

บทที่ 8 การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Analysis)

การวิเคราะห์การถดถอย เป็นวิธีการทางสถิติอย่างหนึ่ง ที่ใช้ในการตรวจสอบลักษณะของความสัมพันธ์ระหว่างตัวแปร ตั้งแต่ 2 ตัวขึ้นไป โดยแบ่งเป็นตัวแปรอิสระ (**Independent variable**) และตัวแปรตาม (**Dependent variable**)

ผลของการศึกษาจะให้ทราบถึง

- (1) ขนาดของความสัมพันธ์ระหว่างตัวแปรอิสระ ที่มีต่อตัวแปรตาม และ
- (2) แบบจำลองความสัมพันธ์ระหว่างตัวแปรอิสระ และตัวแปรตาม

ในการวิเคราะห์การถดถอย มักเรียกตัวแปรอิสระ ว่า ตัวทำนาย (**predictor**) หรือตัวแปรกระตุ้น (**stimulus variable**) ส่วนตัวแปรตาม มักเรียกว่า ตัวแปรตอบสนอง (**response variable**) หรือตัวแปรเกณฑ์ (**criterion variable**)

วัตถุประสงค์ของการประยุกต์ใช้การวิเคราะห์การถดถอย

1. ต้องการศึกษาค่าความสัมพันธ์ระหว่างตัวแปรอิสระ และตัวแปรตามว่ามีความสัมพันธ์กันหรือไม่ และมีความสัมพันธ์กันอย่างไร ในกลุ่มตัวแปรอิสระหลายๆ ตัวนั้น ตัวใดบ้างที่มีความสัมพันธ์กับตัวแปรตาม ตัวใดมีความสัมพันธ์สูง ตัวใดมีความสัมพันธ์น้อย หรือไม่มีความสัมพันธ์ เพื่อที่จะสามารถคาดการณ์ได้ว่าตัวแปรอิสระตัวใดมีอิทธิพลต่อตัวแปรตามมากที่สุด เช่น ศึกษาความสัมพันธ์ระหว่างน้ำหนักทารกแรกเกิด กับอายุ น้ำหนัก และ ส่วนสูงของมารดา
2. ต้องการสร้างแบบจำลองเพื่อทำนายตัวแปรตาม โดยรูปแบบจำลองดังกล่าวอยู่ในลักษณะสมการทางคณิตศาสตร์ เช่น ศึกษาปริมาณการใช้จ่ายที่ส่งผลต่อความดันโลหิต
3. ต้องการทราบความสัมพันธ์ระหว่างตัวแปรอิสระบางตัวที่มีผลต่อตัวแปรตาม โดยควบคุมอิทธิพลของตัวแปรอิสระตัวอื่นๆ ใ้คงที่ เช่น ศึกษาอิทธิพลของความวิตกกังวลที่มีต่อประสิทธิภาพการทำงาน เมื่อควบคุมระยะเวลาในการทำงานติดต่อกันให้คงที่
4. ต้องการสร้างแบบจำลองทางคณิตศาสตร์ที่ดีที่สุด เพื่อนำไปใช้ในการทำนายตัวแปรตาม โดยอาจมีแบบจำลองจำนวนมากให้ตัดสินใจ
5. ต้องการทราบว่าแบบจำลองที่พัฒนาขึ้นมาสำหรับทำนายนั้นจะมีประสิทธิภาพในการทำนายได้อย่างคงเส้นคงวาหรือไม่ เมื่อนำไปใช้กับกลุ่มเป้าหมายต่างๆ กัน

ชนิดของการวิเคราะห์การถดถอย

การวิเคราะห์การถดถอย มีหลายชนิด ขึ้นกับลักษณะของตัวแปรตาม รูปแบบความสัมพันธ์ และการกำหนดตัวแปรอิสระ (ตัวแปรต้น) ซึ่งโดยทั่วไปแบ่งการวิเคราะห์การถดถอยได้เป็น 2 ประเภท คือ

- การวิเคราะห์การถดถอยเชิงเส้น (**Linear regression analysis**) เป็นการวิเคราะห์การถดถอยที่ตัวแปรอิสระส่วนใหญ่เป็นตัวแปรเชิงปริมาณ ส่วนตัวแปรตามเป็นจะต้องเป็นตัวแปรเชิงปริมาณเท่านั้น รูปแบบของความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม สามารถแทนได้ด้วยสมการทางคณิตศาสตร์ที่เป็นเชิงเส้น (**Linear model**)
- การวิเคราะห์การถดถอยแบบไม่เป็นเชิงเส้น (**Non linear regression**) เป็นการวิเคราะห์การถดถอย ที่รูปแบบของความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม สามารถแทนได้ด้วยสมการทางคณิตศาสตร์ที่ไม่เป็นเชิงเส้น (**non - Linear model**)

สำหรับเนื้อหาในบทนี้จะกล่าวถึงเฉพาะการวิเคราะห์การถดถอยเชิงเส้น เท่านั้น

การวิเคราะห์การถดถอยเชิงเส้น มี 2 แบบ คือ

การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย จะประกอบด้วยตัวแปรตาม 1 ตัว และตัวแปรอิสระ เพียง 1 ตัว การวิเคราะห์เป็นการหาความสัมพันธ์ของตัวแปรทั้งสอง และสร้างรูปแบบสมการทางคณิตศาสตร์ที่เป็นการพยากรณ์ค่าของตัวแปรตาม เช่น

การพยากรณ์ระดับ **Carbon monoxide** ในผู้สูบบุหรี่ เมื่อทราบปริมาณการสูบบุหรี่ต่อวัน

การพยากรณ์น้ำหนักของทารก เมื่อทราบอายุของมารดา

การพยากรณ์ผลการสอบปลายภาค เมื่อทราบผลการสอบกลางภาค เป็นต้น

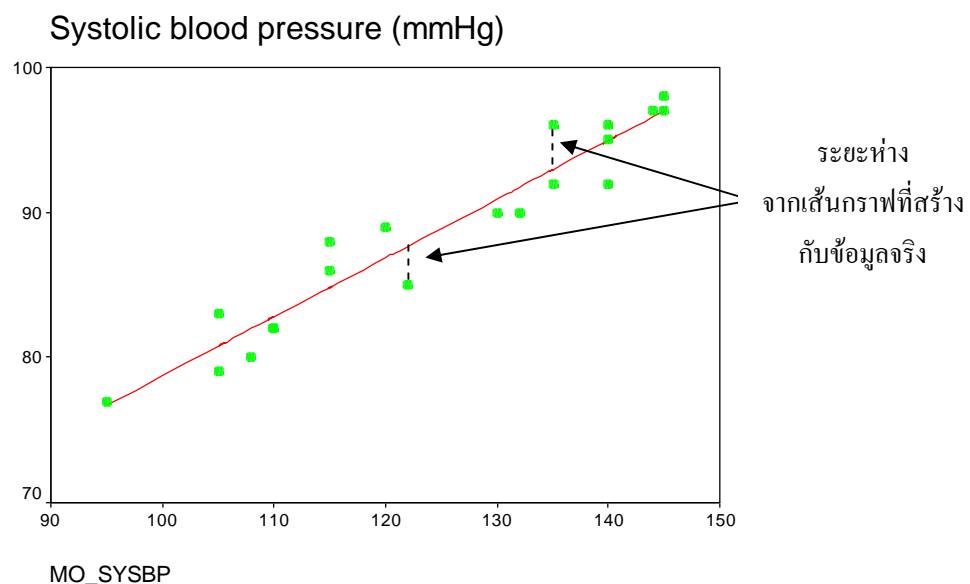
การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ (Multiple Linear Regression)

จะประกอบด้วยตัวแปรตาม 1 ตัว และ ตัวแปรอิสระตั้งแต่ 2 ตัวขึ้นไป การวิเคราะห์เป็นการหาขนาดของความสัมพันธ์ และสร้างรูปแบบสมการทางคณิตศาสตร์ที่เป็นการพยากรณ์ค่าของตัวแปรตาม โดยใช้ตัวแปรอิสระที่ศึกษา เช่น ศึกษาความสัมพันธ์ระหว่างอายุ เพศ ปัญหาในการทำงาน ความขัดแย้งในครอบครัว กับความรู้สึกเก็บกด ของผู้ป่วยในโรงพยาบาลแห่งหนึ่ง เป็นต้น

แนวคิดของการวิเคราะห์การถดถอยเชิงเส้น

ในการวิเคราะห์การถดถอยเชิงเส้น (กรณีการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย) จะเป็นการนำข้อมูลจากตัวแปรที่ทำการศึกษามาวิเคราะห์หาความสัมพันธ์ที่สามารถบอกแนวโน้มของความสัมพันธ์โดยใช้แผนภาพเส้นตรงแทนได้ และจะทำการหาเส้นตรงที่ดีที่สุดเพื่อเป็นตัวแทนของรูปแบบความสัมพันธ์ของตัวแปรที่ศึกษา

