที่เรียนของนักเรียนใกล้เคียงกัน ส่วนในโรงเรียนคาทอลิค ได้ข้อค้นพบว่า ตัวแปรสภาพครอบครัวมีอิทธิพลต่อการออกจากโรงเรียนเพียงอย่างเดียว ตัวแปรเชื้อชาติอีสพานิคมีอิทธิพลต่อการย้ายที่เรียนเพียงอย่างเดียว ตัวแปรความสนใจในการเรียน และตัวแปรการปฏิบัติตนต่อการเสี่ยงมีอิทธิพลต่อการออกจากโรงเรียนในขณะที่ศึกษาอยู่มากกว่าการย้ายที่เรียน ตัวแปรผิวสีดำ ตัวแปรอายุเกิน 16 ปี ตัวแปรสถานภาพทางสังคม และตัวแปรจำนวนบุตรในครอบครัวมีอิทธิพลต่อการออกจากโรงเรียนในขณะที่ศึกษาอยู่ใกล้เคียงกับการย้ายที่เรียน ส่วนตัวแปรเพศหญิงและตัวแปรการขาดเรียนเกิน 10 ครั้งนั้นไม่มีอิทธิพลต่อการออกจากโรงเรียนในขณะที่กำลังศึกษาอยู่หรือการย้ายสถานที่เรียนของนักเรียน

จากการศึกษาของ Garrison (1992) ที่ตรวจสอบความตรงของโมเดลเกี่ยวกับตัวแปรที่พยากรณ์แนวโน้มในการออกจากโรงเรียนในขณะที่กำลังศึกษาอยู่ ของนักเรียนระดับมัธยมศึกษา ในมลรัฐเท็กซัส ประเทศอเมริกา จำนวน 6 โมเดล ได้แก่ โมเดลของนักเรียนระดับ 8, โมเดลของนักเรียนระดับ 9, โมเดลของนักเรียนระดับ 10, โมเดลของนักเรียนระดับ 11, โมเดลของนักเรียนระดับ 12 และ โมเดลของนักเรียนระดับ 8 ถึง 12 โดยใช้การวิเคราะห์การถดถอยโลจิสติกในการตรวจสอบความตรงของโมเดลและคำนวณค่าพยากรณ์ภาพรวมของโมเดลที่สร้างขึ้น ปรากฏว่า โมเดลของนักเรียนระดับ 12 เพียงโมเดลเดียวเท่านั้นที่มีความตรงและกลมกลืนกับข้อมูล และโมเดลดังกล่าวสามารถพยากรณ์ภาพรวมได้ถูกต้องถึงร้อยละ 72.70

Valerie, Christopher and Helen (1993) ได้ศึกษาและติดตามผลเกี่ยวกับนักศึกษาที่สำเร็จการศึกษาในระดับปริญญาตรีจากวิทยาลัยชุมชนแล้ว ว่ามีแนวโน้มที่จะศึกษาต่อในระดับที่สูงขึ้นอีกหรือไม่ อย่างไร ซึ่งเป็นการศึกษาระยะยาว (longitudinal studies) โดยวัดการเปลี่ยนแปลง 4 ครั้ง (เดือนกุมภาพันธ์ ค.ศ. 1980 ถึง เดือนกุมภาพันธ์ ค.ศ. 1986) ตัวแปรตามในการวิจัยได้แก่ สถานภาพทางการศึกษา แบ่งออกเป็น 3 กลุ่มคือ กลุ่มที่ต้องการสำเร็จการศึกษาระดับปริญญาตรีเท่านั้น กลุ่มที่สำเร็จการศึกษาแล้วเข้าศึกษาต่อในระดับปริญญาโท และกลุ่มที่สำเร็จการศึกษาแล้วคาดว่าจะเข้าศึกษาต่อในระดับปริญญาโท ตัวแปรพยากรณ์แบ่งออกเป็น 3 ด้าน ได้แก่ ด้านภูมิหลังของนักศึกษาประกอบด้วยตัวแปรเพศ ตัวแปรสีผิว ตัวแปรเชื้อชาติ และตัวแปรสถานทางสังคมในโรงเรียน ด้านความประพฤติของนักศึกษา ประกอบด้วยตัวแปรการพักอยู่ในวิทยาลัย ตัวแปรการเป็นนักศึกษาภาคปกติ ตัวแปรความสนใจในการเรียน ตัวแปรความพึงพอใจในสังคมของโรงเรียน ตัวแปรผลสัมฤทธิ์ทางการเรียน และ ตัวแปรการได้เรียนในสาขาเกี่ยวกับวิทยาศาสตร์หรือเทคโนโลยีหรือไม่ และด้านสถาบันที่จะศึกษาต่อ ประกอบด้วยตัวแปร การที่สถาบันตั้งอยู่ในรัฐฟลอริด้า ตั้งอยู่ในรัฐแคลิฟอเนียร์ ตัวแปรสถาบันของเอกชน ตัวแปรวิทยาลัยที่มีนักศึกษามากกว่า 15,000 คน ตัวแปรวิทยาลัยที่มีนักศึกษาน้อยกว่า 1,500 คน ตัวแปรสาขาวิชาที่เลือกมีนักศึกษาน้อยกว่าร้อยละ10 ตัวแปรสาขาวิชาที่เลือกมีนักศึกษามากกว่าร้อยละ30 ตัวแปรการเคยได้รับรางวัลทางวิชาการของสถาบัน และ ตัวแปรการเลือกคัดเลือกของสถาบัน นอกจากนี้ยังมีตัวแปรพยากรณ์อีก 2 ตัว คือ ตัวแปรการเคยย้ายที่เรียนในปี ค.ศ.1984 และ ตัวแปรการสำเร็จการศึกษาในปี ค.ศ.1986 โดยใช้การวิเคราะห์ถดถอยโลจิสติกในการคำนวณค่าสัมประสิทธิ์การถดถอยในรูปของออดการิทึมของแต้มต่อเพื่อศึกษาว่าตัวแปรแต่ละตัวนั้นมีอิทธิพลต่อตัวแปรตามเป็นอย่างไร ผลการวิจัยสรุปได้ว่า

