บทที่ 3

การวิเคราะห์ล็อกลิเนียร์และการอนุมาน

การวิเคราะห์ล็อกลิเนียร์ เป็นสถิติที่มีหลักการวิเคราะห์ข้อมูลมาจากการบูรณาการเทคนิควิธีการวิเคราะห์ข้อมูล 3 วิธีด้วยกัน คือ การวิเคราะห์ความแปรปรวน (Analysis Of Variance: ANOVA) การวิเคราะห์การถดถอย (Regression Analysis) และการทดสอบความกลมกลืนแบบไควสแควร์หรือการทดสอบภาวะรูปดี (Goodness of fit) โดยการวิเคราะห์ล็อกลิเนียร์จะมีลักษณะคล้ายกับการวิเคราะห์ความแปรปรวนตรงที่สามารถทราบอิทธิพลหลัก (Main effect) และอิทธิพลปฏิสัมพันธ์ (Interaction effect) ของตัวแปรและยังมีโมเดลในการนำเสนอความสัมพันธ์ระหว่างตัวแปรที่ศึกษา แต่การวิเคราะห์ล็อกลิเนียร์ยังมีข้อแตกต่างจากการวิเคราะห์ความแปรปรวนซึ่งสามารถสรุปได้ดังนี้ (เบ็ญจมาศ แสงอนุเคราะห์, 2541)

1. โมเดลที่ได้จากการวิเคราะห์ความแปรปรวนจะทำนายค่าเฉลี่ยของตัวแปรตามด้วยอิทธิพลของตัวแปรอิสระและเป็นโมเดลแบบบวก (Additive model) ส่วนโมเดลที่ได้จากการวิเคราะห์ล็อกลิเนียร์จะทำนายค่าความถี่คาดหวัง (Expected frequency) ของตัวแปรตามซึ่งอยู่ในรูปของตารางการณ์จรด้วยชุดอิทธิพลของตัวแปรอิสระ และเป็นโมเดลแบบคูณ (Multiplicative model)

2. ในการวิเคราะห์ความแปรปรวน นักวิจัยสามารถสร้างโมเดลได้เพียงโมเดลเดียวเท่านั้น แต่ในการวิเคราะห์ล็อกลิเนียร์ สามารถสร้างโมเดลได้หลายโมเดลเพื่อเลือกใช้โมเดลที่เหมาะสมกับข้อมูลมากที่สุด

3. การวิเคราะห์ความแปรปรวนเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรโดยกำหนดให้ตัวแปรหนึ่งเป็นตัวแปรตามและศึกษาถึงอิทธิพลของชุดของตัวแปรอิสระที่มีต่อตัวแปรตาม ส่วนการวิเคราะห์ล็อกลิเนียร์ สามารถวิเคราะห์ความสัมพันธ์ได้ 2 ลักษณะ คือ การวิเคราะห์ความสัมพันธ์แบบสมมาตร (Symmetrical relationship) เป็นการศึกษาความสัมพันธ์ระหว่างชุดของตัวแปรตั้งแต่ 2 ตัวแปรขึ้นไป โดยไม่ระบุว่าตัวแปรใดเป็นตัวแปรตามและตัวแปรใดเป็นตัวแปรต้น และการวิเคราะห์ความสัมพันธ์แบบอสมมาตร (Asymmetrical relationship) เป็นการศึกษาความสัมพันธ์ระหว่างชุดของตัวแปรตั้งแต่ 2 ตัวแปรขึ้นไป โดยกำหนดให้ตัวแปรหนึ่งเป็นตัวแปรตามและตัวแปรอื่นๆ เป็นตัวแปรต้น จากลักษณะดังกล่าวจะเห็นได้ว่า การวิเคราะห์ความแปรปรวนเป็นการวิเคราะห์ความสัมพันธ์แบบอสมมาตรเท่านั้น

4. ในการวิเคราะห์ความแปรปรวน ตัวแปรตามจะเป็นตัวแปรแบบต่อเนื่อง (Continuous variables) และตัวแปรอิสระเป็นมาตรนามบัญญัติ แต่ในการวิเคราะห์ล็อกลิเนียร์ทั้งตัวแปรตามและตัวแปรอิสระจะเป็นตัวแปรเชิงกลุ่ม โดยอาจเป็นตัวแปรมาตรนามบัญญัติหรือมาตรอันดับก็ได้

ผลจากการบูรณาการนี้ทำให้การวิเคราะห์ล็อกลิเนียร์มีจุดเด่นตรงที่สามารถวิเคราะห์ข้อมูลของตัวแปรที่มีมาตรวัดนามบัญญัติได้พร้อมๆ กันเป็นจำนวนมาก กล่าวคือ สามารถวิเคราะห์ความสัมพันธ์ของตัวแปรอิสระครั้งละหลายตัวแปรได้ หรือข้อมูลที่มีลักษณะเป็นตารางการณ์จรมากกว่า 2 ทาง ซึ่งสามารถเรียกอีกอย่างได้ว่าตารางการณ์จรที่มีหลายมิติ (Multidimensional Contingency tables) ตลอดจนสามารถประมาณพารามิเตอร์ต่างๆ ของตัวแบบ นอกจากนี้ยังใช้ในการทดสอบสมมติฐานทางสถิติ เช่นตรวจสอบความกลมกลืนระหว่างโมเดลกับข้อมูลได้อีกด้วย

จากตารางการณ์จรสองทาง โดยทั่วไปช่องของจำนวนนับในตารางการณ์จรสองทางนั้น จะเป็นตัวแปรเชิงสุ่มปัวส์ซงที่เป็นอิสระต่อกัน (Independent Poisson Random Variable) ซึ่งในการสร้างตัวแบบล็อกลิเนียร์ จำนวนนับในแต่ละช่องจะทำหน้าที่เป็นตัวแปรตาม ส่วนตัวแปรทางด้านแถวและสดมภ์จะทำหน้าที่เป็นตัวแปรอิสระ และได้ว่าค่าเฉลี่ยของจำนวนนับในแต่ละช่องจะเท่ากับค่าคาดหวังของจำนวนนับในช่องนั้นๆ

กำหนดให้ n_i แทนจำนวนนับในช่องที่ i

m_i แทนค่าเฉลี่ยของ n_i ซึ่งเท่ากับ $E(n_i)$, $i = 1, 2, ..., N$

เนื่องจากตัวแบบล็อกลิเนียร์ เป็นส่วนหนึ่งของตัวแบบเชิงเส้นที่วางนัยโดยทั่วไป (GLM) ในการวิเคราะห์ข้อมูลเชิงกลุ่มจะใช้ตัวแปร n_i และ m_i แทนตัวแปร Y_i และ μ_i (หรือ λ_i) ตามลำดับ ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบปัวส์ซง คือ

$$f(n_i;m_i) = \frac{e^{-m_i} m_i^{n_i}}{n_i!}$$

$$= \exp(-m_i)\left(\frac{1}{n_i!}\right)\exp[n_i \ln(m_i)] \quad , \; n_i \geq 0$$

ถ้า Y_i มีรูปแบบการแจกแจงที่ในกลุ่มเอกซ์โปเนนเชียล จะมีฟังก์ชันความน่าจะเป็นดังนี้

$$f(y_i : \theta_i, \phi) = \exp\left[\left(\frac{y_i\theta_i - b(\theta_i)}{a_i(\phi)}\right) + c(y_i, \phi)\right]$$

จากฟังก์ชันของ $f(n_i ; m_i)$ สามารถจัดให้อยู่ในกลุ่มเอ็กซ์โปเนนเชียลได้ดังนี้

$$\log f(n_i; m_i) \quad = n_i \log(m_i) - m_i - \log(n_i!)$$

จึงได้ว่า

$$\theta_i = \log(m_i) \qquad b(\theta_i) = m_i \qquad c(y_i, \phi) = -\log(n_i!) \qquad a(\phi) = 1$$

ดังนั้น Link Function สำหรับการแจกแจงแบบปัวส์ซง คือ $\log(m_i)$ หรือ $\ln(m_i)$ ซึ่งเรียกว่า Log-link

สำหรับการแจกแจงแบบปัวส์ซงนั้นตัวแบบ GLM จะเชื่อมฟังก์ชันของ m_i กับตัวแปรอิสระในลักษณะเชิงเส้นที่มีรูปแบบ คือ

$$\eta = \ln(m_i) = \sum_j \beta_j X_{ij}$$

เราจะเรียกตัวแบบ GLM สำหรับช่องของจำนวนนับของตารางการณ์จรสองทาง หรือ $\ln(m_i) = \sum_j \beta_j X_{ij}$ นี้ว่า ตัวแบบล็อกลิเนียร์ (Log-linear Model)

ลักษณะโดยทั่วไปของข้อมูลที่นำมาใช้ในการวิเคราะห์ด้วยตัวแบบล็อกลิเนียร์ จะมีลักษณะข้อมูลดังตารางที่ 3.1

ตาราง 3.1 ลักษณะโดยทั่วไปของข้อมูลที่นำมาใช้ในการวิเคราะห์ด้วยตัวแบบล็อกลิเนียร์

ตัวแปร X	ตัวแปร Y			
	Y_1	Y_2	…	Y_c
X_1	m_{11}	m_{12}	…	m_{1c}
X_2	m_{21}	m_{22}	…	m_{2c}
$\vdots$	$\vdots$	$\vdots$	m_{ij}	$\vdots$
X_r	m_{r1}	m_{r2}	…	m_{rc}