เส้นตรงที่ดีที่สุดจะมีเพียงเส้นเดียวโดยถือหลักการว่าจะต้องมีผลรวมของระยะห่างกำลังสอง จากเส้นกราฟถึงทุกๆจุดนั้น มีค่าน้อยที่สุด เราเรียกหลักการนี้ว่า วิธีกำลังสองน้อยที่สุด (Method of Least Squares)



จากเส้นตรงดังกล่าว ใช้กระบวนการทางสถิติเพื่อหาค่าคงที่และสัมประสิทธิ์สมการสร้างเป็นแบบจำลองในรูปแบบสมการทางคณิตศาสตร์ เรียก สมการถดถอยเชิงเส้น หรือสมการพยากรณ์ หลังจากได้แบบจำลองแล้ว จึงทำการตรวจสอบความสอดคล้องของแบบจำลอง เพื่อดู

ว่าแบบจำลองที่สร้างขึ้น มีความสอดคล้องและเหมาะสมกับข้อมูลหรือไม่ โดยมีการทดสอบทางสถิติดังต่อไปนี้

1. การทดสอบความเหมาะสมของโมเดล (เป็นการตรวจสอบว่าตัวแปรอิสระ และตัวแปรตามมีความสัมพันธ์เชิงเส้นกันหรือไม่) จะใช้สถิติทดสอบ ANOVA
2. การทดสอบค่าคงที่ และค่าสัมประสิทธิ์ในสมการถดถอย ทีละตัวโดยใช้สถิติทดสอบ t
3. พิจารณาค่าสัมประสิทธิ์สหสัมพันธ์แบบพหุ (Multiple R) และค่าความคลาดเคลื่อนในการพยากรณ์ (Standard Error of Estimate)

กระบวนการดังกล่าวทั้งหมดตั้งแต่เริ่มต้นจนจบจะใช้การคำนวณและการทดสอบนัยสำคัญทางสถิติ ด้วยการคำนวณตัวเลขเอง หรือสามารถใช้โปรแกรมวิเคราะห์ข้อมูลทางสถิติทำการวิเคราะห์ให้ก็ได้

ข้อตกลงเบื้องต้นในการใช้การวิเคราะห์การถดถอยเชิงเส้น

1. ตัวแปรอิสระ และตัวแปรตาม มีมาตรวัดเป็น **Interval** ขึ้นไป (อนุญาตให้ตัวแปรอิสระบางตัวมีมาตรวัดเป็น **Nominal** หรือ **Ordinal** ได้บ้าง โดยจะต้องทำการเปลี่ยนตัวแปรอิสระที่มีมาตรวัดเป็น **Nominal** หรือ **Ordinal** เหล่านั้น เป็นตัวแปรหุ่น แล้วจึงทำการวิเคราะห์การถดถอย โดยใช้ตัวแปรหุ่นที่เกิดขึ้นแทนตัวแปรเดิมที่มี)
2. ข้อมูลของตัวแปรอิสระ และตัวแปรตาม จะต้องสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ
3. ตัวแปรอิสระจะต้องไม่มีความสัมพันธ์กันเอง (การเกิดความสัมพันธ์กันเองของตัวแปรอิสระ เรียกว่า การเกิด **Multicollinearity**)*
4. ข้อมูลจะต้องไม่มีความสัมพันธ์ภายในตัวเอง
5. ค่าความคลาดเคลื่อนที่เกิดจากการพยากรณ์ จะต้อง
 - 5.1 มีการแจกแจงแบบปกติ (Assumption of Normality)
 - 5.2 มีค่าเฉลี่ยเท่ากับ 0 และ
 - 5.3 มีความแปรปรวนคงที่ (Homogeneity of Variance)
 - 5.4 ความคลาดเคลื่อนเป็นอิสระกัน (Assumption of Autocorrelation)

* เฉพาะในการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ

หมายเหตุ

ตัวแปรหุ่น (Dummy variable) เป็นการสร้างตัวแปรใหม่ขึ้นใช้แทนตัวแปรเดิม โดยอาศัยชุดตัวเลข 0 และ 1 เรียงประกอบกันเพื่อใช้แทนลักษณะต่างๆ ของข้อมูลในตัวแปรนั้น ซึ่งตัวแปรที่จะนำมาสร้างตัวแปรหุ่น จะต้องเป็นตัวแปรที่บรรจุข้อมูลเชิงคุณภาพ (มาตรวัดเป็นนามบัญญัติหรืออันดับ)

ตัวอย่างการสร้างตัวแปรหุ่น ของตัวแปร “สถานภาพสมรส”

สถานภาพสมรส (STATUS)	ค่าเดิมของข้อมูล (เป็นรหัส)	ตัวแปรหุ่น	
		STAT1	STAT2
โสด	1	1	0
แต่งงาน	2	0	1
อื่นๆ	3	0	0

การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

แบบจำลองการถดถอยเชิงเส้น

ตัวแปรอิสระ (X) 1 ตัว

ตัวแปรตาม (Y) 1 ตัว

$$Y = \beta_0 + \beta_1 X + \epsilon \quad \text{เป็นสมการถดถอยของประชากร}$$

$$Y' = b_0 + b_1 X \quad \text{เป็นสมการถดถอยของกลุ่มตัวอย่าง}$$

$$Z_Y = B_1 Z_X \quad \text{เป็นสมการถดถอยของกลุ่มตัวอย่าง ในรูปคะแนนมาตรฐาน}$$

เมื่อ X, Z_X เป็น ค่าของตัวแปรอิสระในรูปคะแนนดิบ และคะแนนมาตรฐาน

Y เป็น ค่าของตัวแปรตาม

Y', Z_Y เป็น ค่าพยากรณ์ของตัวแปรตามในรูปคะแนนดิบ และคะแนนมาตรฐาน

β_0 และ β_1 เป็น ค่าคงที่ และสัมประสิทธิ์หน้าตัวแปรอิสระในสมการ (ประชากร)

b_0 และ b_1 เป็น ค่าคงที่ และสัมประสิทธิ์หน้าตัวแปรอิสระในสมการ (กลุ่มตัวอย่าง)

B_1 เป็น สัมประสิทธิ์หน้าตัวแปรอิสระในสมการ (คะแนนมาตรฐาน)

ϵ เป็น ค่าความคลาดเคลื่อน

การหาค่าคงที่และสัมประสิทธิ์ในสมการ

จาก $Y' = b_0 + b_1 X$ เป็นสมการถดถอยของกลุ่มตัวอย่าง

สามารถหาค่าของ b_0 และ b_1 ได้จากสูตร

$$b_0 = \bar{Y} - b_1 \bar{X} \quad \text{และ} \quad b_1 = r_{xy} \frac{S_y}{S_x}$$

เมื่อ \bar{Y}, \bar{X} เป็นค่าเฉลี่ยของตัวแปร Y และ X

r_{xy} เป็นค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร X กับ Y

ซึ่งคำนวณได้จากสูตร
$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

S_Y, S_X เป็นส่วนเบี่ยงเบนมาตรฐานของตัวแปร Y และ X

การหาค่าความคลาดเคลื่อนมาตรฐานของการพยากรณ์ (Standard Error of Estimation)

สามารถหาได้จากสูตร

$$S_{Y.X} = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-2}} \quad \text{หรือ} \quad S_{Y.X} = S_Y \sqrt{\frac{(1-r^2)(n-1)}{n-2}}$$

การทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอย

สมมติฐานของการทดสอบ

$$H_0: \beta = 0$$

$$H_1: \beta \neq 0$$

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

$$t = \frac{b - b}{\sqrt{\frac{S_{Y.X}^2}{(n-1)S_X^2}}}, \quad df = n-2$$

อาณาเขตวิกฤตและการสรุปผล

จะปฏิเสธ H_0 เมื่อค่า t ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง
หรือ t ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่า $-t_{\alpha, n-2}$

ช่วงความเชื่อมั่น (1- α)100% ของการประมาณค่า Y

จากสมการถดถอย $Y' = b_0 + b_1X$ ที่ได้จากกลุ่มตัวอย่าง สามารถใช้ค่า Y' มาประมาณค่า Y ที่เกิดขึ้นได้ เมื่อ $X=X_i$ ดังนี้

ช่วงความเชื่อมั่น (1- α)100% ของ Y

$$Y = Y' \pm t_{\frac{\alpha}{2}, n-2} \cdot S_{Y.X} \sqrt{\left(1 + \frac{1}{n}\right) + \frac{(X_i - \bar{X})^2}{(n-1)S_X^2}}$$

ตัวอย่างการวิเคราะห์ข้อมูล การถดถอยเชิงเส้นอย่างง่าย

ศึกษาความสัมพันธ์ระหว่าง ความดัน Systolic ในมารดาและทารก จำนวน 20 ราย ได้ข้อมูลดังตาราง