ตัวแปรที่มีอิทธิพลต่อกลุ่มที่ต้องการสำเร็จการศึกษาระดับปริญญาตรีเท่านั้นอย่างมีนัยสำคัญทางสถิติได้แก่ ตัวแปรสีผิว (การมีผิวสีดำ) ตัวแปรการเป็นนักศึกษาภาคปกติ ตัวแปรผลสัมฤทธิ์ทางการเรียน ตัวแปรสาขาวิชาที่เลือกมีนักศึกษาน้อยกว่าร้อยละ10 ตัวแปรสาขาวิชาที่เลือกมีนักศึกษามากกว่าร้อยละ30 และตัวแปรการเคยได้รับรางวัลทางวิชาการของสถาบัน ตัวแปรที่มีอิทธิพลต่อกลุ่มที่สำเร็จการศึกษาแล้วเข้าศึกษาต่อในระดับปริญญาโท ได้แก่ ตัวแปรการพักอยู่ในสถาบัน ตัวแปรความสนใจในการเรียน ตัวแปรผลสัมฤทธิ์ทางการเรียน ตัวแปรการสำเร็จการศึกษาในปี ค.ศ.1986 ตัวแปรวิทยาลัยที่มีนักศึกษามากกว่า 15,000 คน ตัวแปรสาขาวิชาที่เลือกมีนักศึกษาน้อยกว่าร้อยละ10 ตัวแปรการเคยได้รับรางวัลทางวิชาการของสถาบัน และตัวแปรการเลือกคัดเลือกของสถาบัน ส่วนตัวแปรที่มีอิทธิพลต่อกลุ่มที่สำเร็จการศึกษาแล้วคาดว่าจะเข้าศึกษาต่อในระดับปริญญาโท ได้แก่ ตัวแปรสีผิว (ผิวสีดำ) ตัวแปรการพักอยู่ในสถาบัน ตัวแปรความสนใจในการเรียน ตัวแปรผลสัมฤทธิ์ทางการเรียน ตัวแปรการสำเร็จการศึกษาในปี ค.ศ. 1986 และตัวแปรการเลือกคัดเลือกของสถาบัน

ตอนที่ 3.2 จากการศึกษาวิจัยในอดีตพบว่า มีตัวแปรจำนวนมากที่มีความสัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนที่สำเร็จการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 ดังตารางที่ 4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 ตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนที่สำเร็จการศึกษาในระดับชั้นประถมศึกษาปีที่ 6

ผู้วิจัย / ตัวแปร	Blake	Waite	Worthington	จรูญ	จิรา วัลย์	ดำรง	เทียม จันทร์	ธวัช	บุญเพ็ง	ประยงค์	ประเสริฐ	ดวงดี	สุนทร
1.อาชีพของผู้ปกครอง	/	/	/	/		/	/	/	/	/		/	/
2.รายได้ของผู้ปกครอง	/	/	/	/	/	/	/	/	/	/	/	/	/
3.ระดับการศึกษาของผู้ปกครอง	/	/	/	/	/	/	/	/	/	/		/	/
4.จำนวนบุตร				/		/		/		/	/		
5.จำนวนบุตรที่ได้รับการศึกษาชั้นสูง				/		/	/			/			/
6.สถานภาพสมรสของบิดามารดา						/		/				/	
7.ความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาชั้นสูง					/	/	/	/	/		/	/	/
8.ความต้องการใช้แรงงาน							/		/		/	/	
9.ความต้องการศึกษาต่อของนักเรียน							/	/	/		/		/
10.ระยะทางระหว่างบ้านถึงโรงเรียน						/		/					/
11.การคมนาคม						/		/					/
12.ผลสัมฤทธิ์ทางการเรียน			/	/		/	/	/	/	/	/	/	/

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 4 พบว่าตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนที่สำเร็จการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 สรุปได้ดังนี้

อาชีพของผู้ปกครอง จากการศึกษาของ จรูญ พานิชผลินไชย (2521) ดำรง แสนสิงห์ (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) ธวัช แก้วอนันต์ (2533) บุญเพ็ง ธานี (2530) ประยงค์ ชูน้อย (2521) สวลี อาชาสุวรรณ (2532) สุนทร โตรัตน์ (2524) Blake (1985) Waite, Rindfuss and tray (1986) และ Worthington and Grant (1971) ได้ข้อค้นพบที่สอดคล้องกันว่า อาชีพของผู้ปกครอง เป็นตัวแปรที่สัมพันธ์กับการศึกษาต่อในระดับชั้นที่สูงขึ้นของนักเรียน กล่าวคือ ผู้ปกครองของนักเรียนที่มีอาชีพที่ทำรายได้ให้กับครอบครัวสูงมีผลต่อโอกาสที่บุตรจะได้รับการศึกษาขั้นสูงกว่าการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 มากกว่าผู้ปกครองของนักเรียนที่มีอาชีพที่ทำรายได้ต่ำ

รายได้ของผู้ปกครอง จากการศึกษาของจรูญ พานิชผลินไชย (2521) จิราวัลย์ มนต์กันภัย (2534) ดำรง แสนสิงห์ (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) ธวัช แก้วอนันต์ (2533) บุญเพ็ง ธานี (2530) ประยงค์ ชูน้อย (2521) ประเสริฐ แก้วเพ็ชร (2528) สวลี อาชาสุวรรณ (2532) สุนทร โตรัตน์ (2524) Blake (1985) Waite, Rindfuss and tray (1986) และ Worthington and Grant (1971) ได้ข้อค้นพบที่สอดคล้องกันว่ารายได้ของผู้ปกครอง เป็นตัวแปรที่สัมพันธ์กับการศึกษาต่อของนักเรียน กล่าวคือผู้ปกครองของนักเรียนที่มีรายได้มาก มีผลต่อโอกาสการศึกษาต่อของบุตรในระดับสูงกว่าการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 มากกว่าผู้ปกครองของนักเรียนที่มีรายได้น้อย