โดยที่ m_{ij} แทนข้อมูลจำนวนนับ (m_{ij} ซึ่งเป็นตัวแปรตาม) ของ X ระดับที่ i และ Y ระดับที่ j

3.1 ตัวแบบล็อกลิเนียร์แบบง่าย (2 มิติ)

สมมติให้ตัวอย่างขนาด n มีการแจกแจงแบบพหุนาม ประกอบด้วยจำนวน $N = rc$ ช่องในตารางการณ์จรสองทางขนาด $(r \times c)$ โดยกำหนดให้ P_{ij} แทนความน่าจะเป็นแบบพหุนาม หรือแทนการแจกแจงร่วมของตัวแปรเชิงกลุ่มสองตัว และตัวแปรทั้งสองจะเป็นอิสระต่อกัน ก็ต่อเมื่อ

$P_{ij} = P_{i.}P_{.j}$ เมื่อ $i = 1,2,...,r$ และ $j = 1,2,...,c$ จะได้ค่าคาดหวังของ m_{ij} คือ $m_{ij} = nP_{i.}P_{.j}$ ในทุกๆ i และ j และสามารถสร้างตัวแบบล็อกลิเนียร์ที่อาศัยเทอม m_{ij} แทนเทอม P_{ij} โดยการประยุกต์จากคุณสมบัติของการแจกแจงแบบปัวส์ซงที่มีขนาด N ช่อง ซึ่งมีค่าคาดหวังเท่ากับ m_{ij} ดังนั้น ตัวแบบล็อกลิเนียร์แบบง่ายในเทอมของ m_{ij} จึงมีตัวแบบที่น่าสนใจดังต่อไปนี้

3.1.1 ตัวแบบที่เป็นอิสระต่อกัน (Independent Model)

จากเทอม $m_i = nP_{i.}P_{.j}$ เราสามารถสร้างตัวแบบขึ้นใหม่โดยใช้มาตรวัดแบบล็อก (Logarithmic Scale) เพื่อทำให้ความเป็นอิสระต่อกันเป็นรูปแบบเชิงบวก (Additive Form) ซึ่งได้รูปแบบดังนี้

$$\ln(m_{ij}) = \ln(n) + \ln(P_{i.}) + \ln(P_{.j}) \tag{3.1}$$

จากสมการ (3.1) จะได้รูปแบบที่คล้ายกับตัวแบบการวิเคราะห์ความแปรปรวน (ANOVA Models) คือ

$$\ln(m_{ij}) = u + u_i^X + u_j^Y \tag{3.2}$$

โดยที่

$$u_i^X = \ln(P_{i.}) - \frac{\sum_h \ln P_{h.}}{r}$$

$$u_j^Y = \ln(P_{.j}) - \frac{\sum_h \ln P_{.h}}{c}$$

$$u = \ln(n) + \frac{\sum_h \ln P_{h.}}{r} + \frac{\sum_h \ln P_{.h}}{c}$$

และพารามิเตอร์ u_i^X กับ u_j^Y จะสอดคล้องกับเงื่อนไขของ $\sum_{i=1}^{r} u_i^X = \sum_{j=1}^{c} u_j^Y = 0$

สมการ (3.2) มีชื่อเรียกว่า ตัวแบบล็อกลิเนียร์ที่เป็นอิสระต่อกันในตารางการณ์จรสองทาง (Log-linear Model of Independence in A Two-Way Contingency Table)

3.1.2 ตัวแบบเต็มรูป (Saturated Model)

สมมติว่าระหว่างตัวแปรต่างๆ มีความไม่เป็นอิสระต่อกัน และ $m_{ij} = nP_{ij} > 0$ กำหนดให้

$$\lambda_{ij} = \ln(m_{ij})$$

$$\lambda_{i\cdot} = \frac{1}{c}\sum_{j=1}^{c}\lambda_{ij}$$

$$\lambda_{\cdot j} = \frac{1}{r}\sum_{i=1}^{r}\lambda_{ij}$$

$$\lambda_{\cdot\cdot} = \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\lambda_{ij} = u$$

โดยที่ u เป็นค่าเฉลี่ยทั้งหมด (Grand Mean) ของ $\ln\left(m_{ij}\right)$

$$u_i^X = \lambda_{i\cdot} - \lambda_{\cdot\cdot}$$

$$u_j^Y = \lambda_{\cdot j} - \lambda_{\cdot\cdot}$$

$$u_{ij}^{XY} = \lambda_{ij} - \lambda_{i\cdot} - \lambda_{\cdot j} - \lambda_{\cdot\cdot}$$

ดังนั้น ตัวแบบล็อกลิเนียร์สำหรับตารางสองทางคือ

$$\ln\left(m_{ij}\right) = u + u_i^X + u_j^Y + u_{ij}^{XY} \qquad (3.3)$$

จากสมการ (3.3) เรียกว่า ตัวแบบเต็มรูป (Saturated Model) ซึ่งเป็นตัวแบบที่สามารถอธิบายเซตของค่าคาดหวังของความถี่ที่มีค่ามากกว่าศูนย์เซตใดเซตหนึ่งได้อย่างสมบูรณ์ และยังเป็นตัวแบบทั่วไปมากที่สุด (Most General Model) ของตารางการณ์จรแบบสองทาง สำหรับเทอมด้านขวาของสมการ (3.3) จะมีตัวแบบที่คล้ายกับตัวแบบการวิเคราะห์ความแปรปรวนสองทาง (Two – Way ANOVA)

จากพารามิเตอร์ u_i^X และ u_i^Y คือ ค่าที่เบี่ยงเบนจากค่าเฉลี่ย และ $\sum_{i=1}^{r} u_i^X = \sum_{j=1}^{c} u_j^Y = 0$ นั้น ทำให้เหลือพารามิเตอร์ที่เป็นอิสระต่อกันจำนวน $(r-1)$ และ $(c-1)$ เทอม นอกจากนี้ยังได้ว่า u_{ij}^{XY} สอดคล้องกับ

$$\sum_{i=1}^{r} u_{ij}^{XY} = \sum_{j=1}^{c} u_{ij}^{XY} = 0$$

ดังนั้น จึงมีจำนวนพารามิเตอร์ของ u_{ij}^{XY} ที่เป็นอิสระต่อกันจำนวน $(r-1)(c-1)$ เทอม

สำหรับตัวแบบ (3.2) เป็นเพียงกรณีพิเศษของตัวแบบ (3.3) เมื่อ $u_{ij}^{XY} = 0$ การเพิ่มเทอม u_{ij}^{XY} หรือพารามิเตอร์ที่แสดงถึงความสัมพันธ์กันนั้น เพื่อบอกความไม่เป็นอิสระต่อกันระหว่างตัวแปรเชิงกลุ่ม X และ Y

สรุป ตัวแบบที่เป็นอิสระต่อกัน (3.2) จะมีจำนวนพารามิเตอร์ที่เป็นอิสระต่อกันเท่ากับ $1+(r-1)+(c-1) = r+c-1$ เทอม ส่วนตัวแบบเต็มรูป (3.3) จะมีจำนวนพารามิเตอร์ที่เป็นอิสระต่อกันเท่ากับ $1+(r-1)+(c-1)+(r-1)(c-1) = rc$ เทอม โดยทั่วไปตารางการณ์จรสองทาง จำนวนพารามิเตอร์ในตัวแบบเต็มรูปจะเท่ากับจำนวนของช่องในตารางข้อมูลนั้น ถ้าตัวแบบมีจำนวนพารามิเตอร์น้อยกว่าจำนวนช่องของตารางข้อมูล จะเรียกตัวแบบดังกล่าวว่า ตัวแบบไม่เต็มรูป (Unsaturated Model)

3.1.3 การตีความหมายของพารามิเตอร์ (Interpretation of Parameters)

การตีความหมายของพารามิเตอร์จะไม่ซับซ้อนสำหรับตัวแปรแบบ Binary เช่น ตัวแบบที่เป็นอิสระต่อกันสำหรับตาราง $(r \times 2)$ จะมีโลจิทของตัวแปร Binary เท่ากับ

$$\ln\left(\frac{P_{1|i}}{P_{2|i}}\right) = \ln\left(\frac{m_{i1}}{m_{i2}}\right)$$

$$= \ln(m_{i1}) - \ln(m_{i2})$$

$$= (u + u_i^X + u_1^Y) - (u + u_i^X + u_2^Y)$$

$$= u_1^Y - u_2^Y$$

โลจิทดังกล่าวจะมีค่าเท่ากับ $u_1^Y - u_2^Y$ ในทุกๆ แถว สำหรับเงื่อนไขของผลรวมพารามิเตอร์เท่ากับ 0 นั้นพบว่า โลจิทที่ i เท่ากับ $2u_i^Y$ เนื่องจาก $u_2^Y = -u_1^Y$ นอกจากนี้ในแต่ละแถวจะได้ว่า $\exp(2u_i^Y)$ คือ ออดส์ (Odds) ของการจัดกลุ่มของสดมภ์ให้เป็นกลุ่ม 1 แทนที่จะเป็นกลุ่ม 2