คนที่	มารดา	ทารก	X^2	Y^2	XY
1	120	89	14400.00	7921.00	10680.00
2	130	90	16900.00	8100.00	11700.00
3	105	83	11025.00	6889.00	8715.00
4	95	77	9025.00	5929.00	7315.00
5	135	92	18225.00	8464.00	12420.00
6	145	98	21025.00	9604.00	14210.00
7	110	82	12100.00	6724.00	9020.00
8	122	85	14884.00	7225.00	10370.00
9	135	96	18225.00	9216.00	12960.00
10	140	95	19600.00	9025.00	13300.00
11	108	80	11664.00	6400.00	8640.00
12	105	79	11025.00	6241.00	8295.00
13	115	86	13225.00	7396.00	9890.00
14	145	97	21025.00	9409.00	14065.00
15	140	92	19600.00	8464.00	12880.00
16	115	88	13225.00	7744.00	10120.00
17	110	82	12100.00	6724.00	9020.00
18	132	90	17424.00	8100.00	11880.00
19	140	96	19600.00	9216.00	13440.00
20	144	97	20736.00	9409.00	13968.00
Total	2491	1774	315033.00	158200.00	222888.00

สร้างสมการถดถอยเชิงเส้น และทำการทดสอบสัมประสิทธิ์การถดถอย (กำหนด $\alpha .05$)

สร้างสมการถดถอย

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

$$r_{xy} = \frac{20(222888) - (2491)(1774)}{\sqrt{[20(315033) - 2491^2][20(158200) - 1774^2]}} = 0.963$$

$$\bar{X} = 124.55 \quad \text{และ} \quad S_x = 15.859$$

$$\bar{Y} = 88.70 \quad \text{และ} \quad S_y = 6.674$$

$$\text{หา} \quad b_1 = r_{xy} \frac{S_y}{S_x} = 0.963 \times \frac{6.674}{15.859} = 0.405$$

$$\text{หา} \quad b_0 = \bar{Y} - b_1 \bar{X} = 88.70 - (0.405)(124.55) = 38.257$$

สมการที่ได้คือ $Y' = 38.257 + 0.405X$

เป็นสมการพยากรณ์ความดัน **Systolic** ของทารก (**Y**) เมื่อทราบความดัน **Systolic** ของมารดา (**X**)

หาค่าความคลาดเคลื่อนมาตรฐานของการพยากรณ์

จากสูตร

$$S_{Y.X} = S_y \sqrt{\frac{(1-r^2)(n-1)}{n-2}} = 6.674 \sqrt{\frac{[1-(0.963)^2][20-1]}{20-2}} = 1.848$$

ทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอย

$$H_0: \beta = 0$$

$$H_1: \beta \neq 0$$

กำหนด $\alpha = .05$ $df = n-2 = 18$ ค่าวิกฤต $t = 2.10$

สถิติทดสอบ

$$t = \frac{b - b_0}{\sqrt{\frac{S_{Y.X}^2}{(n-1)S_x^2}}} = \frac{0.405 - 0}{\sqrt{\frac{1.848^2}{(20-1)15.859^2}}} = \frac{0.405}{0.027} = 15.150$$

อาณาเขตวิกฤตและการสรุปผล

ค่า t ที่คำนวณ (15.15) มีค่าน้อยกว่าค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง (210) ตกในอาณาเขตวิกฤต ปฏิเสธ H_0 แสดงว่า ที่ระดับนัยสำคัญ .05 ค่าสัมประสิทธิ์ที่คำนวณได้จากกลุ่มตัวอย่างสามารถใช้แทนสัมประสิทธิ์ในสมการถดถอยของประชากรได้
นั่นคือ สมการถดถอยที่คำนวณได้จากกลุ่มตัวอย่างนี้สามารถใช้แทนสมการถดถอยของประชากรได้

ช่วงความเชื่อมั่น (1- α)100% ของ Y

จากตัวอย่างข้างต้น หากต้องการประมาณค่าความดัน **Systolic** ของทารก เมื่อทราบว่าความดัน **Systolic** ของมารดา คือ 125 (ที่ระดับความเชื่อมั่น 95%)

จากสมการ $Y' = 38.257 + 0.405X$ เมื่อ $X = 125$ $Y' = 38.257 + 0.405(125) = 88.882$

$$\text{ช่วงความเชื่อมั่น (1-}\alpha\text{)100% ของ Y} \quad Y = Y' \pm t_{\frac{\alpha}{2}, n-2} \cdot S_{Y.X} \sqrt{\left(1 + \frac{1}{n}\right) + \frac{(X_i - \bar{X})^2}{(n-1)S_X^2}}$$

$$Y = 88.882 \pm (2.10)(1.848) \sqrt{\left(1 + \frac{1}{20}\right) + \frac{(125 - 124.55)^2}{(20-1)15.859^2}} = 88.882 \pm 3.977$$

$$= 84.905 \text{ ถึง } 92.859$$

เมื่อความดัน **Systolic** ของมารดาเป็น 125 ความดัน **Systolic** ของทารก จะประมาณ 84.905 ถึง 92.859 ด้วยระดับความเชื่อมั่น 95%

การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ (Multiple Linear Regression)

แบบจำลองการถดถอยเชิงเส้น

ตัวแปรอิสระ (X) n ตัว

ตัวแปรตาม (Y) 1 ตัว

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad \text{เป็นสมการถดถอยของประชากร}$$

$$Y' = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \quad \text{เป็นสมการถดถอยของกลุ่มตัวอย่าง}$$

$$Z'_Y = B_1 Z_{X_1} + B_2 Z_{X_2} + \dots + B_n Z_{X_n} \quad \text{เป็นสมการถดถอยของกลุ่มตัวอย่าง ในรูป
คะแนนมาตรฐาน}$$

การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ เป็นการศึกษาความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระหลาย ๆ ตัว กับตัวแปรตาม ดังนั้นในการสร้างแบบจำลองสมการพยากรณ์ จะพิจารณาจากการมีตัวแปรอยู่ในระบบสมการ ซึ่งเรียกว่า การนำตัวแปรเข้าระบบสมการ ที่นิยมมีด้วยกัน 4 วิธี คือ

- All Enter** ถือว่าตัวแปรอิสระทุกตัวมีอิทธิพลต่อตัวแปรตาม จึงนำตัวแปรอิสระทุกตัวเข้าระบบสมการพร้อมๆ กันในทีเดียว
- Forward** กำหนดให้เริ่มต้นสร้างสมการยังไม่มีตัวแปรใดอยู่ในระบบสมการ จากนั้นให้เริ่มทำการสร้างระบบสมการโดยนำตัวแปรอิสระที่มีขนาดของอิทธิพลสูงสุด (โดยพิจารณาจากค่า **Partial F** ไม่ได้ดูจากค่าสัมประสิทธิ์สหสัมพันธ์) เข้าไปสร้างสมการกับตัวแปรตามก่อน จากนั้นจึงนำตัวแปรอิสระที่เหลือที่มีขนาดของอิทธิพลรองลงไปเข้าทีละตัว และจะหยุดการนำตัวแปรอิสระเข้าระบบสมการหากพบว่าตัวแปรนั้นมีขนาดของอิทธิพลน้อย (ไม่มีนัยสำคัญ) หรือไม่มีอิทธิพลเลย
- Backward** กำหนดให้เมื่อเริ่มสร้างสมการมีตัวแปรอิสระทุกตัวอยู่ครบในระบบสมการ จากนั้นให้ทำการดึงตัวแปรอิสระที่มีขนาดของอิทธิพลน้อยที่สุด (ไม่มีนัยสำคัญ) ออกจากสมการทีละตัว (โดยพิจารณาจากค่า **Partial F**) จนกระทั่งเหลือตัวแปรในระบบสมการเฉพาะที่มีอิทธิพลต่อตัวแปรตาม
- Stepwise** เป็นวิธีการที่นำตัวแปรอิสระเข้าสมการทีละตัวเช่นเดียวกับ **Forward** และเมื่อตัวแปรนั้นเข้าไปอยู่ในระบบสมการแล้ว จะทำการตรวจสอบย้อนกลับโดยวิธี **Backward** อีกทีหนึ่ง ในทุกครั้งที่มีการนำตัวแปรอิสระเข้าสมการ

สำหรับการประมาณค่าคงที่ และสัมประสิทธิ์ในสมการถดถอยเชิงเส้นแบบพหุ จะใช้วิธีการประมาณที่เรียกว่า การประมาณโดยวิธีภาวะน่าจะเป็นสูงสุด (**Maximum Likelihood Estimation : MLE**) โดยวิธีดังกล่าวนี้สามารถประมาณค่าคงที่ และสัมประสิทธิ์ในสมการได้ทั้งในกรณีที่ตัวแปรอิสระมีจำนวน 1 ตัว หรือมากกว่า 1 ตัว โดยมีข้อกำหนดเพิ่มเติมสำหรับตัวแปรอิสระแต่ละตัว คือ ตัวแปรอิสระจะต้องไม่มีความสัมพันธ์กันเอง (**intercollinearity** หรือเกิด **Multicollinearity**) การหาค่าคงที่ และสัมประสิทธิ์ในสมการ จะใช้วิธีการคำนวณโดยรูปแบบเมตริกซ์ ซึ่งในที่นี้จะไม่ขอกล่าวถึง