ระดับการศึกษาของผู้ปกครอง จากการศึกษาของ จรูญ พานิชผลินไชย (2521) จิราวัลย์ มนต์กันภัย (2534) ดำรง แสนสิงห์ (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) ธวัช แก้วอนันต์ (2533) บุญเพ็ง ธานี (2530) ประยงค์ ชูน้อย (2521) สวลี อาชาสุวรรณ (2532) สุนทร โตรัตน์ (2524) Blake (1985) Waite, Rindfuss and tray (1986) และ Worthington and Grant (1971) ได้ข้อค้นพบที่สอดคล้องกันว่าระดับการศึกษาของผู้ปกครอง เป็นตัวแปรที่สัมพันธ์กับการศึกษาต่อของนักเรียน กล่าวคือผู้ปกครองของนักเรียนที่มีระดับการศึกษาสูงย่อมเห็นคุณค่าของการศึกษา ทำให้นักเรียนที่มีผู้ปกครองที่มีระดับการศึกษาสูงมีโอกาสที่จะศึกษาต่อในระดับที่สูงกว่าการศึกษาในระดับชั้นประถมศึกษาปีที่ 6

จำนวนบุตร จากการศึกษาของ จรูญ พานิชผลินไชย (2521) ดำรง แสนสิงห์ (2534) ธวัช แก้วอนันต์ (2533) ประยงค์ ชูน้อย (2521) และ ประเสริฐ แก้วเพ็ชร (2528) พบว่า นักเรียนที่มีจำนวนพี่น้องมากหรืออาศัยอยู่ในครอบครัวที่มีขนาดใหญ่มีโอกาสศึกษาต่อน้อยกว่านักเรียนที่มีพี่น้องจำนวนน้อย หรืออาศัยอยู่ในครอบครัวที่มีขนาดเล็ก

จำนวนบุตรที่ได้รับการศึกษาขั้นสูง จากการศึกษาของจรูญ พานิชผลินไชย (2521) ดำรง แสนสิงห์ (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) ประยงค์ ชูน้อย (2521) และ สุนทร โตรัตน์(2524) พบว่านักเรียนที่มีพี่ศึกษาอยู่ในระดับที่สูงกว่าชั้นประถมศึกษาปีที่ 6 จะศึกษาต่อมากกว่านักเรียนที่ไม่มีพี่ศึกษาต่อในระดับที่สูงกว่าชั้นประถมศึกษาปีที่ 6

สถานภาพสมรสของบิดามารดา จากการศึกษาของ ดำรง แสนสิงห์ (2534) ธวัช แก้วอนันต์ (2533) และสวลี อาชาสุวรรณ (2532) พบว่า สถานภาพสมรสของบิดามารดามีความสัมพันธ์กับความต้องการศึกษาต่อในระดับชั้นที่สูงขึ้นของนักเรียน โดยที่นักเรียนที่มีผู้ปกครองหรือบิดามารดาอยู่ด้วยกันจะมีโอกาสในการได้ศึกษาต่อสูงกว่านักเรียนที่ผู้ปกครองหรือบิดามารดาหย่าร้างกัน

ความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง จากการศึกษาของจิรา วัลย์ มนต์กันภัย (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) บุญเพ็ง ธานี (2530) ประเสริฐ แก้วเพ็ชร (2528) และ สุนทร โตรัตน์ (2524) พบว่าความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง เป็นตัวแปรสำคัญในการตัดสินใจที่จะให้บุตรศึกษาต่อหรือไม่ศึกษาต่อ นอกจากนี้ดำรง แสนสิงห์ (2534) ธวัช แก้วอนันต์ (2533) และสวลี อาชาสุวรรณ (2532) ยังพบว่า ความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูงเป็นตัวแปรที่มีอิทธิพลในการจำแนกกลุ่มนักเรียนที่ได้ศึกษาต่อและไม่ได้ศึกษาต่อ

ความต้องการใช้แรงงาน จากการศึกษาของ เทียมจันทร์ จาตุกัญญาประทีป (2523) บุญเพ็ง ธานี (2530) ประเสริฐ แก้วเพ็ชร (2528) และ สวลี อาชาสุวรรณ (2532) ได้ข้อค้นพบสอดคล้องกันว่า การที่ผู้ปกครองมีความต้องการที่จะใช้แรงงานจากบุตรเพื่อให้ช่วยทำงานในครอบครัวหรือหารายได้มาจุนเจือครอบครัวนั้น เป็นสาเหตุสำคัญที่ทำให้ผู้ปกครองไม่ส่งบุตรศึกษาต่อในระดับการศึกษาที่สูงขึ้น

ความต้องการศึกษาต่อของนักเรียน จากการศึกษาของธวัช แก้วอนันต์ (2533) พบว่า ความต้องการศึกษาต่อของนักเรียนเป็นตัวแปรที่มีความสำคัญในการจำแนกกลุ่มของนักเรียนที่ได้ศึกษาต่อและไม่ได้ศึกษาต่อ ส่วนเทียมจันทร์ จาตุกัญญาประทีป (2523) บุญเพ็ง ธานี (2530) ประเสริฐ แก้วเพ็ชร (2528) และ สุนทร โตรัตน์ (2524) พบว่า นักเรียนที่มีความต้องการที่จะศึกษาต่อในระดับการศึกษาที่สูงขึ้นจะมีโอกาสได้ศึกษาต่อมากกว่า นักเรียนที่ไม่ต้องการที่จะศึกษาต่อในระดับที่สูงขึ้น

ระยะทางระหว่างบ้านถึงโรงเรียน จากการศึกษาของ ดำรง แสนสิงห์ (2534) ธวัช แก้วอนันต์ (2533) และ สุนทร โตรัตน์ (2524) พบว่าระยะทางระหว่างบ้านถึงโรงเรียนเป็นปัจจัยที่ส่งผลต่อการศึกษาต่อในระดับที่สูงขึ้นของนักเรียนในชั้นประถมศึกษาปีที่ 6 โดยที่นักเรียนที่มีที่พักอยู่ใกลจากโรงเรียนที่ต้องการศึกษาต่อเป็นระยะทางมากๆจะไม่ศึกษาต่อ ในขณะที่นักเรียนที่มีที่พักอยู่ใกล้กับโรงเรียนมักจะศึกษาต่อ

การคมนาคม จากการศึกษาของ ดำรง แสนสิงห์ (2534) ธวัช แก้วอนันต์ (2533) และ สุนทร โตรัตน์ (2524) สอดคล้องกันว่าการคมนาคมเป็นตัวแปรที่ส่งผลต่อการศึกษาต่อและไม่ศึกษาต่อในระดับมัธยมศึกษาของนักเรียนที่สำเร็จการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 โดยที่การคมนาคมที่สะดวกสบายมีอิทธิพลต่อการศึกษาต่อในระดับที่สูงขึ้นของนักเรียนมากกว่าการคมนาคมที่ค่อนข้างลำบาก