ความสัมพันธ์ระหว่างอัตราส่วนออดส์ และพารามิเตอร์ที่แสดงความสัมพันธ์โดยยกตัวอย่างตาราง (2×2) โดยพิจารณาจากตัวแบบเต็มรูป พบว่า

$$\ln(\theta) = \ln\left(\frac{m_{11}m_{22}}{m_{12}m_{21}}\right)$$

$$= \ln(m_{11}) + \ln(m_{22}) - \ln(m_{12}) - \ln(m_{21})$$

$$= (u + u_1^X + u_1^Y + u_{11}^{XY}) + (u + u_2^X + u_2^Y + u_{22}^{XY})$$
$$- (u + u_1^X + u_2^Y + u_{12}^{XY}) - (u + u_2^X + u_1^Y + u_{21}^{XY})$$

$$= u_{11}^{XY} + u_{22}^{XY} - u_{12}^{XY} - u_{21}^{XY}$$

ภายใต้เงื่อนไข $\sum_{i=1}^{r} u_{ij}^{XY} = \sum_{j=1}^{c} u_{ij}^{XY} = 0$ จะทำให้ $u_{11}^{XY} = u_{22}^{XY} = -u_{12}^{XY} = -u_{21}^{XY}$

จะได้ว่า $\ln(\theta) = 4u_{11}^{XY}$

ดังนั้น อัตราส่วนออดส์สำหรับตาราง (2×2) จึงเท่ากับ Antilog ของ 4 เท่าของพารามิเตอร์ที่แสดงความสัมพันธ์ในตัวแบบเต็มรูป

3.1.4 ตัวแบบความน่าจะเป็นของเซลล์ (Model of Cell Probabilities)

ตัวแบบที่มีการเลือกตัวอย่างแบบปัวส์ซง จำนวนนับของเซลล์ที่ได้จึงเป็นตัวแปรเชิงสุ่มแบบปัวส์ซงที่เป็นอิสระต่อกันด้วยค่าเฉลี่ย m_{ij} พารามิเตอร์โดยธรรมชาติของการแจกแจงแบบปัวส์ซอง คือ ล็อกของค่าเฉลี่ย [$\ln(m_{ij})$]

จากเงื่อนไขของตัวอย่างขนาด n ตัวแบบล็อกลิเนียร์แบบปัวส์ซงสำหรับเซต m_{ij} จะมีนิพจน์ที่สมมูลกับตัวแบบล็อกลิเนียร์แบบพหุนามสำหรับเซต $P_{ij} = \dfrac{m_{ij}}{\sum\sum m_{ab}}$ ซึ่งเป็นความน่าจะเป็นของตัวอย่างที่ i และ j ดังสมการต่อไปนี้

จากสมการ $\ln(m_{ij}) = u + u_i^X + u_j^Y + u_{ij}^{XY}$

$$m_{ij} = \exp(u + u_i^X + u_j^Y + u_{ij}^{XY})$$

$$\therefore \quad P_{ij} = \frac{\exp(u + u_i^X + u_j^Y + u_{ij}^{XY})}{\sum_a \sum_b \exp(u + u_a^X + u_b^Y + u_{ab}^{XY})}$$

โดยที่ $P_{ij} \geq 0$ และ $\sum\sum P_{ij} = 1$

3.2 ตัวแบบล็อกลิเนียร์เชิงชั้น

3.2.1 ลักษณะและคุณสมบัติของตารางการณ์จรสามทาง

ตารางการณ์จรสามทาง (Three-way contingency table) คือตารางแจกแจงความถี่ของข้อมูลตามคุณลักษณะตัวแปรเชิงกลุ่มที่สนใจจำนวน 3 ตัวแปร สมมติว่าตัวแปรเชิงกลุ่มที่สนใจศึกษามี 3 ตัวแปร คือตัวแปร X, Y และ Z โดยที่ตัวแปร X แบ่งออกเป็น r กลุ่ม ตัวแปร Y แบ่งออกเป็น c กลุ่ม และตัวแปร Z แบ่งออกเป็น l กลุ่ม ซึ่งแต่ละตัวแปรอาจอยู่ในมาตรวัดแบบนามบัญญัติหรือแบบเรียงลำดับหรือแบบอันตรภาคชั้นก็ได้ และหากจัดหมวดหมู่ตามลักษณะของปัจจัยดังกล่าวจะสามารถจำแนกได้ถึง $r \times c \times l$ กลุ่ม ซึ่งแสดงถึงขนาดของตารางการณ์จรหรือจำนวนช่องทั้งหมด

ให้ X_{ijk} เป็นตัวแปรสุ่มซึ่งใช้แทนความถี่ของจำนวนนับที่สอดคล้องกับกลุ่มที่ i ของตัวแปร X เมื่อ $i = 1, 2, ..., r$ กลุ่มที่ j ของตัวแปร Y เมื่อ $j = 1, 2, ..., c$ และกลุ่มที่ k ของตัวแปร Z เมื่อ $k = 1, 2, ..., l$ หรือเป็นจำนวนหน่วยตัวอย่างที่ตกอยู่ในช่องที่ ijk โดยสัญลักษณ์ ijk แทนตำแหน่งของช่องในตารางการณ์จร ดังแสดงในตารางที่ 3.2

ตาราง 3.2 ลักษณะทั่วไปของตารางการณ์จรสามทาง ขนาด $r \times c \times l$ ที่ประกอบด้วยความถี่ในแต่ละเซลล์

ตัวแปร X	ตัวแปร Z									รวม
	Z_1 ตัวแปร Y			… ตัวแปร Y			Z_l ตัวแปร Y			
	Y_1	…	Y_c	Y_1	…	Y_c	Y_1	…	Y_c	
X_1	x_{111}	…	x_{1c1}	x_{11k}	…	x_{1ck}	x_{11l}	…	x_{1cl}	$n_{1\cdot\cdot}$
X_2	x_{211}	…	x_{2c1}	x_{21k}	…	x_{2ck}	x_{21l}	…	x_{2cl}	$n_{2\cdot\cdot}$
⋮	⋮	x_{ij1}	⋮	⋮	x_{ijk}	⋮	⋮	x_{ijl}	⋮	⋮
X_r	x_{r11}	…	x_{rc1}	x_{r1k}	…	x_{rck}	x_{r1l}	…	x_{rcl}	$n_{r\cdot\cdot}$
รวม	$n_{\cdot 11}$	$n_{\cdot j1}$	$n_{\cdot c1}$	$n_{\cdot 1k}$	$n_{\cdot jk}$	$n_{\cdot ck}$	$n_{\cdot 1l}$	$n_{\cdot jl}$	$n_{\cdot cl}$	$n_{\cdot\cdot\cdot} = n$
		$n_{\cdot\cdot 1}$			$n_{\cdot\cdot k}$			$n_{\cdot\cdot l}$		

$n_{\cdot 1 \cdot}$ $n_{\cdot j \cdot}$ $n_{\cdot c \cdot}$

เมื่อ $n_{i\cdot\cdot} = \sum_{j=1}^{c}\sum_{k=1}^{l} x_{ijk}$, $n_{\cdot j\cdot} = \sum_{i=1}^{r}\sum_{k=1}^{l} x_{ijk}$, $n_{\cdot\cdot k} = \sum_{i=1}^{r}\sum_{j=1}^{c} x_{ijk}$

และ $\sum_{j=1}^{c}\sum_{k=1}^{l} n_{i\cdot\cdot}, = \sum_{i=1}^{r}\sum_{k=1}^{l} n_{\cdot j\cdot} = \sum_{i=1}^{r}\sum_{j=1}^{c} n_{\cdot\cdot k} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l} x_{ijk} = n_{\cdots} = n$

ให้ p_{ijk} แทนความน่าจะเป็นที่ตัวแปรสุ่ม X_{ijk} จะตกอยู่ในเซลล์ ijk หรือความน่าจะเป็นที่ตัวแปรสุ่ม X_{ijk} จะมีลักษณะสอดคล้องกับตัวแปร X ในกลุ่มที่ i ตัวแปร Y ในกลุ่มที่ j และตัวแปร Z ในกลุ่มที่ k ทั้งนี้ p_{ijk} ก็สามารถแสดงในตารางการณ์จรเช่นเดียวกับ X_{ijk} แต่มีคุณสมบัติที่แตกต่างกันดังนี้

$$p_{i\cdot\cdot} = \sum_{j=1}^{c}\sum_{k=1}^{l} p_{ijk}, \qquad p_{\cdot j\cdot} = \sum_{i=1}^{r}\sum_{k=1}^{l} p_{ijk}, \qquad p_{\cdot\cdot k} = \sum_{i=1}^{r}\sum_{j=1}^{c} p_{ijk}$$

และ $\sum_{j=1}^{c}\sum_{k=1}^{l} p_{i\cdot\cdot} = \sum_{i=1}^{r}\sum_{k=1}^{l} p_{\cdot j\cdot} = \sum_{i=1}^{r}\sum_{j=1}^{c} p_{\cdot\cdot k} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l} p_{ijk} = p_{\cdots} = 1$

จาก p_{ijk} ตัวแปรทั้งสามจะเป็นอิสระต่อกันก็ต่อเมื่อ $p_{ijk} = p_{i\cdot\cdot}.p_{\cdot j\cdot}.p_{\cdot\cdot k}$ จะได้ค่าคาดหวังคือ

$$m_{ijk} = np_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k} = np_{ijk} = n\left(\frac{x_{ijk}}{n}\right) = x_{ijk} \quad \text{ในทุกๆ } i, j \text{ และ } k$$