ผลการวิเคราะห์โดยโปรแกรมสำเร็จรูป (โปรแกรม SPSS)

การตรวจสอบการแจกแจงของข้อมูล ทั้งตัวแปรอิสระและตัวแปรตาม

*** การตรวจสอบการแจกแจงของข้อมูลจะอยู่ในการวิเคราะห์เรื่อง สถิติทดสอบ K-S ซึ่งจะไม่ขอกว่าถึงในที่นี้ แต่จะแสดงผลการวิเคราะห์ เพื่อให้ทราบถึงลักษณะของข้อมูลว่าผ่านข้อตกลงเบื้องต้นหรือไม่

One-Sample Kolmogorov-Smirnov Test

		mo_sysbp	sys_bp
N		20	20
Normal Parameters ^{a,b}	Mean	124.55	88.70
	Std. Deviation	15.859	6.674
Most Extreme Differences	Absolute	.145	.127
	Positive	.126	.103
	Negative	-.145	-.127
Kolmogorov-Smirnov Z		.649	.570
Asymp. Sig. (2-tailed)		.794	.902

a. Test distribution is Normal.

b. Calculated from data.

H_0 : ความดัน Systolic ของมารดา (mo_sysbp) มีการแจกแจงแบบปกติ

H_1 : ความดัน Systolic ของมารดา (mo_sysbp) ไม่มีการแจกแจงแบบปกติ

จากผลการวิเคราะห์ ค่า Sig= .794 มีค่ามากกว่า ระดับนัยสำคัญ .05 (ตกใน H_0) แสดงว่า ความดัน Systolic ของมารดา (mo_sysbp) มีการแจกแจงแบบปกติ

H_0 : ความดัน Systolic ของทารก (sys_bp) มีการแจกแจงแบบปกติ

H_1 : ความดัน Systolic ของทารก (sys_bp) ไม่มีการแจกแจงแบบปกติ

จากผลการวิเคราะห์ ค่า Sig= .902 มีค่ามากกว่า ระดับนัยสำคัญ .05 (ตกใน H_0) แสดงว่า ความดัน Systolic ของทารก มีการแจกแจงแบบปกติ

การตรวจสอบค่าความคลาดเคลื่อนที่เกิดจากการพยากรณ์

การตรวจสอบ : ความคลาดเคลื่อนที่เกิดจากการพยากรณ์ มีค่าเฉลี่ยเท่ากับ 0

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	76.73	96.99	88.70	6.426	20
Residual	-2.960	3.169	.000	1.802	20
Std. Predicted Value	-1.863	1.289	.000	1.000	20
Std. Residual	-1.599	1.712	.000	.973	20

a. Dependent Variable: sys_bp

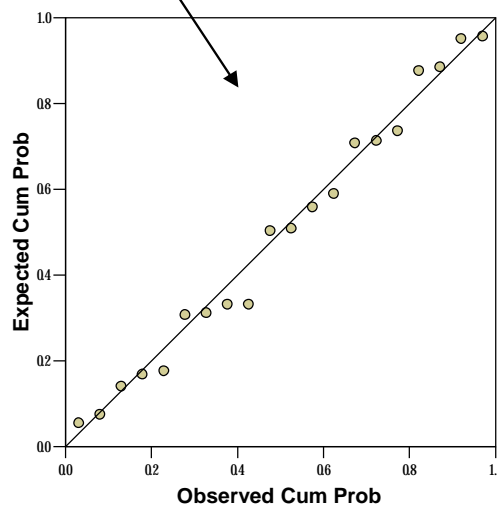
ค่าเฉลี่ยความคลาดเคลื่อน (Residual) = .000 ซึ่งมีค่าใกล้เคียง หรือเท่ากับศูนย์

การตรวจสอบ : ความคลาดเคลื่อนที่เกิดจากการพยากรณ์ จะต้องมีการแจกแจงแบบปกติ

จะใช้ **Normal Prob. Plot** ในการตรวจสอบ
จุด **☉** แสดงถึงค่า **Prob.** ของข้อมูล ส่วนเส้นตรงที่ลากทะแยง เป็นแนวเส้นแสดงการแจกแจงแบบปกติ ดังนั้นหาก **☉** มีการเรียงตัวใกล้เคียงเส้นตรงจะถือว่ามี การแจกแจงแบบปกติ

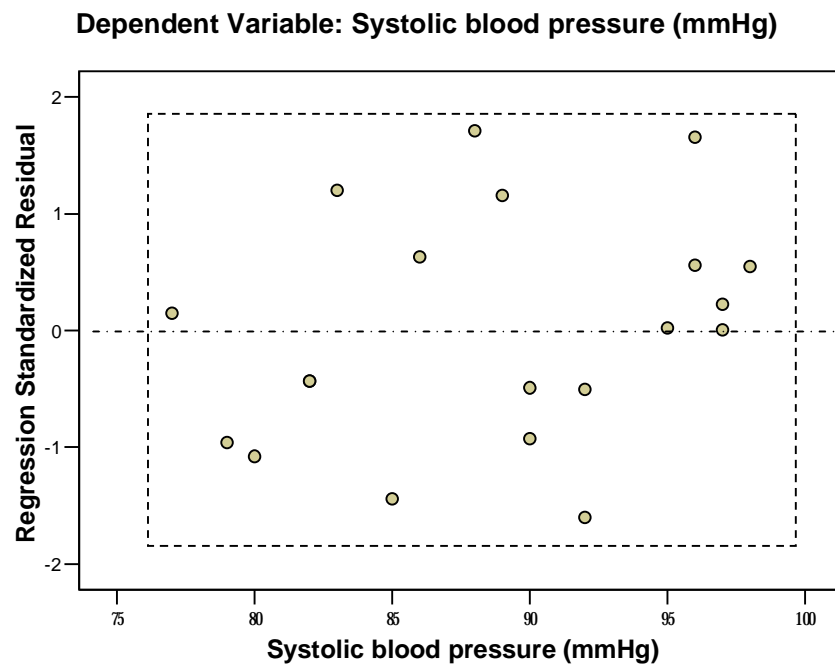
Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Systolic blood pressure (mmHg)



การตรวจสอบ : ความคลาดเคลื่อนที่เกิดจากการพยากรณ์ มีความแปรปรวนคงที่

Scatterplot



จากกราฟจะเห็นว่า การกระจายของเหนือและใต้ระดับ **0.0** มีพื้นที่ใกล้เคียงกันและเป็นแนวตั้งเหลี่ยมผืนผ้าได้ แสดงว่าความคลาดเคลื่อน มีความแปรปรวนคงที่

การตรวจสอบ : ความคลาดเคลื่อนเป็นอิสระต่อกัน (ไม่เกิด **Autocorrelation**)

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.963 ^a	.927	.923	1.851	2.630

a. Predictors: (Constant), MO_SYSBP

b. Dependent Variable: Systolic blood pressure (mmHg)

การตรวจสอบความเป็นอิสระของ **error** จะใช้ค่า **Durbin-Watson (d)** เป็นตัวตรวจสอบ

H_0 : no autocorrelation

H_1 : autocorrelation

จะปฏิเสธ H_0 เมื่อ ค่า d ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับ d_L หรือ ค่า d ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับ $4d_U$ จากตารางค่าวิกฤต **Durbin-watson**