ผลสัมฤทธิ์ทางการเรียน จากการศึกษาของจรูญ พานิชผลินไชย (2521) ดำรง แสนสิงห์ (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) ธวัช แก้วอนันต์ (2533) บุญเพ็ง ธานี (2530) ประยงค์ ชูน้อย (2521) ประเสริฐ แก้วเพชร (2528) สวลี อาชาสุวรรณ (2532) สุนทร โตรัตน์ (2524) และ Worthington and Grant (1971) ได้ข้อค้นพบสอดคล้องกันว่า ผลสัมฤทธิ์ทางการเรียนหรือระดับผลการเรียนของนักเรียนเป็นตัวแปรที่มีอิทธิพลต่อการได้เข้าศึกษาต่อในระดับการศึกษาที่สูงขึ้นของนักเรียน

จากงานวิจัยทั้ง 13 เรื่องที่ศึกษาตัวแปรที่สัมพันธ์กับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนทั้งหมดนี้ ผู้วิจัยพบว่า ตัวแปรตามในการวิจัยเหล่านี้ไม่ได้วัดจากพฤติกรรมของนักเรียนที่ศึกษาต่อและไม่ศึกษาต่อโดยตรง แต่วัดจากความคาดหวังหรือโอกาสในการศึกษาต่อเป็นส่วนใหญ่ ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงไม่ได้ศึกษาพฤติกรรมการศึกษาต่อหรือไม่ศึกษาต่อของนักเรียนโดยตรงเช่นกัน หากแต่เลือกศึกษาความคาดหวังในการศึกษาต่อแทนพฤติกรรมด้วย

เหตุผล 2 ประการ ประการแรก จากทฤษฎีทางจิตวิทยาที่ว่า พฤติกรรมเป็นผลจากการที่คนเราเลือกปฏิกิริยาตอบสนองกับสิ่งเร้าตามกระบวนการจูงใจ (motivation process) ประกอบด้วย ความต้องการและแรงขับ (needs and drives) หรือแรงขับและแรงจูงใจ (drives and motives) เพื่อนำไปสู่การแสดงออกทางพฤติกรรมที่มีเป้าหมาย (motivated behavior) และที่เป็นเป้าหมาย (goals) โดยที่กระบวนการจูงใจเป็นผลมาจากความคาดหวัง (expectation) นั่นเอง (Bandura, 1969, 1977, 1986; Luthan, 1985, 1988) ดังนั้น การศึกษาความคาดหวังย่อมได้ผลใกล้เคียงกับการศึกษาพฤติกรรม

นอกจากนี้ แนวคิดของกลุ่มทฤษฎีการเรียนรู้ทางสังคม (social learning theory) ได้กล่าวเอาไว้ว่า สิ่งที่ทำให้มนุษย์เกิดการเรียนรู้ได้นั้นคือ ความสัมพันธ์ระหว่างสิ่งต่างๆ ไม่ว่าจะเป็นการเรียนรู้ที่เกิดจากความสัมพันธ์ระหว่างเหตุการณ์กับเหตุการณ์ หรือการเรียนรู้ที่เกิดจากความสัมพันธ์ระหว่างพฤติกรรมกับผลที่เกิดขึ้นจากพฤติกรรมนั้น มนุษย์จึงเกิดการเรียนรู้ว่า เมื่อมีเหตุการณ์หนึ่งเกิดขึ้นแล้วจะต้องมีเหตุการณ์หนึ่งเกิดตามขึ้นมา ซึ่งสิ่งนั้นก็คือความคาดหวัง และความคาดหวังนี้จะทำให้มนุษย์ตัดสินใจทำหรือไม่ทำพฤติกรรมหนึ่งเพื่อให้เกิดผลตามที่ปรารถนา ดังนั้นจึงอาจกล่าวได้ว่า สิ่งที่กำหนดพฤติกรรมของมนุษย์คือความรู้สึกนึกคิดที่เป็นความคาดหวังและการตัดสินใจกระทำตามความคาดหวัง และพฤติกรรมของมนุษย์เป็นผลของความคาดหวัง (Bandura, 1978; Greenough, Blake and Wallace, 1987; David and Newstorm, 1985) ผู้วิจัยจึงเชื่อว่า พฤติกรรมที่จะศึกษาต่อหรือไม่ศึกษาต่อของนักเรียนหลังจากที่จบการศึกษาภาคบังคับนั้นเป็นผลมาจากความคาดหวัง

เหตุผลประการที่สองที่ผู้วิจัยเลือกศึกษาความคาดหวังในการศึกษาต่อของนักเรียน เนื่องจากความเชื่อที่ว่า หากได้มีการศึกษาและทราบความคาดหวังในการศึกษาของนักเรียน ครูและผู้ที่เกี่ยวข้องจะมีโอกาสเปลี่ยนความคาดหวังของนักเรียนจากไม่ศึกษาต่อเป็นศึกษาต่อได้ทันเหตุการณ์กว่าการวิจัยที่ศึกษาพฤติกรรมโดยตรง

ตัวแปรที่เป็นสาเหตุในโมเดลความสัมพันธ์เชิงสาเหตุของความคาดหวังในการศึกษาต่อของนักเรียนครั้งนี้เป็นตัวแปรที่ผู้วิจัยคาดว่าน่าจะมีอิทธิพลทั้งทางตรงและทางอ้อมต่อความคาดหวังในการศึกษาต่อโดยพิจารณาคัดเลือกจากตัวแปรที่มีความสัมพันธ์กับการศึกษาต่อของนักเรียนสูงๆเท่านั้น ซึ่งจากตารางที่ 4 จะพบว่าตัวแปรที่มีความสัมพันธ์กับการศึกษาต่อที่มีความถี่สูงได้แก่ รายได้ของผู้ปกครอง (13) ระดับการศึกษาของผู้ปกครอง (12) อาชีพของผู้ปกครอง (11) ผลสัมฤทธิ์ทางการเรียน (10) และความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการ

ศึกษาชั้นสูง (8) จากเหตุผลดังกล่าวผู้วิจัยจึงพิจารณาคัดเลือกตัวแปรอิสระเข้าสู่โมเดลความสัมพันธ์เชิงสาเหตุที่เป็นสมมุติฐานการวิจัยเพียง 5 ตัวเท่านั้น

การนำเสนอแนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้เสนอตามลำดับของตัวแปรในโมเดลความสัมพันธ์เชิงสาเหตุผลตามแผนภาพที่ 15 เรียงลำดับดังนี้

อาชีพของผู้ปกครอง

อาชีพของผู้ปกครองเป็นตัวแปรสำคัญที่มีอิทธิพลต่อการศึกษาต่อในระดับที่สูงขึ้นของนักเรียน เนื่องจากการศึกษาเป็นรากฐานสำคัญในการประกอบอาชีพของบุคคล บุคคลที่จบการศึกษาในระดับสูงมักจะมีโอกาสในการประกอบอาชีพที่ดีกว่าบุคคลที่จบการศึกษาในระดับการศึกษาที่ต่ำกว่า ดังนั้นผู้ที่ประกอบอาชีพที่ดีจึงมักจะเห็นความสำคัญของการศึกษาเป็นอย่างยิ่ง จากการศึกษาของ ดำรง แสนสิงห์ (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) บุญเพ็ง ธานี (2530) ประยงค์ ชูน้อย (2521) สวลี อาชาสุวรรณ (2532) และ สุนทร โตรัตน์ (2524) ได้ข้อค้นพบว่า อาชีพของผู้ปกครองจะส่งผลต่อโอกาสในการศึกษาต่อในระดับมัธยมศึกษาตอนต้นของนักเรียนที่สำเร็จการศึกษาระดับชั้นประถมศึกษาปีที่ 6 ซึ่งสอดคล้องกับ Blake (1985) Waite, Rindfuss and tray (1986) และ Worthington and Grant (1971) ที่พบว่า อาชีพของผู้ปกครองเป็นตัวแปรที่สัมพันธ์กับการศึกษาต่อของนักเรียน กล่าวคือผู้ปกครองของนักเรียนที่ประกอบอาชีพที่ดีมีรายได้สูงมีผลต่อโอกาสที่บุตรจะได้รับการศึกษาขั้นสูงกว่าการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 และ จรูญ พานิชผลินไชย (2521) และ ธวัช แก้วอนันต์ (2533) พบว่า อาชีพของผู้ปกครองมีผลต่อรายได้ของผู้ปกครองอีกด้วย เนื่องจากงานวิจัยพบว่าผู้ปกครองที่ประกอบอาชีพที่ดีย่อมมีรายได้มากกว่าผู้ปกครองที่ประกอบอาชีพที่ด้อยกว่า นอกจากนี้จากการศึกษาของวราภรณ์ วิหคโต (2536) ยังพบว่าอาชีพของผู้ปกครองส่งผลต่อผลสัมฤทธิ์ทางการเรียนอีกด้วย กล่าวคือ นักเรียนที่มีผู้ปกครองประกอบอาชีพที่ดีส่วนใหญ่จะมีผลการเรียนอยู่ในระดับที่ดีกว่านักเรียนที่มีผู้ปกครองประกอบอาชีพที่ด้อยกว่า

ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงคาดว่า อาชีพของผู้ปกครองน่าจะมีอิทธิพลทางตรงต่อความคาดหวังในการศึกษาต่อของนักเรียนและน่าจะมีอิทธิพลทางอ้อมโดยผ่านทางรายได้ของผู้ปกครอง และผลสัมฤทธิ์ทางการเรียน

ระดับการศึกษาของผู้ปกครอง

ผู้ปกครองที่มีระดับการศึกษาสูงมักจะเห็นความสำคัญของการศึกษาของบุตร ดังรายงานการวิจัยของ Jamison and Lockheed (1987) และ Rehberg and Westby (1967) ที่พบว่าการสำเร็จการศึกษาและระดับการศึกษาของบิดาเป็นปัจจัยที่กำหนดการศึกษาต่อของบุตรสอดคล้องกับงานวิจัยของ Chemichovsky (1985) เช่นกันว่า ระดับการศึกษาของผู้ปกครองมีผลต่อการได้รับการศึกษาของบุตรมากที่สุด ส่วน Waite, Rindfuss and Tray (1986) ได้ศึกษาถึงความมุ่งหวังของมารดาเกี่ยวกับการศึกษาของบุตรพบว่าระดับการศึกษาของมารดามีผลต่อความมุ่งหวังทางการศึกษาของมารดาที่จะให้บุตรได้รับการศึกษาขั้นสูง นอกจากนี้ระดับการศึกษาของผู้ปกครองยังมีผลต่อรายได้ของผู้ปกครองด้วย ดังที่ Baldridge (1969) และ Blau and Duncan (1967) อ้างใน นฤมล บุลนิม (2529) พบว่า ผู้ที่มีระดับการศึกษาสูงมักจะมีรายได้หรือค่าตอบแทนในวิชาชีพสูงกว่าผู้ที่มีระดับการศึกษาต่ำ และจากการศึกษาของวรารภรณ์ วิหคโต (2536) พบว่าระดับการศึกษาของผู้ปกครองส่งผลต่อผลสัมฤทธิ์ทางการเรียนด้วย

ด้วยเหตุผลดังกล่าวผู้วิจัยจึงคาดว่า ระดับการศึกษาของผู้ปกครองน่าจะมีอิทธิพลทางตรงต่อความคาดหวังในการศึกษาต่อของนักเรียนและน่าจะมีอิทธิพลทางอ้อมโดยผ่านทางรายได้ความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง และผลสัมฤทธิ์ทางการเรียน

รายได้ของผู้ปกครอง

การศึกษาในระดับประถมศึกษานั้นเป็นการจัดแบบให้เปล่าจากรัฐ แต่การศึกษาต่อในระดับมัธยมศึกษานั้นผู้ปกครองของนักเรียนต้องรับภาระค่าใช้จ่ายในการศึกษาเองทั้งหมด ทำให้นักเรียนที่สำเร็จการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 แล้วไม่มีโอกาสได้ศึกษาต่อในระดับที่สูงขึ้น จึงกล่าวได้ว่า รายได้ของผู้ปกครองมีความสำคัญต่อการศึกษาต่อเป็นอย่างยิ่ง จากการศึกษาของเทียมจันทร์ จาตุกัญญาประทีป (2523) พบว่า รายได้ของผู้ปกครองเป็นตัวแปรตัวหนึ่งที่ทำให้นักเรียนมีโอกาสได้ศึกษาต่อและไม่ได้ศึกษาต่อ โดยนักเรียนที่ได้ศึกษาต่อนั้นผู้ปกครองจะมีรายได้สูงกว่านักเรียนที่ไม่ได้ศึกษาต่อซึ่งสอดคล้องกับผลการวิจัยของ บุญเพ็ง ธานี (2530) และ ประเสริฐ แก้วเพชร (2528)