ดังนั้นจะสามารถสร้างตารางการณ์จรของ m_{ijk} ได้เช่นเดียวกับ X_{ijk} และมีคุณสมบัติดังนี้

$$m_{i\cdot\cdot} = \sum_{j=1}^{c}\sum_{k=1}^{l} m_{ijk} = \sum_{j=1}^{c}\sum_{k=1}^{l} np_{ijk} = np_{i\cdot\cdot} = n_{i\cdot\cdot}$$

$$m_{\cdot j\cdot} = \sum_{i=1}^{r}\sum_{k=1}^{l} m_{ijk} = \sum_{i=1}^{r}\sum_{k=1}^{l} np_{ijk} = np_{\cdot j\cdot} = n_{\cdot j\cdot}$$

$$m_{\cdot\cdot k} = \sum_{i=1}^{r}\sum_{j=1}^{c} m_{ijk} = \sum_{i=1}^{r}\sum_{j=1}^{c} np_{ijk} = np_{\cdot\cdot k} = n_{\cdot\cdot k}$$

และ $m_{\cdots} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l} m_{ijk} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l} np_{ijk} = n\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l} p_{ijk} = n$

จากตารางการณ์จรสามทาง เราอาจสนใจศึกษาตัวแปรเพียง 2 ตัวแปรใดๆ ในกลุ่มย่อยใดกลุ่มย่อยหนึ่งของตัวแปรที่ 3 ก็ได้ เช่นต้องการศึกษาตัวแปร X และตัวแปร Y ภายใต้กลุ่มย่อยที่ 1 ของตัวแปร Z ซึ่งตารางการณ์จรดังกล่าวจะกลายเป็นตารางการณ์จรสองทางที่เรียกว่า ตารางการณ์

จรบางส่วน (Partial table) และจะเรียกความสัมพันธ์ของตัวแปรเชิงกลุ่มระหว่างตัวแปร X และตัวแปร Y ว่าเป็นความสัมพันธ์ร่วม (Partial association) และหากระดับความสัมพันธ์ระหว่างตัวแปร X และตัวแปร Y มีความแตกต่างในแต่ละกลุ่มย่อยของตัวแปร Z แล้ว จะสรุปได้ว่าเกิดความสัมพันธ์ร่วม 3 ตัวแปร (Three-way interaction) ระหว่าง X, Y และ Z

ตาราง 3.3 ตารางการณ์จรบางส่วนที่สนใจศึกษาตัวแปร X และตัวแปร Y ภายใต้กลุ่มย่อยที่ 1 ของตัวแปร Z

ตัวแปร X	ตัวแปร Z_1			รวม
	ตัวแปร Y_1	...	ตัวแปร Y_c	
X_1	x_{111}	…	x_{1c1}	$n_{1\cdot1}$
X_2	x_{211}	…	x_{2c1}	$n_{2\cdot1}$
$\vdots$	$\vdots$	x_{ij1}	$\vdots$	$\vdots$
X_r	x_{r11}	…	x_{rc1}	$n_{r\cdot1}$
รวม	$n_{\cdot11}$	$n_{\cdot j1}$	$n_{\cdot c1}$	$n_{\cdot\cdot1}$

<u>ความผิดพลาดในการใช้ตารางการณ์จรสองทางแทนตารางการณ์จรสามทาง</u>

ในบางครั้งถ้าเราสนใจศึกษาตัวแปรเพียง 2 ตัวแปรใดๆ เราอาจทำการยุบ (Collapse) ตัวแปรที่เราไม่สนใจนั้นทิ้งเสีย กล่าวคือเป็นการยุบตารางการณ์จรสามทาง (ตารางที่ 3.2) ให้เหลือเพียงสองทาง เรียกว่าตารางการณ์จรแบบชายขอบ (Marginal table) ซึ่งมีจำนวนช่องเท่ากับผลคูณของจำนวนกลุ่มย่อยของตัวแปรทั้งสอง เช่นตารางการณ์จรแบบชายขอบของตัวแปร X และตัวแปร Y เมื่อไม่สนใจตัวแปร Z จะแสดงได้ดังตารางที่ 3.4 และจะเรียกความสัมพันธ์ของตัวแปรเชิงกลุ่มระหว่างตัวแปร X และตัวแปร Y ว่าเป็นความสัมพันธ์ชายขอบ (Marginal association)

ตาราง 3.4 ตารางการณ์จรแบบชายขอบที่สนใจศึกษาเพียงตัวแปร X และตัวแปร Y โดยไม่สนใจศึกษาตัวแปร Z

ตัวแปร X	ตัวแปร Y			รวม
	Y_1	…	Y_c	
X_1	$x_{11\cdot}$	…	$x_{1c\cdot}$	$n_{1\cdot\cdot}$
X_2	$x_{21\cdot}$	…	$x_{2c\cdot}$	$n_{2\cdot\cdot}$
⋮	⋮	$x_{ij\cdot}$	⋮	⋮
X_r	$x_{r1\cdot}$	…	$x_{rc\cdot}$	$n_{r\cdot\cdot}$
รวม	$n_{\cdot 1\cdot}$	$n_{\cdot j\cdot}$	$n_{\cdot c\cdot}$	$n_{\cdot\cdot\cdot} = n$

อย่างไรก็ตามการยุบตัวแปรจากตารางการณ์จรสามทาง ให้เหลือเพียงสองทางนั้น อาจทำให้ข้อสรุปจากการวิเคราะห์ข้อมูลมีความผิดพลาดได้ ซึ่งถือว่าเป็นกรณีศึกษาที่เรียกว่า ปฏิทรรศน์ของซิมป์สัน (Simpson’s paradox) (Agresti: 2002) แสดงได้ดังนี้

ตาราง 3.5 ความถี่ในตารางการณ์จรสามทาง กรณีเกิดปฏิทรรศน์ของซิมป์สัน

ปัจจัย	น้ำหนักตัวเกินมาตรฐาน		น้ำหนักตัวไม่เกินมาตรฐาน		รวม
	ชาย	หญิง	ชาย	หญิง	
เป็นโรคหัวใจ	20	10	15	50	95
ไม่เป็นโรคหัวใจ	50	40	5	40	135
รวม	70	50	20	90	230

จากตารางที่ 3.5 เมื่อพิจารณาตามตัวแปร “น้ำหนักตัวเกินมาตรฐาน” จะพบว่าความน่าจะเป็นที่ผู้ชายจะเป็นโรคหัวใจ เท่ากับ 20/70 = 0.286 และความน่าจะเป็นที่ผู้หญิงจะเป็นโรคหัวใจเท่ากับ 10/50 = 0.2 เมื่อพิจารณาตามตัวแปร “น้ำหนักตัวไม่เกินมาตรฐาน” จะพบว่าความน่าจะเป็นที่ผู้ชายจะเป็นโรคหัวใจ เท่ากับ 15/20 = 0.75 และความน่าจะเป็นที่ผู้หญิงจะเป็นโรคหัวใจ เท่ากับ 50/90 = 0.556 ดังนั้นจะพบว่าผู้ชายจะมีโอกาสเป็นโรคหัวใจมากกว่าผู้หญิงในแต่ละตัวแปรด้านน้ำหนัก ถ้าทำการยุบตัวแปรด้านน้ำหนัก ให้เหลือเพียงตัวแปรเพศและตัวแปรการเป็นโรคหัวใจเท่านั้น จะได้ผลดังตารางที่ 3.6

ตาราง 3.6 ความถี่ในตารางการณ์จรสองทางที่ยุบตัวแปรด้านน้ำหนัก กรณีเกิดปฏิทรรศน์ของซิมป์สัน

ปัจจัย	ชาย	หญิง	รวม
เป็นโรคหัวใจ	35	60	95
ไม่เป็นโรคหัวใจ	55	80	135
รวม	90	140	230

จากตารางที่ 3.6 เมื่อทำการยุบตัวแปรด้านน้ำหนักแล้ว จะพบว่าความน่าจะเป็นที่ผู้ชายจะเป็นโรคหัวใจ เท่ากับ 35/90 = 0.389 ในขณะที่ความน่าจะเป็นที่ผู้หญิงจะเป็นโรคหัวใจ เท่ากับ 60/140 = 0.429 ซึ่งมีโอกาสการเป็นโรคหัวใจมากกว่าผู้ชาย ซึ่งผลสรุปจากตารางที่ 3.6 ขัดแย้งกับตารางที่ 3.5 ดังนั้นการยุบตัวแปรในกรณีนี้จึงให้ข้อสรุปผิดไปจากความจริง

3.2.1.1 รูปแบบของความเป็นอิสระต่อกันระหว่างตัวแปรเชิงกลุ่มในตารางการณ์จรสามทาง

1) ความเป็นอิสระอย่างสมบูรณ์ (Mutual or Complete Independence)

ตัวแปรเชิงกลุ่มในตารางการณ์จรสามทาง จะเป็นอิสระต่อกันอย่างสมบูรณ์ก็ต่อเมื่อ

$$p_{ijk} = p_{i\cdot\cdot} \cdot p_{\cdot j\cdot} \cdot p_{\cdot\cdot k} \qquad \text{ทุก } i, j \text{ และ } k$$