จะยอมรับ H_0 เมื่อ ค่า $d_U < d < 4d_U$

นอกนั้นไม่สามารถสรุปได้

กำหนด $\alpha = .01$

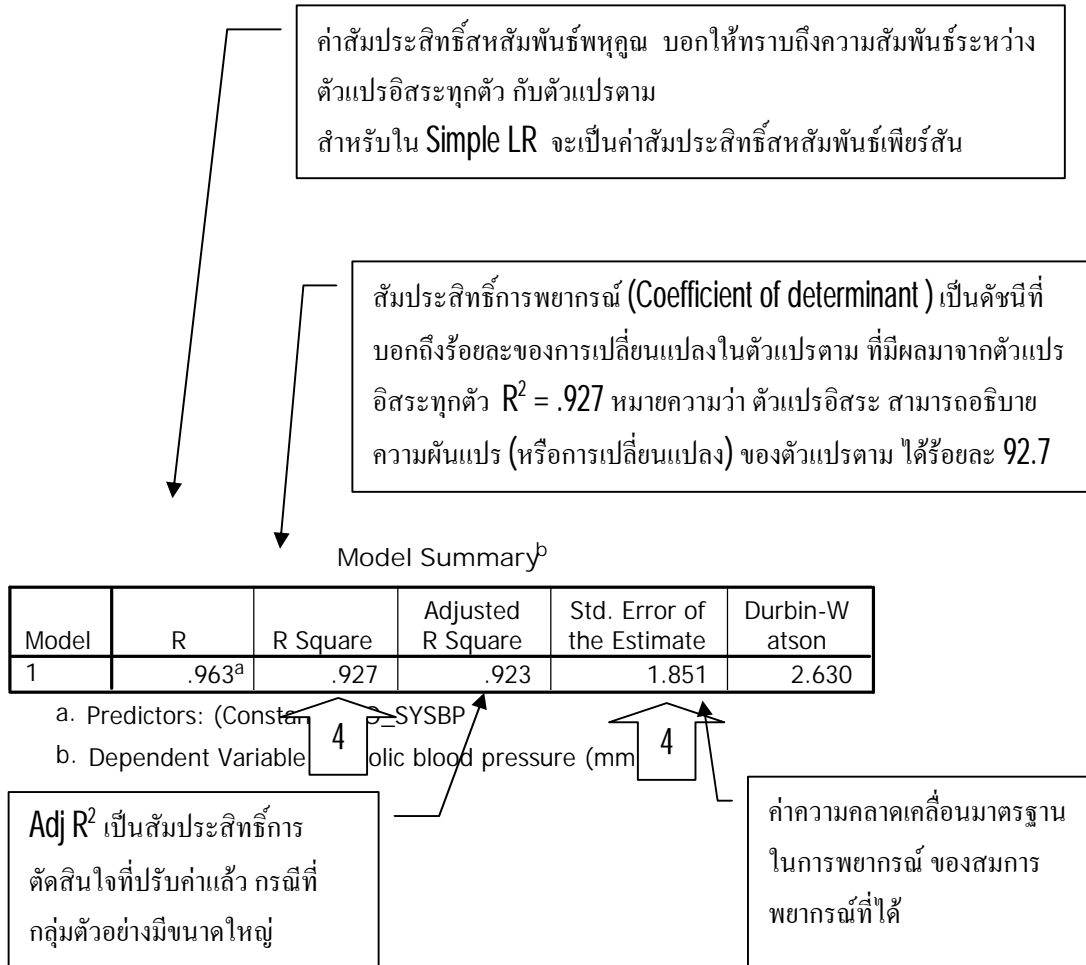
เปิดตาราง ใช้ $p = 2$ (จำนวนตัวแปรอิสระ และตัวแปรตาม) และ $n = 20$ ค่า $d_L = 0.95$ และ $d_U = 1.15$ พบว่า $0.95 < d(2.630) < 4(1.15)$ ตกใน H_0 แสดงว่า ความคลาดเคลื่อน (error) ของการพยากรณ์ มีความเป็นอิสระต่อกัน

ลำดับของการอ่านผลการวิเคราะห์การถดถอย

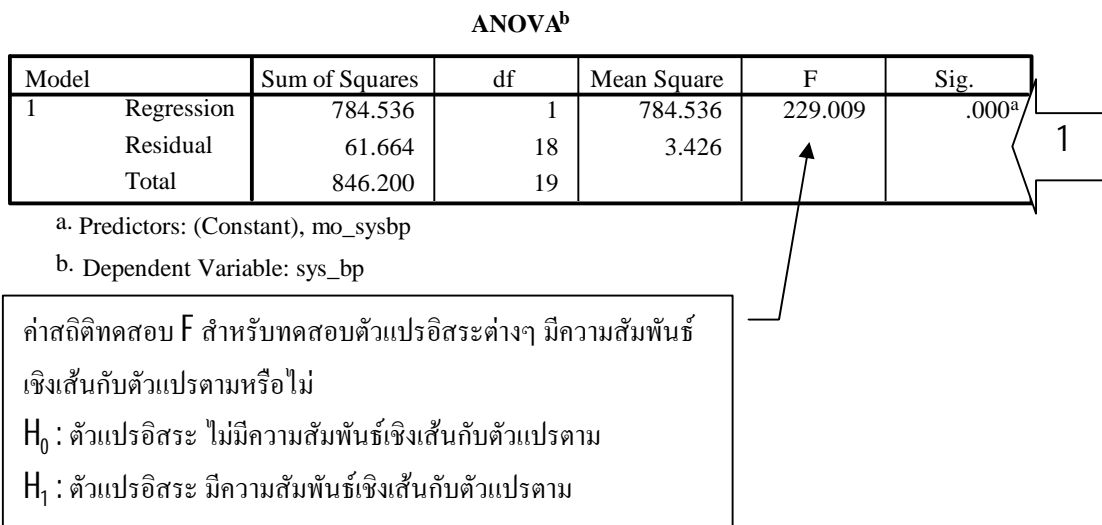
เมื่อผ่านการตรวจสอบข้อตกลงเบื้องต้นในการวิเคราะห์การถดถอย แล้ว ลำดับของการอ่านผลการวิเคราะห์ เป็นดังนี้

1. การตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระกับตัวแปรตาม (**F test** ในตาราง **ANOVA**)
2. การตรวจสอบความมีนัยสำคัญของค่าคงที่ และสัมประสิทธิ์แต่ละตัวในสมการ (ในตาราง **Coefficients**)
3. การเขียนรูปแบบของสมการถดถอยจากค่าสัมประสิทธิ์ที่ได้ (จากตาราง **Coefficients**)
4. การสรุปสมการถดถอยที่ได้ (ตาราง **Model Summary**)

ลำดับที่ 4 นำค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ สัมประสิทธิ์การพยากรณ์ และค่าความคลาดเคลื่อนมาตรฐานมาสรุปร่วมกับสมการถดถอยที่ได้



ลำดับที่ 1 ตรวจสอบความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระกับตัวแปรตาม



ลำดับที่ 2 ตรวจสอบความมีนัยสำคัญของค่าคงที่ และสัมประสิทธิ์ในสมการ

(B) แสดงค่าสัมประสิทธิ์การถดถอย ของตัวแปรอิสระ ที่อยู่ในสมการ (รูปคะแนนดิบ) (Std. Error) แสดงค่าความคลาดเคลื่อนมาตรฐานของสัมประสิทธิ์การถดถอย บอกถึงการกระจายของค่าสัมประสิทธิ์ หากมีการกระจายมากแสดงว่ามีความแม่นยำหรือเชื่อถือได้น้อย

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	38.236	3.360		11.379	.000
	mo_sysbp	.405	.027	.963	15.133	.000

a. Dependent Variable: sys_bp

(Beta) ค่าสัมประสิทธิ์การถดถอย ในสมการที่อยู่ในรูปคะแนนมาตรฐาน แสดงถึงขนาดของอิทธิพลของตัวแปรอิสระนั้นๆ หากเป็นการวิเคราะห์การถดถอยเชิงพหุ ซึ่งมีตัวแปรอิสระหลายตัว สามารถนำค่า Beta มาเปรียบเทียบกันได้

ค่าสถิติ t สำหรับทดสอบความมีนัยสำคัญของค่าคงที่ และสัมประสิทธิ์แต่ละตัวในสมการ
 $H_0 : \beta_i = 0$
 $H_1 : \beta_i \neq 0$

สมการพยากรณ์จะเป็น

$$\text{Sys_bp}' = 38.236 + .405 \text{ mo_sysbp}$$

ลำดับที่ 3 จากค่าคงที่ และสัมประสิทธิ์ในสมการที่ได้ นำมาเขียนรูปแบบสมการ

สมการพยากรณ์จะเป็น $\text{Sys_bp}' = 38.236 + .405 \text{ mo_sysbp}$

การเขียนรายงานผลการวิเคราะห์

โดยทั่วไปในการเสนอรายงานผลการวิเคราะห์ มักนำเสนอในรูปแบบตารางแสดงการวิเคราะห์การถดถอย และอธิบายสมการพยากรณ์ ดังนี้

ตารางที่ ... การวิเคราะห์การถดถอยเพื่อพยากรณ์ความดัน **Systolic** ของทารก จากความดัน **Systolic** ของมารดา (mmHg)

ค่าคงที่/ตัวแปร	b	SE _b	β	t	p-value
ค่าคงที่	38.236	3.360		11.379	.000
ความดัน Systolic ของมารดา	0.405	0.027	0.963	15.133	.000