จากผลการศึกษาของ Worthington and Grant (1971) และ Wright and Bean (1974) สอดคล้องกันว่า นักเรียนที่ผู้ปกครองมีฐานะทางเศรษฐกิจดีจะมีผลสัมฤทธิ์ทางการเรียนสูงกว่า

นักเรียนที่มีผู้ปกครองที่มีฐานะทางเศรษฐกิจต่ำ และมีความมุ่งหวังที่จะให้บุตรได้รับการศึกษาขั้นสูงมากกว่านักเรียนที่มีผู้ปกครองมีฐานะทางเศรษฐกิจต่ำ สอดคล้องกับผลการศึกษาของคณะกรรมการการศึกษาแห่งชาติ (2526) ที่พบว่า ฐานะทางเศรษฐกิจของครอบครัวมีผลต่อความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูงอีกด้วย กล่าวคือ ผู้ปกครองที่มีรายได้สูงก็มักจะมีความมุ่งหวังให้บุตรได้รับการศึกษาสูงๆขึ้นไปและฐานะทางเศรษฐกิจจะมีผลต่อผลสัมฤทธิ์ทางการเรียน แต่จากการศึกษาของพิศเพลิน เขียวหวาน (2520) สุรศักดิ์ อมรรัตนศักดิ์ (2521) และ ประนอม ทวีกาญจน์ (2526) อ้างถึงใน นฤมล บุลนิม (2529) พบว่า ฐานะทางเศรษฐกิจของผู้ปกครองมีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนน้อยมากหรือเกือบไม่มีเลย

ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงเชื่อว่า รายได้ของผู้ปกครองน่าจะมีอิทธิพลทางตรงต่อความคาดหวังในการศึกษาต่อของนักเรียนและน่าจะมีอิทธิพลทางอ้อมโดยผ่านทางความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง แต่ไม่มีอิทธิพลทางอ้อมผ่านทางผลสัมฤทธิ์ทางการเรียน

ความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง

การที่บุตรจะมีโอกาสในการศึกษาต่อมากหรือน้อยเพียงใดนั้นขึ้นอยู่กับผู้ปกครองเป็นสำคัญเนื่องจากผู้ปกครองเป็นผู้ที่คอยรับผิดชอบดูแลในเรื่องต่างๆของบุตร หากผู้ปกครองเห็นความสำคัญของการศึกษาโอกาสที่บุตรจะได้รับการศึกษาย่อมมีมาก และหากผู้ปกครองมีความมุ่งหวังที่จะให้บุตรได้รับการศึกษาขั้นสูงจะมีผลต่อการศึกษาต่อของบุตรด้วย ดังผลการศึกษาของเทียมจันทร์ จาตุกัญญาประทีป (2523) สวลี อาชาสุวรรณ (2532) และ สุนทร โคตรัตน์ (2524) ที่พบว่า ผู้ปกครองของนักเรียนที่ได้ศึกษาต่อจะมีระดับความต้องการที่จะให้บุตรศึกษาต่อมากกว่าผู้ปกครองของนักเรียนที่ไม่ได้ศึกษาต่อ สอดคล้องกับงานวิจัยของ Kandel and Lesser (1969) ที่ว่าความมุ่งหวังของผู้ปกครองมีความสัมพันธ์โดยตรงต่อการศึกษาของบุตร และจากงานวิจัยของ Sewell and Shah (1968) อ้างถึงใน พราวมาศ นนทยาธร (2520) ได้ศึกษาความสัมพันธ์ระหว่างระดับการศึกษาของบิดามารดาที่มีต่อความมุ่งหวังที่จะให้บุตรได้รับการศึกษาขั้นสูงและผลสัมฤทธิ์ทางการเรียนของบุตร พบว่าระดับการศึกษาของบิดามารดามีผลต่อความมุ่งหวังที่จะให้บุตรได้รับการศึกษาขั้นสูง กล่าวคือ ยิ่งบิดามารดามีระดับการศึกษาสูงขึ้นเท่าใดบุตรก็จะได้รับความมุ่งหวังจากบิดามารดาให้ได้รับการศึกษาสูงยิ่งขึ้น บุตรจะได้รับการศึกษาต่อในขั้นที่สูงขึ้นและได้รับผลสัมฤทธิ์ทางการเรียนสูง

ด้วยเหตุผลดังกล่าวผู้วิจัยจึงคาดว่าความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูงน่าจะมีอิทธิพลทางตรงต่อคาดหวังในการศึกษาต่อของนักเรียนและน่าจะมีอิทธิพลทางอ้อมโดยผ่านทางผลสัมฤทธิ์ทางการเรียนอีกด้วย

ผลสัมฤทธิ์ทางการเรียน

จากการศึกษาของ จรูญ พานิชผลินไชย (2521) ดำรง แสนสิงห์ (2534) เทียมจันทร์ จาตุกัญญาประทีป (2523) ชวัช แก้วอนันต์ (2533) บุญเพ็ง ชานี (2530) ประยงค์ ชูน้อย (2521) ประเสริฐ แก้วเพ็ชร (2528) สวลี อาชาสุวรรณ (2532) สุนทร โคตรัตน์ (2524) และ Worthington and Grant (1971) สอดคล้องกันว่า นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนในระดับชั้นประถมศึกษาดีจะมีโอกาสได้ศึกษาต่อ ในระดับมัธยมศึกษาสูงกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนในระดับชั้นประถมศึกษาปานกลาง หรือ ต่ำ นอกจากนี้ เทียมจันทร์ จาตุกัญญาประทีป (2523) ยังพบอีกว่า นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูงจะมีความมุ่งหวังในการศึกษาต่อสูงกว่านักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำ

ผู้วิจัยจึงคาดว่าผลสัมฤทธิ์ทางการเรียน น่าจะมีอิทธิพลทางตรงต่อคาดหวังในการศึกษาต่อของนักเรียน