ซึ่งหมายความว่าความน่าจะเป็นที่ตัวแปรสุ่ม X_{ijk} จะตกอยู่ในเซลล์ ijk ใดๆ ของตัวแปร X, Y และ Z จะมีค่าเท่ากับผลคูณของความน่าจะเป็นชายขอบของตัวแปร X, Y และ Z นั้นๆ โดยรูปแบบของความเป็นอิสระอย่างสมบูรณ์นี้มีได้เพียงรูปแบบเดียวเท่านั้น ไม่มีกรณีย่อย

2) ความเป็นอิสระร่วม (Joint or Partial Independence)

ตัวแปรเชิงกลุ่มในตารางการณ์จรสามทาง จะเป็นอิสระต่อร่วมก็ต่อเมื่อ ความน่าจะเป็นที่ตัวแปรสุ่ม X_{ijk} จะตกอยู่ในเซลล์ ijk ใดๆ ของตัวแปร X, Y และ Z จะมีค่าเท่ากับผลคูณของความน่าจะเป็นชายขอบของ 2 ตัวแปรที่มีความสัมพันธ์กัน กับความน่าจะเป็นชายขอบของตัวแปรที่ 3 ที่เป็นอิสระจาก 2 ตัวแปรนั้น เช่น ตัวแปร X, Y จะเป็นอิสระร่วมกับตัวแปร Z ก็ต่อเมื่อ

$$p_{ijk} = p_{ij\cdot} \cdot p_{\cdot\cdot k} \qquad \text{ทุก } i, j \text{ และ } k$$

นอกจากนี้ในกรณีอื่นก็จะหาได้โดยใช้หลักการเดียวกัน ดังนี้

ตัวแปร X, Z จะเป็นอิสระร่วมกับตัวแปร Y ก็ต่อเมื่อ

$$p_{ijk} = p_{i\cdot k} \cdot p_{\cdot j\cdot} \qquad \text{ทุก } i, j \text{ และ } k$$

ตัวแปร Y, Z จะเป็นอิสระร่วมกับตัวแปร X ก็ต่อเมื่อ

$$p_{ijk} = p_{\cdot jk} \cdot p_{i\cdot\cdot} \qquad \text{ทุก } i, j \text{ และ } k$$

3) ความเป็นอิสระอย่างมีเงื่อนไข (Conditional Independence)

ตัวแปรเชิงกลุ่มในตารางการณ์จรสามทาง จะเป็นอิสระต่อกันอย่างมีเงื่อนไขก็ต่อเมื่อความน่าจะเป็นที่ตัวแปรสุ่ม X_{ijk} จะตกอยู่ในเซลล์ ijk ใดๆ ของตัวแปร X, Y และ Z จะแสดงได้ดังนี้ เช่น กรณีตัวแปร X และ Y เป็นอิสระต่อกันอย่างมีเงื่อนไขในตัวแปร Z พิจารณาได้ดังนี้

$$p_{ij|k} = \frac{p_{ijk}}{p_{\cdot\cdot k}}$$

เมื่อกำหนดตัวแปร X และ Y เป็นอิสระต่อกันอย่างมีเงื่อนไขในตัวแปร Z จึงได้ว่า

$$p_{i\cdot|k} \cdot p_{\cdot j|k} = \frac{p_{ijk}}{p_{\cdot\cdot k}}$$

ดังนั้น ตัวแปร X และ Y จะเป็นอิสระต่อกันอย่างมีเงื่อนไขในตัวแปร Z ก็ต่อเมื่อ

$$p_{ijk} = p_{\cdot\cdot k}\left(\frac{p_{i\cdot k}}{p_{\cdot\cdot k}}\right)\left(\frac{p_{\cdot jk}}{p_{\cdot\cdot k}}\right)$$

$$= \frac{p_{i\cdot k} \cdot p_{\cdot jk}}{p_{\cdot\cdot k}} \qquad \text{ทุก } i, j \text{ และ } k$$

นอกจากนี้ในกรณีอื่นก็จะหาได้โดยใช้หลักการเดียวกัน ดังนี้

ตัวแปร X และ Z จะเป็นอิสระอย่างมีเงื่อนไขในตัวแปร Y ก็ต่อเมื่อ

$$p_{ijk} = \frac{p_{ij\cdot} \cdot p_{\cdot jk}}{p_{\cdot j\cdot}} \qquad \text{ทุก } i, j \text{ และ } k$$

ตัวแปร Y และ Z จะเป็นอิสระอย่างมีเงื่อนไขในตัวแปร X ก็ต่อเมื่อ

$$p_{ijk} = \frac{p_{ij\cdot} \cdot p_{i\cdot k}}{p_{i\cdot\cdot}} \qquad \text{ทุก } i, j \text{ และ } k$$

4) ความเป็นอิสระที่ไม่มีตัวแปรคู่ใดเป็นอิสระต่อกัน (Mutual Dependence)

เป็นรูปแบบที่ไม่มีพจน์ความสัมพันธ์ร่วมของ 3 ตัวแปร (No-three-way interaction) ความเป็นอิสระในลักษณะนี้มีความซับซ้อนไม่มีรูปแบบที่ชัดเจน แต่มีหลักกว้างๆ เพียงว่า ตัวแปรทุกคู่จะต้องมีความสัมพันธ์กันและระดับความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปรใดๆ จะต้องมีค่า

เท่ากันในทุกระดับของตัวแปรที่ 3 ซึ่งรูปแบบความสัมพันธ์ทุกรูปแบบที่กล่าวมาข้างต้นก็ล้วนแล้วแต่มีระดับความสัมพันธ์ 2 ตัวแปรใดๆ เท่ากันในทุกระดับของปัจจัยที่ 3 เช่นเดียวกัน

จากรูปแบบความสัมพันธ์พบว่ามีเพียง 3 รูปแบบแรกเท่านั้น มีลักษณะเป็น รูปแบบปิด (Closed form) กล่าวคือสามารถหารูปแบบที่ชัดเจนได้ ในขณะที่รูปแบบความเป็นอิสระที่มี ไม่มีตัวแปรคู่ใดเป็นอิสระต่อกันนั้นไม่สามารถหารูปแบบความสัมพันธ์ที่ชัดเจนได้

3.2.1.2 การตีความหมายของตัวแบบพารามิเตอร์ (Interpretation of Model Parameters)

การตีความหมายของพารามิเตอร์ของตัวแบบล็อกลิเนียร์ จะใช้อัตราส่วนออดส์มาอธิบายตารางแบบชายขอบและความสัมพันธ์บางส่วน เช่นตารางแบบชายขอบของ X กับ Y (X–Y Marginal Table) เมื่อกำหนดให้ระดับ k ของ Z คงที่ จะมีอัตราส่วนออดส์จำนวน $(r-1)(c-1)$ ค่า ซึ่งมีสัญลักษณ์ดังนี้

$$\theta_{ij(k)} = \frac{P_{ijk}\ P_{i+1,j+1,k}}{P_{i,j+1,k}\ P_{i+1,j,k}} \quad \text{โดยที่ } 1 \le i \le r-1 \text{ และ } 1 \le j \le c-1$$

ซึ่งอธิบายถึง ความสัมพันธ์แบบมีเงื่อนไข (Conditional Association) ของ X และ Y ในทำนองเดียวกันความสัมพันธ์แบบมีเงื่อนไขของ X และ Z เมื่อกำหนดให้ระดับ j ของ Y คงที่ก็จะอธิบายได้ด้วยอัตราส่วนออดส์จำนวน $(r-1)(l-1)$ ค่า หรือ $(\theta_{i(j)k})$ และความสัมพันธ์แบบมีเงื่อนไขของ Y และ Z เมื่อกำหนดให้ระดับที่ i ของ X คงที่ ก็จะอธิบายได้ด้วยอัตราส่วนออดส์จำนวน $(c-1)(l-1)$ ค่า (หรือ $\theta_{(i)jk}$)

การหาค่าของพารามิเตอร์ของตัวแบบล็อกลิเนียร์สามารถหาได้จากฟังก์ชันของอัตราส่วนออดส์แบบมีเงื่อนไข เช่นการหาพารามิเตอร์สำหรับตารางการณ์จรขนาด 2×2×2 สามารถหาได้โดยการแทนค่า $\ln(m_{ijk})$ ของตัวแบบเต็มรูปสำหรับตารางการณ์จรสามทาง ลงในล็อกอัตราส่วนออดส์แบบมีเงื่อนไข (Log Conditional Odds Ratios) จะได้ผลลัพธ์ดังที่ Agresti (2002) แสดงไว้คือ

$$u_{111}^{XYZ} = \frac{1}{8}\ln\left(\frac{\theta_{11(1)}}{\theta_{11(2)}}\right)$$

$$= \frac{1}{8}\ln\left(\frac{\theta_{1(1)1}}{\theta_{1(2)1}}\right)$$

$$= \frac{1}{8}\ln\left(\frac{\theta_{(1)11}}{\theta_{(2)11}}\right)$$

สำหรับเงื่อนไขที่เกี่ยวกับ $\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l} u_{ijk}^{XYZ} = 0$ นั้นในแต่ละเทอมของ u_{ijk}^{XYZ} จะมีค่าเป็นศูนย์ก็ต่อเมื่อ อัตราส่วนออดส์ระหว่างตัวแปร 2 ตัวมีค่าคงที่ ณ ทุกระดับของตัวแปรที่ 3 เช่น จากตารางการณ์จรขนาด 2×2×k พบว่า $u_{ijk}^{XYZ} = 0$ เมื่อ $\theta_{11(1)} = ... = \theta_{11(k)}$ ดังนั้น

$$u_{11}^{XY} = \frac{1}{4}\ln\left(\theta_{11(k)}\right) \qquad \text{สำหรับ } k = 1, 2, ..., l$$