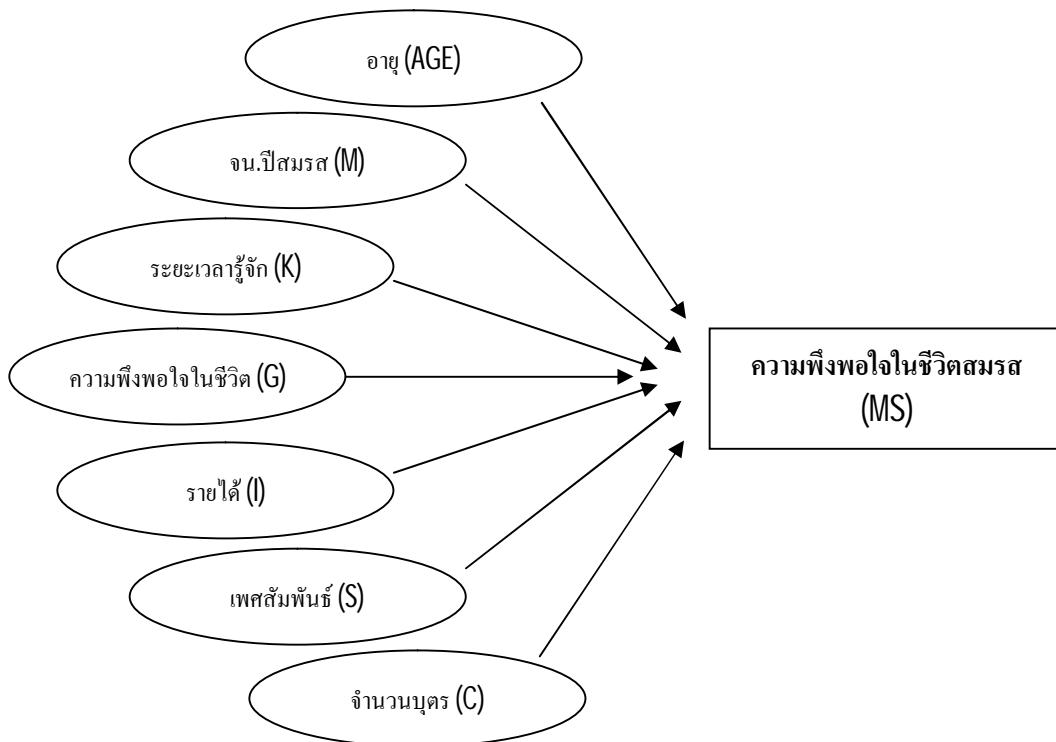
SE_{est} = ±1.851
R = .963; R² = .927; F = 229.009; p-value = .000

ความดัน **Systolic** ของทารก มีความสัมพันธ์กับความดัน **Systolic** ของมารดา โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ **.963** และสามารถพยากรณ์ความดัน **Systolic** ของทารก ได้ร้อยละ **92.7** อย่างมีนัยสำคัญทางสถิติที่ระดับ **.001** โดยมีความคลาดเคลื่อนมาตรฐานในการพยากรณ์ เท่ากับ **±1.851** และสมการพยากรณ์ความดัน **Systolic** ของทารก เมื่อทราบความดัน **Systolic** ของมารดา เป็น $Y' = 38.236 + 0.405X$

ตัวอย่างการวิเคราะห์ Multiple Linear Regression โดยใช้โปรแกรม SPSS

นักวิจัยผู้หนึ่ง ต้องการศึกษาลึถึงปัจจัยที่มีอิทธิพลต่อ ความพึงพอใจในชีวิตสมรส จึงได้ทำการสุ่มตัวอย่างชาย หรือหญิงที่มีสถานภาพสมรส มาจำนวน 82 ราย ทำการบันทึกข้อมูล อายุของกลุ่มตัวอย่าง เพศ จำนวนปีที่ใช้ชีวิตสมรส ระยะเวลาที่รู้จักกันก่อนแต่งงาน(เดือน) รายได้ของครอบครัว จำนวนบุตร คะแนนความพึงพอใจในชีวิต คะแนนความพึงพอใจด้านเพศสัมพันธ์ และคะแนนความพึงพอใจในชีวิตสมรส โดยใช้แบบวัดความพึงพอใจที่สร้างขึ้น

เมื่อสำรวจข้อมูลได้แล้ว ผู้วิจัยทำการวิเคราะห์โดยใช้โปรแกรมสำเร็จรูปทางสถิติ ได้ทำการตรวจสอบการแจกแจงของตัวแปรทุกตัว พบว่ามีการแจกแจงแบบปกติ จากนั้นทำการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ โดยกำหนดให้ตัวแปรตาม คือ คะแนนความพึงพอใจในชีวิตสมรส (MS) ส่วนตัวแปรอิสระที่ศึกษา ได้แก่ อายุ (AGE) จำนวนปีที่ใช้ชีวิตสมรส (M) ระยะเวลาที่รู้จักกัน ก่อนแต่งงาน (เดือน) (K) คะแนนความพึงพอใจในชีวิต (G) รายได้ของครอบครัว (หน่วย : พันบาท) (I) คะแนนความพึงพอใจด้านเพศสัมพันธ์ (S) และ จำนวนบุตร (C)



การตรวจสอบข้อตกลงเบื้องต้นในการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ

ในการตรวจสอบข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ จะมีการตรวจสอบเช่นเดียวกันกับการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (ข้อตกลง 1,24 และ 5) แต่จะต้องเพิ่มการตรวจสอบว่าตัวแปรอิสระไม่มีความสัมพันธ์กันเอง (ข้อ 3) ในที่นี้จะขอกว่าเฉพาะการตรวจสอบตัวแปรอิสระไม่มีความสัมพันธ์กันเอง (ไม่เกิด **Multicollinearity**) เท่านั้น ส่วนการตรวจสอบที่เหลือ จะมีกระบวนการ และการอ่านผลการวิเคราะห์เช่นเดียวกันกับการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย

การตรวจสอบ : ตัวแปรอิสระจะต้องไม่มีความสัมพันธ์กันเอง (ไม่เกิด **Multicollinearity**)

การตรวจสอบ **Multicollinearity** จะใช้ค่า **Variance inflation factor (VIF)** หรือค่า **Tolerance** หรือค่า **Eigen Value** ตัวใด ตัวหนึ่งก็ได้ โดยมีเกณฑ์การตรวจสอบดังนี้

Variance inflation factor (VIF)

ค่า **VIF** ที่เหมาะสมไม่ควรเกิน 4 หรือ 5 หากเกินกว่านี้แสดงว่าตัวแปรอิสระมีความสัมพันธ์กันเอง (ตำราบางแนะนำว่า หาก **VIF > 10** แสดงว่าเกิด **Multicollinearity**)

Tolerance

หากค่า **Tolerance < 0.2** แสดงว่าเกิด **Multicollinearity**

Eigen Value

หากค่า **Eigen Value** ตัวที่มากที่สุด มีค่า ≥ 10 แสดงว่าเกิด **Multicollinearity**

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	9.422	.294		32.039	.000		
	M	-.215	.009	-.314	-24.562	.000	.331	3.024
	AGE	.002	.006	.004	.365	.716	.540	1.850
	K	.111	.003	.355	39.414	.000	.665	1.504
	G	.235	.005	.383	44.861	.000	.739	1.352
	I	.121	.002	.597	59.337	.000	.533	1.875
	S	2.813	.051	.597	55.500	.000	.466	2.148
	C	.174	.029	.048	6.024	.000	.844	1.185

a. Dependent Variable: MS

จากผลการวิเคราะห์จะพบว่า ค่า **VIF** สูงสุดที่ได้มีค่า **3.024** ซึ่งไม่เกิน 4 หรือค่า **Tolerance** ที่มีค่าน้อยที่สุด **.331** ซึ่งไม่ต่ำกว่า **.2** และค่า **Eigenvalue** ที่มากที่สุด **7.279** ซึ่งไม่เกิน **10** แสดงว่าตัวแปรอิสระไม่มีความสัมพันธ์กัน (ไม่เกิด **Multicollinearity**)

การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ นำตัวแปรเข้าระบบสมการ โดยวิธี **All Enter**

Regression

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	C, K, AGE ^a G, I, S, M	.	Enter

a. All requested variables entered.

b. Dependent Variable: MS

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.998 ^a	.996	.996	.265

a. Predictors: (Constant), C, K, AGE, G, I, S, M

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1297.293	7	185.328	2639.875	.000 ^a
	Residual	5.195	74	.070		
	Total	1302.488	81			

a. Predictors: (Constant), C, K, AGE, G, I, S, M

b. Dependent Variable: MS

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	9.422	.294		32.039	.000
	M	-.215	.009	-.314	-24.562	.000
	AGE	.002	.006	.004	.365	.716
	K	.111	.003	.355	39.414	.000
	G	.235	.005	.383	44.861	.000
	I	.121	.002	.597	59.337	.000
	S	2.813	.051	.597	55.500	.000
	C	.174	.029	.048	6.024	.000

a. Dependent Variable: MS

การเสนอรายงานผลการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ วิธี All Enter

โดยทั่วไปในการเสนอผลการวิเคราะห์ มักนำเสนอในรูปแบบตารางแสดงการวิเคราะห์การถดถอย และอธิบายสมการพยากรณ์ ดังนี้

ตารางที่ ... การวิเคราะห์การถดถอยเพื่อพยากรณ์ความพึงพอใจในชีวิตสมรส โดยนำปัจจัยทุกด้านเข้าในสมการ