ตอนที่ 3.3 งานวิจัยเกี่ยวกับการประยุกต์ใช้การวิเคราะห์การถดถอยโลจิสติกร่วมกับการวิเคราะห์อิทธิพล

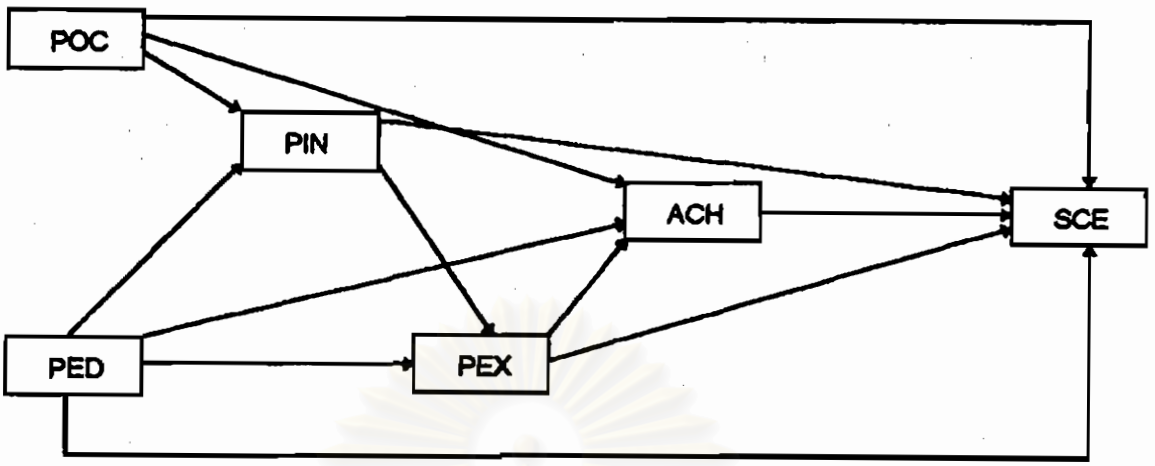
งานวิจัยเกี่ยวกับ การประยุกต์ใช้การวิเคราะห์การถดถอยโลจิสติกร่วมกับการวิเคราะห์อิทธิพลเท่าที่ผ่านมาในอดีต พบว่ามี 1 เรื่องคือ ความสัมพันธ์เชิงสาเหตุของตัวแปรที่มีอิทธิพลต่อการเลือกสถานศึกษาสำหรับบุตรของผู้ปกครองนักเรียนชั้นมัธยมศึกษาตอนต้นในภาคกลาง: การวิเคราะห์เส้นทาง (วรรณี แกมเกตุ, 2536) มีวัตถุประสงค์เพื่อสร้างและทดสอบตัวแบบ (model) ความสัมพันธ์เชิงสาเหตุของตัวแปรเกี่ยวกับภูมิหลังของผู้ปกครองและตัวแปรเกี่ยวกับนักเรียนที่มีอิทธิพลต่อการเลือกสถานศึกษาสำหรับบุตรของผู้ปกครองนักเรียนชั้นมัธยมศึกษาตอนต้นในภาคกลาง โดยใช้การวิเคราะห์เส้นทางหรือการวิเคราะห์อิทธิพล (path analysis) แบบดั้งเดิม กล่าวคือ ในการหาค่า R^2 (ordinary squared multiple correlation coefficient) ค่าสัมประสิทธิ์การพยากรณ์ และค่าสถิติที ของตัวแบบความสัมพันธ์แบบเต็มรูปโดยการวิเคราะห์การถดถอยนั้น หากพบว่า มีการวิเคราะห์การถดถอยที่มีตัวแปรตามเป็นค่าไม่ต่อเนื่อง จะใช้การวิเคราะห์การ

ถดถอยโลจิสติกในการคำนวณค่าสถิติดังกล่าว ซึ่งแตกต่างจากการวิจัยในครั้งนี้ เนื่องจากการวิจัยในครั้งนี้ใช้การวิเคราะห์อิทธิพลด้วยลิสเรล มิได้ใช้การวิเคราะห์อิทธิพลแบบดั้งเดิม กล่าวคือ ผู้วิจัยใช้การวิเคราะห์การถดถอยโลจิสติกในการตรวจสอบความกลมกลืนของโมเดลการถดถอยโลจิสติกกับข้อมูลเชิงประจักษ์ แล้วนำสมการถดถอยโลจิสติกมาใช้เพื่อพยากรณ์ความน่าจะเป็นของความคาดหวังในการศึกษาต่อของนักเรียน ผลที่ได้จากการวิเคราะห์การถดถอยโลจิสติกคือ ค่าขนาดอิทธิพลของตัวแปรอิสระหรือตัวแปรพยากรณ์ที่มีต่อออกการิทึมของอัตราส่วนแต้มต่อความคาดหวังในการศึกษาต่อของนักเรียนผู้วิจัยจะนำค่าอิทธิพลดังกล่าวไปประยุกต์ใช้กับการวิเคราะห์อิทธิพลเพื่อตรวจสอบความตรงของโมเดลความสัมพันธ์เชิงสาเหตุของตัวแปรต่างๆที่สร้างขึ้นกับข้อมูลเชิงประจักษ์และศึกษาเปรียบเทียบค่าผลรวมอิทธิพล อิทธิพลทางตรง และอิทธิพลทางอ้อมของตัวแปรที่เป็นสาเหตุในโมเดลที่มีผลต่อความคาดหวังในการศึกษาต่อของนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งผลจากการวิเคราะห์ในครั้งนี้น่าจะตอบคำถามวิจัยได้อย่างครบถ้วนและลึกซึ้งมากกว่าการประยุกต์ใช้การวิเคราะห์การถดถอยโลจิสติกร่วมกับการวิเคราะห์อิทธิพลแบบดั้งเดิม