ซึ่งสอดคล้องกับกรณีของตาราง 2 มิติ [$\ln\left(\theta_{ij}\right) = 4u_{11}^{XY}$] คือ พารามิเตอร์ที่แสดงความสัมพันธ์นั้นเป็นสัดส่วนกับล็อกของอัตราส่วนออดส์

สรุป ตัวแบบล็อกลิเนียร์สามารถแสดงคุณสมบัติบางอย่างได้โดยใช้อัตราส่วนออดส์ เช่น ความเป็นอิสระต่อกันแบบมีเงื่อนไขของ X และ Y ซึ่งสมมูลกับ $\theta_{ij(k)} = 1$, $i = 1, 2, ..., r-1$ $j = 1, 2, ..., c-1$ และ $k = 1, 2, ..., l-1$

3.2.1.3 เงื่อนไขที่พอเพียงบางประการ

เงื่อนไขที่พอเพียงที่ทำให้อัตราส่วนออดส์ สำหรับ X และ Y มีค่าเท่ากันทั้งในตารางบางส่วน และตารางแบบชายขอบ เมื่ออัตราส่วนออดส์ดังกล่าวเท่ากัน จะช่วยให้เราศึกษาความสัมพันธ์ของ X และ Y ได้สะดวกขึ้นโดยการยุบมิติ (Collapsing) ของ Z โดยที่ Z อาจเป็นตัวแปรเดียวหรือหลายตัวแปร (Multidimensional) เงื่อนไขเหล่านั้นจึงเรียกว่าเงื่อนไขการยุบได้ (Collapsibility Conditions) โดยเฉพาะบางเงื่อนไขที่เกี่ยวกับเซตของอัตราส่วนออดส์ จะมีค่าเท่ากัน คือ

$$\theta_{ij}^{XY} = \theta_{ij(1)} = \theta_{ij(2)} = ... = \theta_{ij(k)} \quad \text{เมื่อ } 1 \le i \le r-1 \text{ และ } 1 \le j \le c-1$$

และถ้าข้อหนึ่งข้อใดหรือทั้ง 2 ข้อ ต่อไปนี้เป็นจริง

1. $\theta_{i(j)k} = 1$ เมื่อ $1 \le i \le r-1$ และ $1 \le j \le c-1$ และ $1 \le k \le l-1$
2. $\theta_{(i)jk} = 1$ เมื่อ $1 \le i \le r-1$ และ $1 \le j \le c-1$ และ $1 \le k \le l-1$

กล่าวคือ ความสัมพันธ์แบบชายขอบ และความสัมพันธ์บางส่วนของ X กับ Y (XY – Marginal and Partial Association) จะมีค่าเท่ากัน ถ้า X และ Z เป็นอิสระต่อกันแบบมีเงื่อนไข หรือ ถ้า Y และ Z เป็นอิสระต่อกันแบบมีเงื่อนไข และตัวแบบล็อกลิเนียร์สำหรับ m_{ij} ก็จะมีเทอมความสัมพันธ์ของพารามิเตอร์ u_{ij}^{XY} เท่ากัน และมีอัตราส่วนออดส์เท่ากันด้วย

3.2.1.4 ตัวแบบล็อกลิเนียร์เชิงชั้นสำหรับตารางการณ์จรสามทาง
(Hierarchical log-linear models for three-way contingency table)

ตัวแบบล็อกลิเนียร์มีรูปแบบโครงสร้างทั่วไปคล้ายคลึง กับตัวแบบการวิเคราะห์ความแปรปรวน โดยตัวแบบล็อกลิเนียร์สำหรับตารางการณ์จรสามทาง มีพื้นฐานมาจากแนวคิดในเรื่องรูปแบบของความเป็นอิสระต่อกันระหว่างตัวแปรเชิงกลุ่ม ดังนั้นจึงปรากฏตัวแบบย่อยขึ้นตามรูปแบบความเป็นอิสระ รวมทั้งยังมีตัวแบบเต็มรูปซึ่งเป็นตัวแบบที่รวมอิทธิพลต่างๆ ที่เป็นไปได้ทั้งหมดและตัวแบบย่อยที่มีรูปแบบปิดของความน่าจะเป็นในแต่ละช่อง (p_{ijk}) ซึ่งมีดังต่อไปนี้

1. ตัวแบบเต็มรูป (Saturated Model)

เป็นตัวแบบที่รวมพจน์ของอิทธิพลหรือความสัมพันธ์ทั้งหมดที่เป็นไปได้ จะใช้สัญลักษณ์ [XYZ] แทนตัวแบบเต็มรูป ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ij}^{XY} + u_{ik}^{XZ} + u_{jk}^{YZ} + u_{ijk}^{XYZ} \quad (3.4)$$

โดย m_{ijk} แทนความถี่คาดหวังของเซลล์ที่ ijk ในตารางการณ์จรซึ่งตรงกับกลุ่มย่อยที่ i ของตัวแปร X , กลุ่มย่อยที่ j ของตัวแปร Y และกลุ่มย่อยที่ k ของตัวแปร Z โดยที่ $m_{ijk} > 0$

u แทนค่าคงที่หรือเป็นค่าเฉลี่ยทั้งหมด (Grand Mean)

u_i^X แทนอิทธิพลหลักของตัวแปร X ในกลุ่มย่อยที่ i (X_i main effect)

u_j^Y แทนอิทธิพลหลักของตัวแปร Y ในกลุ่มย่อยที่ j (Y_j main effect)

u_k^Z แทนอิทธิพลหลักของตัวแปร Z ในกลุ่มย่อยที่ k (Z_k main effect)

u_{ij}^{XY} แทนอิทธิพลร่วมของตัวแปร X และ Y ในกลุ่มย่อยที่ i และ j ตามลำดับ ($(XY)_{ij}$ partial interaction)

u_{ik}^{XZ} แทนอิทธิพลร่วมของตัวแปร X และ Z ในกลุ่มย่อยที่ i และ k ตามลำดับ (($(XZ)_{ik}$ partial interaction)

u_{jk}^{YZ} แทนอิทธิพลร่วมของตัวแปร Y และ Z ในกลุ่มย่อยที่ j และ k ตามลำดับ ($(YZ)_{jk}$ partial interaction)

u_{ijk}^{XYZ} แทนอิทธิพลร่วมของตัวแปร X, Y และ Z ในกลุ่มย่อยที่ i, j และ k ตามลำดับ (Three - factor interaction)

ในอีกนัยหนึ่ง พารามิเตอร์ในตัวแบบก็หมายถึง ความแตกต่างจากค่าเฉลี่ยของตัวแบบนั่นเอง ซึ่งในทุกตัวแบบก็จะใช้หลักการเดียวกัน ดังนี้

$$u = \frac{1}{rcl}\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}\ln m_{ijk} = \lambda_{...}$$

$$u_i^X = \frac{1}{cl}\sum_{j=1}^{c}\sum_{k=1}^{l}\ln m_{ijk} - u = \lambda_{i..} - \lambda_{...}$$

$$u_j^Y = \frac{1}{rl}\sum_{i=1}^{r}\sum_{k=1}^{l}\ln m_{ijk} - u = \lambda_{.j.} - \lambda_{...}$$

$$u_k^Z = \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\ln m_{ijk} - u = \lambda_{..k} - \lambda_{...}$$

$$u_{ij}^{XY} = \frac{1}{l}\sum_{k=1}^{l}\ln m_{ijk} - u_i^X - u_j^Y - u = \lambda_{ij.} - \lambda_{i..} - \lambda_{.j.} - \lambda_{...}$$

$$u_{ik}^{XZ} = \frac{1}{c}\sum_{j=1}^{c}\ln m_{ijk} - u_i^X - u_k^Z - u = \lambda_{i.k} - \lambda_{i..} - \lambda_{..k} - \lambda_{...}$$

$$u_{jk}^{YZ} = \frac{1}{r}\sum_{i=1}^{r}\ln m_{ijk} - u_j^Y - u_k^Z - u = \lambda_{.jk} - \lambda_{.j.} - \lambda_{..k} - \lambda_{...}$$

และ

$$u_{ijk}^{XYZ} = \ln m_{ijk} - u_i^X - u_j^Y - u_k^Z - u_{ij}^{XY} - u_{ik}^{XZ} - u_{jk}^{YZ} - u$$

$$= \lambda_{ijk} - \lambda_{ij.} - \lambda_{i.k} + \lambda_{.jk} + \lambda_{i..} + \lambda_{.j.} + \lambda_{..k} + \lambda_{...}$$

ด้วยเหตุนี้จึงมีคุณสมบัติผลรวมของพารามิเตอร์เท่ากับศูนย์ (Zero–sum constraint) ดังนี้

$$\sum_{i=1}^{r}u_i^X = \sum_{j=1}^{c}u_j^Y = \sum_{k=1}^{l}u_k^Z = \sum_{i=1}^{r}u_{ij}^{XY} = \sum_{j=1}^{c}u_{ij}^{XY} = \sum_{i=1}^{r}u_{ik}^{XZ} = \sum_{k=1}^{l}u_{ik}^{XZ} = \sum_{j=1}^{c}u_{jk}^{YZ} = \sum_{k=1}^{l}u_{jk}^{YZ}$$

$$= \sum_{i=1}^{r}u_{ijk}^{XYZ} = \sum_{j=1}^{c}u_{ijk}^{XYZ} = \sum_{k=1}^{l}u_{ijk}^{XYZ} = 0$$