ตัวแปร	b	SE _b	β	t	p-value
จำนวนปีที่ใช้ชีวิตสมรส	-.215	.009	-.314	-24.562	.000
อายุ	.002	.006	.004	.365	.716
ระยะเวลาที่รู้จักกันก่อนแต่งงาน	.111	.003	.355	39.414	.000
ความพึงพอใจในชีวิต	.235	.005	.383	44.861	.000
รายได้ของครอบครัว	.121	.002	.597	59.337	.000
ความพึงพอใจด้านเพศสัมพันธ์	2.813	.051	.597	55.500	.000
จำนวนบุตร	.174	.029	.048	6.024	.000

ค่าคงที่ 9.422; SE_{est} = ±.26
R = .998; R² = .996; F = 2639.875; p-value = .000

จากตารางที่ ... จะเห็นว่าปัจจัยทั้ง 7 ด้านมีความสัมพันธ์กับความพึงพอใจในชีวิตสมรสในระดับสูงมาก โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ทุกคู่เป็น .998 และสามารถร่วมกันพยากรณ์ความพึงพอใจในชีวิตสมรส ได้ร้อยละ 99.6 อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยมีความคลาดเคลื่อนมาตรฐานในการพยากรณ์ เท่ากับ ±.26

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอยของตัวพยากรณ์ พบว่ามีเพียงปัจจัยด้านอายุเท่านั้นที่ไม่สามารถพยากรณ์ความพึงพอใจในชีวิตสมรสได้ (p-value .716) ส่วนปัจจัยด้านจำนวนปีที่ใช้ชีวิตสมรส ระยะเวลาที่รู้จักกันก่อนแต่งงาน ความพึงพอใจในชีวิต ความพึงพอใจด้านเพศสัมพันธ์ รายได้ครอบครัว และจำนวนบุตร สามารถพยากรณ์ความพึงพอใจในชีวิตสมรส ได้อย่างมีนัยสำคัญทางสถิติที่ระดับ .001

สมการพยากรณ์ความพึงพอใจในชีวิตสมรส เมื่อนำปัจจัยทุกด้านเข้าสมการในรูปแบบคะแนนดิบ เป็นดังนี้

$$MS' = 9.422 - .215M + .002AGE + .111K + .235G + .121I + 2.813S + .174C$$

การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ นำตัวแปรเข้าระบบสมการ โดยวิธี *Stepwise*

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	K	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	I	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	S	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
4	G	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
5	M	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
6	C	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: MS

Model Summary^g

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.680 ^a	.463	.456	2.957	.463	68.982	1	80	.000	
2	.882 ^b	.777	.772	1.916	.314	111.449	1	79	.000	
3	.936 ^c	.877	.872	1.436	.099	62.745	1	78	.000	
4	.979 ^d	.957	.955	.848	.081	146.574	1	77	.000	
5	.997 ^e	.994	.994	.320	.037	466.177	1	76	.000	
6	.998 ^f	.996	.996	.263	.002	36.860	1	75	.000	1.935

- a. Predictors: (Constant), K
- b. Predictors: (Constant), K, I
- c. Predictors: (Constant), K, I, S
- d. Predictors: (Constant), K, I, S, G
- e. Predictors: (Constant), K, I, S, G, M
- f. Predictors: (Constant), K, I, S, G, M, C
- g. Dependent Variable: MS

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	603.082	1	603.082	68.982	.000 ^a
	Residual	699.406	80	8.743		
	Total	1302.488	81			
2	Regression	1012.367	2	506.184	137.834	.000 ^b
	Residual	290.121	79	3.672		
	Total	1302.488	81			
3	Regression	1141.705	3	380.568	184.624	.000 ^c
	Residual	160.783	78	2.061		
	Total	1302.488	81			
4	Regression	1247.113	4	311.778	433.537	.000 ^d
	Residual	55.375	77	.719		
	Total	1302.488	81			
5	Regression	1294.726	5	258.945	2535.352	.000 ^e
	Residual	7.762	76	.102		
	Total	1302.488	81			
6	Regression	1297.283	6	216.214	3115.840	.000 ^f
	Residual	5.204	75	.069		
	Total	1302.488	81			

a. Predictors: (Constant), K

b. Predictors: (Constant), K, I

c. Predictors: (Constant), K, I, S

d. Predictors: (Constant), K, I, S, G

e. Predictors: (Constant), K, I, S, G, M

f. Predictors: (Constant), K, I, S, G, M, C

g. Dependent Variable: MS

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	30.006	.531		56.472	.000		
	K	.213	.026	.680	8.306	.000	1.000	1.000
2	(Constant)	23.911	.672		35.565	.000		
	K	.208	.017	.665	12.520	.000	.999	1.001
	I	.114	.011	.561	10.557	.000	.999	1.001
3	(Constant)	18.464	.852		21.664	.000		
	K	.150	.014	.480	10.404	.000	.743	1.345
	I	.106	.008	.521	12.988	.000	.984	1.017
	S	1.734	.219	.368	7.921	.000	.733	1.365
4	(Constant)	12.539	.702		17.858	.000		
	K	.122	.009	.390	13.807	.000	.692	1.446
	I	.089	.005	.437	17.696	.000	.906	1.104
	S	2.128	.133	.452	15.957	.000	.689	1.451
	G	.188	.016	.307	12.107	.000	.859	1.164
5	(Constant)	10.122	.287		35.231	.000		
	K	.114	.003	.363	33.906	.000	.683	1.465
	I	.122	.002	.597	50.169	.000	.553	1.807
	S	2.709	.057	.575	47.520	.000	.535	1.868
	G	.235	.006	.383	37.613	.000	.756	1.322
	M	-.205	.009	-.298	-21.591	.000	.411	2.434
6	(Constant)	9.471	.260		36.432	.000		
	K	.111	.003	.355	39.648	.000	.665	1.503
	I	.122	.002	.597	60.862	.000	.553	1.807
	S	2.815	.050	.598	56.172	.000	.471	2.124
	G	.234	.005	.383	45.580	.000	.756	1.322
	M	-.214	.008	-.312	-26.871	.000	.396	2.525
	C	.174	.029	.048	6.071	.000	.844	1.184

a. Dependent Variable: MS

Excluded Variable^g

Model	Beta In	t	Sig.	Partial Correlation	Collinearity Statistics			
					Tolerance	VIF	Minimum Tolerance	
1	AGE	.315 ^a	4.209	.000	.428	.988	1.012	.988
	M	.412 ^a	5.899	.000	.553	.969	1.032	.969
	G	.336 ^a	4.514	.000	.453	.973	1.027	.973
	I	.561 ^a	10.557	.000	.765	.999	1.001	.999
	S	.444 ^a	5.446	.000	.522	.744	1.343	.744
2	C	.001 ^a	.011	.991	.001	1.000	1.000	1.000
2	AGE	.036 ^b	.574	.568	.065	.728	1.373	.728
	M	.068 ^b	.952	.344	.107	.554	1.806	.554
	G	.208 ^b	4.109	.000	.422	.914	1.095	.914
	S	.368 ^b	7.921	.000	.668	.733	1.365	.733
	C	-.074 ^b	-1.388	.169	-.155	.983	1.018	.982
3	AGE	-.091 ^c	-1.890	.062	-.211	.654	1.529	.654
	M	-.118 ^c	-2.077	.041	-.230	.467	2.143	.467
	G	.307 ^c	12.107	.000	.810	.859	1.164	.689
	C	.031 ^c	.739	.462	.084	.881	1.135	.657
4	AGE	-.094 ^d	-3.458	.001	-.369	.654	1.529	.622
	M	-.298 ^d	-21.591	.000	-.927	.411	2.434	.411
	C	.008 ^d	.304	.762	.035	.876	1.142	.629
5	AGE	.005 ^e	.434	.666	.050	.541	1.849	.340
	C	.048 ^e	6.071	.000	.574	.844	1.184	.396
6	AGE	.004 ^f	.365	.716	.042	.540	1.850	.331

- a. Predictors in the Model: (Constant), K
- b. Predictors in the Model: (Constant), K, I
- c. Predictors in the Model: (Constant), K, I, S
- d. Predictors in the Model: (Constant), K, I, S, G
- e. Predictors in the Model: (Constant), K, I, S, G, M
- f. Predictors in the Model: (Constant), K, I, S, G, M, C
- g. Dependent Variable: MS