สมมุติฐานของการวิจัย

การวิจัยครั้งนี้ผู้วิจัยตั้งสมมุติฐานการวิจัยไว้ 3 ข้อ ดังนี้

1.จากแนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้นำมาเป็นเหตุผลสนับสนุนการเชื่อมโยงเส้นทางอิทธิพลระหว่างตัวแปรในโมเดลเพื่อสร้างสมมุติฐานความสัมพันธ์เชิงสาเหตุของความคาดหวังในการศึกษาต่อของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยพิจารณาจากตัวแปรตามคือ ความคาดหวังในการศึกษาต่อก่อน ต่อจากนั้นจึงเป็นการพิจารณาตัวแปรที่เกี่ยวข้องย้อนกลับไปยังตัวแปรสาเหตุทีละตัวแปร ซึ่งได้โมเดลสมมุติฐานความสัมพันธ์เชิงสาเหตุของความคาดหวังในการศึกษาต่อของนักเรียนชั้นประถมศึกษาปีที่ 6 ดังภาพที่ 14



ภาพที่ 14 โมเดลสมมุติฐานความสัมพันธ์เชิงสาเหตุของความคาดหวังในการศึกษาต่อของนักเรียน

ตัวแปรภายนอก (exogenous variables)

POC = อาชีพของผู้ปกครอง (parents ' occupation)

PED = ระดับการศึกษาของผู้ปกครอง (parents ' education)

ตัวแปรภายใน (endogenous variables)

PIN = รายได้ของผู้ปกครอง (parents ' income)

PEX = ความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง (parents ' expectation)

ACH = ผลสัมฤทธิ์ทางการเรียน (achievement)

SCE = ความคาดหวังในการศึกษาต่อ (students ' continuing education expectation)

2. การศึกษาเปรียบเทียบความเหมาะสมของวิธีวิเคราะห์อิทธิพลระหว่าง การใช้ค่าดอการิทึมของอัตราส่วนแต้มต่อที่ได้จากการวิเคราะห์การถดถอยโลจิสติกเป็นตัวแปรผลในโมเดล กับการใช้ค่าพยากรณ์ที่ได้จากการวิเคราะห์การถดถอยพหุคูณเป็นตัวแปรผลในโมเดล ผู้วิจัยตั้งสมมุติฐานว่า การวิเคราะห์อิทธิพลโดยใช้ค่าดอการิทึมของอัตราส่วนแต้มต่อที่ได้จากการวิเคราะห์การถดถอยโลจิสติกเป็นตัวแปรผลในโมเดลมีความเหมาะสมมากกว่าการใช้ค่าพยากรณ์ที่ได้จากการวิเคราะห์การถดถอยพหุคูณเป็นตัวแปรผลในโมเดล เนื่องจากความแตกต่างในการประมาณค่าพารามิเตอร์ระหว่างเทคนิควิธีวิเคราะห์ทั้งสอง กล่าวคือ ในการวิเคราะห์การถดถอย

โลจิสติก จะประมาณค่าพารามิเตอร์ในโมเดลด้วยวิธีไลค์ลิฮู้ดสูงสุด (maximum likelihood) อันเป็นการคำนวณทวนซ้ำ (iterative algorithm) เพื่อให้ได้ค่าประมาณของพารามิเตอร์ที่ใกล้เคียงกับข้อมูลเชิงประจักษ์มากที่สุด แต่ในการประมาณค่าพารามิเตอร์ของการวิเคราะห์การถดถอยพหุคูณจะประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุด (ordinary least squares) โดยคัดเลือกสัมประสิทธิ์การถดถอยที่ทำให้ผลรวมกำลังสองของความคลาดเคลื่อนในการพยากรณ์มีค่าน้อยที่สุด โดยที่ความคลาดเคลื่อนคำนวณจากผลต่างระหว่างคะแนนที่วัดได้กับคะแนนที่เป็นค่าพยากรณ์ ซึ่งค่าที่ได้จากการวิเคราะห์การถดถอยโลจิสติกน่าจะมีความเหมาะสมกว่า เพราะการประมาณค่าพารามิเตอร์ในการวิเคราะห์โมเดลลิสเรลนั้นโดยส่วนใหญ่แล้ว จะประมาณค่าพารามิเตอร์ในโมเดลด้วยวิธีไลค์ลิฮู้ดสูงสุดเช่นเดียวกันกับการประมาณค่าพารามิเตอร์ในโมเดลการวิเคราะห์การถดถอยโลจิสติก (นงลักษณ์ วิรัชชัย, 2538; Hosmer and Lemmeshow, 1989 and Norusis, 1990)

3. จากการวิจัยของดำรง แสนสิงห์ (2534) และธวัช แก้วอนันต์ (2533) เกี่ยวกับการวิเคราะห์จำแนก (discriminant analysis) ปัจจัยที่เกี่ยวข้องกับการศึกษาต่อและไม่ศึกษาต่อของนักเรียนพบว่าตัวแปรที่สามารถจำแนกกลุ่มได้คือ อาชีพของผู้ปกครอง ระดับการศึกษาของผู้ปกครอง รายได้ของผู้ปกครอง จำนวนบุตร จำนวนบุตรที่ได้รับการศึกษาขั้นสูง ความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง ระยะทางระหว่างบ้านถึงโรงเรียน การคมนาคม ความต้องการศึกษาต่อของนักเรียน และผลสัมฤทธิ์ทางการเรียน และจากตารางที่ 4 ซึ่งแสดงตัวแปรที่สัมพันธ์กับการศึกษาต่อของนักเรียนที่สำเร็จการศึกษาในระดับชั้นประถมศึกษาปีที่ 6 ในรูปของความถี่ ดังนี้ รายได้ของผู้ปกครอง (13) ระดับการศึกษาของผู้ปกครอง (12) อาชีพของผู้ปกครอง (11) ผลสัมฤทธิ์ทางการเรียน (10) และความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูง (8)

ดังนั้นผู้วิจัยจึงตั้งสมมุติฐานของการวิจัยข้อที่ 3 ว่า ตัวแปรที่เป็นสาเหตุแต่ละตัวในโมเดลความสัมพันธ์เชิงสาเหตุที่จะสามารถพยากรณ์ความน่าจะเป็นของการศึกษาต่อที่คาดหวังของนักเรียนชั้นประถมศึกษาปีที่ 6 ในรูปของอัตราส่วนแต้มต่อได้ดีที่สุด คือ รายได้ของผู้ปกครอง รองลงมาได้แก่ ระดับการศึกษาของผู้ปกครอง อาชีพของผู้ปกครอง ผลสัมฤทธิ์ทางการเรียน และความมุ่งหวังของผู้ปกครองที่จะให้บุตรได้รับการศึกษาขั้นสูงสุด ตามลำดับ