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions						
				(Constant)	K	I	S	G	M	C
1	1	1.789	1.000	.11	.11					
	2	.211	2.911	.89	.89					
2	1	2.654	1.000	.01	.04	.02				
	2	.289	3.031	.03	.90	.09				
	3	.057	6.824	.95	.06	.90				
3	1	3.616	1.000	.00	.02	.01	.00			
	2	.293	3.513	.01	.72	.06	.00			
	3	.073	7.034	.09	.10	.90	.09			
	4	.018	14.077	.90	.16	.03	.90			
4	1	4.561	1.000	.00	.01	.00	.00	.00		
	2	.307	3.853	.00	.68	.04	.00	.01		
	3	.076	7.728	.03	.07	.94	.04	.03		
	4	.044	10.169	.00	.05	.01	.31	.44		
	5	.011	20.287	.96	.19	.01	.65	.52		
5	1	5.479	1.000	.00	.01	.00	.00	.00	.00	
	2	.320	4.135	.00	.66	.02	.00	.00	.01	
	3	.102	7.329	.04	.05	.12	.02	.04	.20	
	4	.050	10.511	.00	.10	.41	.15	.09	.26	
	5	.040	11.659	.01	.00	.36	.10	.32	.25	
	6	.008	25.420	.95	.18	.09	.73	.55	.28	
6	1	6.309	1.000	.00	.00	.00	.00	.00	.00	.00
	2	.338	4.318	.00	.61	.01	.00	.00	.00	.04
	3	.165	6.176	.00	.03	.04	.00	.00	.06	.59
	4	.094	8.180	.03	.08	.09	.04	.04	.13	.12
	5	.047	11.584	.00	.06	.59	.06	.05	.39	.05
	6	.039	12.749	.01	.01	.18	.12	.46	.12	.05
	7	.007	29.310	.96	.21	.08	.78	.45	.30	.14

- a. Dependent Variable: MS

การเสนอรายงานผลการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ วิธี Stepwise

การเสนอผลการวิเคราะห์ จะเสนอโดยยึดตามวัตถุประสงค์ของการวิจัย โดยทั่วไปมักนำเสนอเป็นรูปแบบดังนี้

แบบที่ **1** นำเสนอในขั้นตอนที่มีตัวแปรอิสระในสมการถดถอย จำนวนมากที่สุด

แบบที่ **2** นำเสนอโดยพิจารณาทุกขั้นตอนที่มีการนำตัวแปรอิสระเข้าในระบบสมการ โดยจะนำเสนอตารางสรุปเพื่อให้เห็นการเพิ่มของค่า R และ R^2 ในทุกครั้งที่มีการนำตัวแปรอิสระเข้าสมการ พร้อมนำเสนอขั้นตอนที่มีตัวแปรอิสระในสมการถดถอย จำนวนมากที่สุดเป็นตารางสรุป

แบบที่ 1

ตารางที่ ... การวิเคราะห์การถดถอยเพื่อพยากรณ์ความพึงพอใจในชีวิตสมรส โดยนำปัจจัยทุก
ด้านเข้าในสมการ

ตัวแปร	b	SE _b	β	t	p-value
ระยะเวลาที่รู้จักกันก่อนแต่งงาน	.111	.003	.355	39.648	.000
รายได้ของครอบครัว	.122	.002	.597	60.862	.000
ความพึงพอใจด้านเพศสัมพันธ์	2.815	.050	.598	56.172	.000
ความพึงพอใจในชีวิต	.234	.005	.383	45.580	.000
จำนวนปีที่ใช้ชีวิตสมรส	-.214	.008	-.312	-26.871	.000
จำนวนบุตร	.174	.029	.048	6.071	.000

ค่าคงที่ 9.471 ; SE_{est} = ±.26
R = .998; R² = .996; F = 3115.840; p-value = .000

จากตารางที่ ... จะเห็นว่าระยะเวลาที่รู้จักกันก่อนแต่งงาน รายได้ครอบครัว ความพึงพอใจด้านเพศสัมพันธ์ ความพึงพอใจในชีวิต จำนวนปีที่ใช้ชีวิตสมรส และจำนวนบุตร ความสัมพันธ์แบบพหุคูณกับความพึงพอใจในชีวิตสมรส อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยมีค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณเป็น .998 และสามารถพยากรณ์ความพึงพอใจในชีวิตสมรส ได้ร้อยละ 99.6 โดยมีความคลาดเคลื่อนมาตรฐานในการพยากรณ์ เท่ากับ ±.26

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอยของตัวพยากรณ์ พบว่าความพึงพอใจด้านเพศสัมพันธ์ และรายได้ของครอบครัว สามารถพยากรณ์ความพึงพอใจในชีวิตสมรสได้สูงสุด และมีนัยสำคัญทางสถิติที่ระดับ .001 โดยมีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนดิบ และค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐาน (b,β) เป็น 2.815 กับ .598 และ .122 กับ .597 ตามลำดับ รองลงมาเป็นความพึงพอใจในชีวิตสมรส สามารถพยากรณ์ความพึงพอใจในชีวิตสมรส อย่างมีนัยสำคัญทางสถิติที่ระดับ .001 โดยมีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนดิบ และค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐาน เป็น .234 และ .383 ส่วนจำนวนบุตร สามารถพยากรณ์ความพึงพอใจในชีวิตสมรสได้น้อยที่สุด โดยมีนัยสำคัญทางสถิติที่ระดับ .001 มีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนดิบ และค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐาน เป็น .174 และ .048

สมการพยากรณ์ความพึงพอใจในชีวิตสมรส สามารถแสดงในรูปคะแนนดิบ เป็นดังนี้

$$MS' = 9.471 + .111K + .122I + 2.815S + .234G - .214M + .174C$$

แบบที่ 2

ตารางที่ ... แสดงการเพิ่มสัมประสิทธิ์สหสัมพันธ์พหุคูณ (R) และสัมประสิทธิ์การพยากรณ์ (R²)
เมื่อเพิ่มปัจจัยเข้าไปทีละตัว

ตัวแปร	R	R ²	R ² change	SE _{est}	F	p-value
K	.680	.463	.463	2.96	68.982	.000
K I	.882	.777	.314	1.92	137.834	.000
K I S	.936	.877	.099	1.44	184.624	.000
K I S G	.979	.957	.081	.85	433.537	.000
K I S G M	.997	.994	.037	.32	2535.352	.000
K I S G M C	.998	.996	.002	.26	3115.840	.000

จากตาราง พบว่าระยะเวลาที่รู้จักกันก่อนแต่งงาน เป็นตัวแปรตัวแรกที่สามารถพยากรณ์ความพึงพอใจในชีวิตสมรสได้ ต่อมาเป็น รายได้ของครอบครัว ความพึงพอใจด้านเพศสัมพันธ์ ความพึงพอใจในชีวิต จำนวนปีที่ใช้ชีวิตสมรส และจำนวนบุตร ตามลำดับ มีค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณเป็น **.998** อย่างมีนัยสำคัญทางสถิติที่ระดับ **.001** และสามารถพยากรณ์ความพึงพอใจในชีวิตสมรส ได้ร้อยละ **99.6** โดยมีความคลาดเคลื่อนมาตรฐานในการพยากรณ์ เท่ากับ **± 26**

ตารางที่ ... แสดงค่าสัมประสิทธิ์การถดถอยพหุคูณ ค่าคงที่ และค่าความคลาดเคลื่อนมาตรฐานในการพยากรณ์

ตัวแปร	b	β	t	p-value
ระยะเวลาที่รู้จักกันก่อนแต่งงาน (เดือน)	.111	.355	39.648	.000
รายได้ของครอบครัว	.122	.597	60.862	.000
ความพึงพอใจด้านเพศสัมพันธ์	2.815	.598	56.172	.000
ความพึงพอใจในชีวิต	.234	.383	45.580	.000
จำนวนปีที่ใช้ชีวิตสมรส	-.214	-.312	-26.871	.000
จำนวนบุตร	.174	.048	6.071	.000

ค่าคงที่ 9.471 ; SE_{est} = ±.26

จากตารางที่ ... จะเห็นว่าค่าสัมประสิทธิ์การถดถอยของตัวพยากรณ์ ความพึงพอใจด้านเพศสัมพันธ์ และรายได้ของครอบครัว สามารถพยากรณ์ความพึงพอใจในชีวิตสมรสได้สูงสุด และมีนัยสำคัญทางสถิติที่ระดับ **.001** โดยมีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนดิบ และค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐาน (b, β) เป็น **2.815** กับ **.598** และ **.122** กับ **.597** ตามลำดับ รองลงมาเป็นความพึงพอใจในชีวิตสมรส สามารถพยากรณ์ความพึงพอใจในชีวิตสมรส อย่างมีนัยสำคัญทางสถิติที่ระดับ **.001** โดยมีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนดิบ และค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐาน เป็น **.234** และ **.383** ส่วนจำนวนบุตรสามารถพยากรณ์ความพึงพอใจในชีวิตสมรสได้น้อยที่สุด โดยมีนัยสำคัญทางสถิติที่ระดับ **.001** มีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนดิบ และค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐาน เป็น **.174** และ **.048** โดยมีความคลาดเคลื่อนมาตรฐานในการพยากรณ์ เท่ากับ ± 26

สมการพยากรณ์ความพึงพอใจในชีวิตสมรส สามารถแสดงในรูปคะแนนดิบ เป็นดังนี้

$$MS' = 9.471 + .111K + .122I + 2.815S + .234G - .214M + .174C$$