ค่าประมาณความถี่คาดหวัง (Estimated expected frequency) สำหรับตัวแบบเต็มรูป คือ

$$\hat{m}_{ijk} = n\hat{p}_{ijk} = n\left(\frac{x_{ijk}}{n}\right) = x_{ijk}$$

เนื่องจากไม่มีข้อกำหนดใดๆ เกี่ยวกับความเป็นอิสระระหว่างตัวแปรต่างๆ

เนื่องจาก $\sum_{i=1}^{r} u_i^X = 0$ ดังนั้น จึงมีพารามิเตอร์ $\{u_i^X\}$ จำนวน $(r-1)$ ตัวที่เป็นอิสระต่อกันเชิงเส้น (Linearly Independent) และเนื่องจาก $\sum_{i=1}^{r} u_{ij}^{XY} = \sum_{j=1}^{c} u_{ij}^{XY} = 0$ ดังนั้น จึงมีพารามิเตอร์ u_{ij}^{XY} จำนวน $(r-1)(c-1)$ ตัวที่เป็นอิสระต่อกันเชิงเส้น และทำนองเดียวกันนี้สำหรับพารามิเตอร์อื่นๆ จะได้ว่า จำนวนพารามิเตอร์ที่เป็นอิสระต่อกันเชิงเส้น (รวม u) หรือมีจำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบเต็มรูป คือ

$1+(r-1)+(c-1)+(l-1)+(r-1)(c-1)+(r-1)(l-1)+(c-1)(l-1)+(r-1)(c-1)(l-1)$
$= rcl$ ตัวหรือก็คือจำนวนช่องในตารางการณ์จรสามทางนั่นเอง

ตัวแบบที่จะกล่าวถึงต่อไปนี้ ล้วนเป็นตัวแบบที่ลดรูปจากตัวแบบเต็มรูปทั้งสิ้น โดยในการศึกษาครั้งนี้จะพิจารณาเฉพาะตัวแบบลดรูปที่เป็นมีลักษณะเป็น**ตัวแบบเชิงชั้น (Hierarchical Model)** กล่าวคือ เป็นตัวแบบที่รวมเอาอิทธิพลร่วมของตัวแปรหลายตัว และจะต้องมีอิทธิพลร่วมของเซตย่อยของตัวแปรเหล่านั้นอยู่ในตัวแบบด้วย เนื่องจากในทางปฏิบัติเรามักจะพิจารณาตัวแบบเชิงชั้นเป็นหลัก ทั้งนี้เนื่องมาจากเหตุผลในด้านการตีความหมายของตัวแบบ เช่นถ้าในตัวแบบประกอบด้วยอิทธิพลร่วม u_{ij}^{XY} แล้วทำให้การพิจารณาว่าอิทธิพล u_i^X และ u_j^Y จะอยู่ในตัวแบบหรือไม่นั้น ไม่ใช่สิ่งจำเป็นเพราะเมื่อมีอิทธิพลร่วมแล้วก็จะต้องมีอิทธิพลหลักอยู่ในตัวแบบด้วย ตัวอย่างของตัวแบบล็อกลิเนียร์เชิงชั้น (Hierarchical Log-linear Models) เช่น

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ij}^{XY}$$

จะเห็นได้ว่า เมื่อมีอิทธิพลร่วมหลายตัวแปรคือ u_{ij}^{XY} ก็จะมีอิทธิพลหลักคือ u_i^X และ u_j^Y อยู่ในตัวแบบด้วย

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ij}^{XY} + u_{ik}^{XZ}$$

จะเห็นได้ว่า เมื่อมีอิทธิพลร่วมหลายตัวแปรคือ u_{ij}^{XY} และ u_{ik}^{XZ} ก็จะมีอิทธิพลหลักคือ u_i^X, u_j^Y และ u_k^Z อยู่ในตัวแบบด้วย

ตัวอย่างของตัวแบบล็อกลิเนียร์ที่ไม่เป็นตัวแบบเชิงชั้น (Nonhierarchical Models) เช่น

$$\ln m_{ijk} = u + u_i^X + u_k^Z + u_{ij}^{XY}$$

จะเห็นได้ว่า เมื่อมีอิทธิพลร่วมหลายตัวแปร คือ u_{ij}^{XY} แต่ไม่มีอิทธิพลหลัก คือ u_j^Y .อยู่ในตัวแบบ

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_{ij}^{XY} + u_{ik}^{XZ}$$

จะเห็นได้ว่า เมื่อมีอิทธิพลร่วมหลายตัวแปรคือ u_{ij}^{XY} และ u_{ik}^{XZ} แต่ไม่มีอิทธิพลหลักคือ u_k^Z ในตัวแบบ

2. ตัวแบบความเป็นอิสระอย่างสมบูรณ์ (Model of Mutual Independence)

เป็นตัวแบบที่ไม่มีพจน์ของอิทธิพลหรือความสัมพันธ์ร่วมระหว่างตัวแปรใดๆ เราจะใช้สัญลักษณ์ [X][Y][Z] แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z \tag{3.5}$$

โดยตัวแบบความเป็นอิสระอย่างสมบูรณ์นี้มีได้เพียงตัวแบบเดียวเท่านั้น ไม่มีตัวแบบย่อย และตัวแปรแต่ละคู่เป็นอิสระต่อกันแบบมีเงื่อนไขและเป็นอิสระต่อกันแบบชายขอบด้วย ค่าประมาณของความถี่คาดหวังสำหรับตัวแบบความเป็นอิสระอย่างสมบูรณ์ คือ

$$\begin{aligned}\hat{m}_{ijk} &= n\hat{p}_{ijk} = n\hat{p}_{i\cdot\cdot}\hat{p}_{\cdot j\cdot}\hat{p}_{\cdot\cdot k} \\ &= n\left(\frac{n_{i\cdot\cdot}}{n}\right)\left(\frac{n_{\cdot j\cdot}}{n}\right)\left(\frac{n_{\cdot\cdot k}}{n}\right) = \frac{n_{i\cdot\cdot}n_{\cdot j\cdot}n_{\cdot\cdot k}}{n^2}\end{aligned}$$

จำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบความเป็นอิสระอย่างสมบูรณ์คือ $1 + (r-1) + (c-1) + (l-1) = r + c + l - 2$ ตัว

3. ตัวแบบความเป็นอิสระร่วม (Models of joint independence)

โดยตัวแบบความเป็นอิสระร่วมนี้ มีตัวแบบย่อย ได้แก่

1) ตัวแบบที่ปัจจัย Z เป็นอิสระร่วมกับปัจจัย X และ Y

เราจะใช้สัญลักษณ์ [Z][XY] แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ij}^{XY} \tag{3.6}$$

ค่าประมาณของความถี่คาดหวังสำหรับตัวแบบในกรณีนี้ คือ

$$\begin{aligned}\hat{m}_{ijk} &= n\hat{p}_{ijk} = n\hat{p}_{ij\cdot}\hat{p}_{\cdot\cdot k} \\ &= n\left(\frac{n_{ij\cdot}}{n}\right)\left(\frac{n_{\cdot\cdot k}}{n}\right) = \frac{n_{ij\cdot}n_{\cdot\cdot k}}{n}\end{aligned}$$

จำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบ ในกรณีนี้คือ

$1 + (r-1) + (c-1) + (l-1) + (r-1)(c-1) = rc + l - 1$ ตัว

2) ตัวแบบที่ปัจจัย Y เป็นอิสระร่วมกับปัจจัย X และ Z

เราจะใช้สัญลักษณ์ [Y][XZ] แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ik}^{XZ} \tag{3.7}$$

ค่าประมาณของความถี่คาดหวังสำหรับตัวแบบในกรณีนี้ คือ

$$\hat{m}_{ijk} = n\hat{p}_{ijk} = n\hat{p}_{i\cdot k}\hat{p}_{\cdot j\cdot}$$

$$= n\left(\frac{n_{i\cdot k}}{n}\right)\left(\frac{n_{\cdot j\cdot}}{n}\right) = \frac{n_{i\cdot k}n_{\cdot j\cdot}}{n}$$

จำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบ ในกรณีนี้คือ

$1 + (r-1) + (c-1) + (l-1) + (r-1)(l-1) = rl + c - 1$ ตัว

3) ตัวแบบที่ปัจจัย X เป็นอิสระร่วมกับปัจจัย Y และ Z

เราจะใช้สัญลักษณ์ [X][YZ] แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{jk}^{YZ} \tag{3.8}$$

ค่าประมาณของความถี่คาดหวังสำหรับตัวแบบในกรณีนี้ คือ

$$\hat{m}_{ijk} = n\hat{p}_{ijk} = n\hat{p}_{\cdot jk}\hat{p}_{i\cdot\cdot}$$

$$= n\left(\frac{n_{\cdot jk}}{n}\right)\left(\frac{n_{i\cdot\cdot}}{n}\right) = \frac{n_{\cdot jk}n_{i\cdot\cdot}}{n}$$

จำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบ ในกรณีนี้คือ

$1 + (r-1) + (c-1) + (l-1) + (c-1)(l-1) = cl + r - 1$ ตัว

4. ตัวแบบความเป็นอิสระอย่างมีเงื่อนไข (Models of conditional independence)

โดยตัวแบบความเป็นอิสระอย่างมีเงื่อนไขนี้ มีตัวแบบย่อย ได้แก่

1) ตัวแบบที่ตัวแปร X และ Y เป็นอิสระต่อกันอย่างมีเงื่อนไขในตัวแปร Z

เราจะใช้สัญลักษณ์ [XZ][YZ] แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ik}^{XZ} + u_{jk}^{YZ} \tag{3.9}$$

ค่าประมาณของความถี่คาดหวังสำหรับตัวแบบในกรณีนี้ คือ

$$\hat{m}_{ijk} = n\hat{p}_{ijk} = n\left(\frac{\hat{p}_{i\cdot k}\hat{p}_{\cdot jk}}{\hat{p}_{\cdot\cdot k}}\right)$$

$$= n\left[\frac{\left(\frac{n_{i\cdot k}}{n}\right)\left(\frac{n_{\cdot jk}}{n}\right)}{\left(\frac{n_{\cdot\cdot k}}{n}\right)}\right] = \frac{n_{i\cdot k}n_{\cdot jk}}{n_{\cdot\cdot k}}$$

จำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบ ในกรณีนี้คือ

$1+(r-1)+(c-1)+(l-1)+(r-1)(l-1)+(c-1)(l-1) = l(r+c-1)$ ตัว

2) ตัวแบบที่ตัวแปร X และ Z เป็นอิสระต่อกันอย่างมีเงื่อนไขในตัวแปร Y

เราจะใช้สัญลักษณ์ [XY][YZ]แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ij}^{XY} + u_{jk}^{YZ} \tag{3.10}$$

ค่าประมาณของความถี่คาดหวังสำหรับตัวแบบในกรณีนี้ คือ

$$\hat{m}_{ijk} = n\hat{p}_{ijk} = n\left(\frac{\hat{p}_{ij\cdot}\hat{p}_{\cdot jk}}{\hat{p}_{\cdot j\cdot}}\right)$$

$$= n\left[\frac{\left(\frac{n_{ij\cdot}}{n}\right)\left(\frac{n_{\cdot jk}}{n}\right)}{\left(\frac{n_{\cdot j\cdot}}{n}\right)}\right] = \frac{n_{ij\cdot}n_{\cdot jk}}{n_{\cdot j\cdot}}$$

จำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบ ในกรณีนี้คือ

$1+(r-1)+(c-1)+(l-1)+(r-1)(c-1)+(c-1)(l-1) = c(r+l-1)$ ตัว

3) ตัวแบบที่ตัวแปร Y และ Z เป็นอิสระต่อกันอย่างมีเงื่อนไขในตัวแปร X

เราจะใช้สัญลักษณ์ [XY][XZ]แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ij}^{XY} + u_{ik}^{XZ} \tag{3.11}$$

ค่าประมาณของความถี่คาดหวังสำหรับตัวแบบในกรณีนี้ คือ

$$\hat{m}_{ijk} = n\hat{p}_{ijk} = n\left(\frac{\hat{p}_{ij\cdot}\hat{p}_{i\cdot k}}{\hat{p}_{i\cdot\cdot}}\right)$$

$$= \quad n\left[\frac{\left(\frac{n_{ij\cdot}}{n}\right)\left(\frac{n_{i\cdot k}}{n}\right)}{\left(\frac{n_{i\cdot\cdot}}{n}\right)}\right] = \frac{n_{ij\cdot}n_{i\cdot k}}{n_{i\cdot\cdot}}$$

จำนวนพารามิเตอร์ที่ไม่ทราบค่าทั้งหมดของตัวแบบ ในกรณีนี้คือ

$1+(r-1)+(c-1)+(l-1)+(r-1)(c-1)+(r-1)(l-1) = r(c+l-1)$ ตัว

5. ตัวแบบที่ไม่มีตัวแปรคู่ใดเป็นอิสระต่อกัน (Mutual dependence)

เราจะใช้สัญลักษณ์ [XY][XZ][YZ] แทนตัวแบบ ซึ่งแสดงในรูปของพารามิเตอร์ได้ดังนี้

$$\ln m_{ijk} = u + u_i^X + u_j^Y + u_k^Z + u_{ij}^{XY} + u_{ik}^{XZ} + u_{jk}^{YZ} \tag{3.12}$$

เนื่องจากตัวแบบที่ไม่มีตัวแปรคู่ใดเป็นอิสระต่อกันเป็นรูปแบบเปิด (No closed-form) ซึ่งเป็นตัวแบบที่ไม่มีรูปแบบที่แน่นอน การคำนวณค่าคาดหวังจึงเป็นกรณีพิเศษ ซึ่งสามารถศึกษาเพิ่มเติมได้จากหนังสือ Categorical Data Analysis ของ Agresti.(2002)

ตัวพยากรณ์เชิงเส้นของตัวแบบลอกลิเนียร์เชิงชั้นข้างต้น มีโครงสร้างคล้ายตัวแบบสำหรับการวิเคราะห์ความแปรปรวนแบบแฟคทอเรียล 3 ทาง เทอม u_i^X, u_j^Y, u_k^Z ในตัวแบบ $\ln(m_{ijk})$ นั้นเปรียบเสมือนอิทธิพลหลัก ส่วนเทอม $u_{ij}^{XY}, u_{ik}^{XZ}, u_{jk}^{YZ}$ เป็นเทอมของอิทธิพลร่วมของ 2 ตัวแปร และ เทอม u_{ijk}^{XYZ} เป็นเทอมของอิทธิพลร่วมของ 3 ตัวแปร

3.2.1.5 การตีความหมายของตัวแบบล็อกลิเนียร์เชิงซ้อนสำหรับตารางการณ์จรสามทาง

การตีความหมายของตัวแบบข้างต้น ต้องอาศัยเงื่อนไขของการยุบมิติในหัวข้อ 3.2.1.3 เราจึงจะสามารถทราบถึงคุณสมบัติต่างๆ ของตัวแบบล็อกลิเนียร์เชิงชั้นของตารางสามทาง นอกจากนี้ยังสามารถตีความหมายของตัวแบบโดยใช้ตัวแบบโลจิทที่สอดคล้องกันด้วย เช่น Agresti (2002) ได้ตีความหมายของตัวแบบต่อไปนี้

1. ตัวแบบเต็มรูป [XYZ]

เป็นตัวแบบลอกลิเนียร์ทั่วไปสำหรับตารางการณ์จรสามทาง ซึ่งประกอบด้วยอิทธิพลของตัวประกอบ 3 ตัว ตัวแปรแต่ละคู่อาจจะไม่เป็นอิสระแบบมีเงื่อนไขต่อกัน และอัตราส่วนออดส์ของแต่ละคู่อาจแตกต่างกันไปตามตัวแปรทั้ง 3 ตัวแบบ [XYZ] นี้อธิบายเซตของ m_{ijk} ทั้งหมดที่มีค่ามากกว่าศูนย์ โดยการตีความหมายของพารามิเตอร์จะต่างกันออกไปบ้าง คือ

$$u = \ln\left(m_{rcl}\right)$$

$$u_i^X = \ln\left(\frac{m_{icl}}{m_{rcl}}\right)$$

$$u_j^Y = \ln\left(\frac{m_{rjl}}{m_{rcl}}\right)$$

$$u_k^Z = \ln\left(\frac{m_{rck}}{m_{rcl}}\right)$$

$$u_{ij}^{XY} = \ln\left(\frac{m_{ijl}m_{rcl}}{m_{icl}m_{rjl}}\right)$$

$$u_{ik}^{XZ} = \ln\left(\frac{m_{ick}m_{rcl}}{m_{icl}m_{rcl}}\right)$$

$$u_{jk}^{YZ} = \ln\left(\frac{m_{rjl}m_{rcl}}{m_{rjl}m_{rcl}}\right)$$

$$u_{ijk}^{XYZ} = \ln\left[\frac{\left(\frac{m_{ijk}m_{rck}}{m_{ick}m_{rjk}}\right)}{\left(\frac{m_{ijl}m_{rcl}}{m_{icl}m_{rjl}}\right)}\right]$$

โดยที่เทอมของ 2 ตัวประกอบ คือ ล็อกของอัตราส่วนออดส์สำหรับผลคูณของช่องในระดับต้นของทุก ตัวแปร ส่วนเทอมของ 3 ตัวประกอบ คือ ล็อกของอัตราส่วนระหว่างอัตราส่วนออดส์ (Log of Ratios of Odds Ratios) โดยพารามิเตอร์ทั้งหลายสอดคล้องกับเงื่อนไขของ

$$u_i^X = u_j^Y = u_k^Z = u_{rj}^{XY} = u_{rc}^{XY} = \ldots = u_{ijl}^{XYZ} = u_{rcl}^{XYZ} = 0$$

2. ตัวแบบความเป็นอิสระอย่างสมบูรณ์ [X][Y][Z]

เป็นตัวแบบที่ตัวแปรทั้ง 3 เป็นอิสระต่อกัน มีผลทำให้ตัวแปร Y เป็นอิสระร่วมกับตัวแปร Z ตลอดจนตัวแปร X และตัวแปร Y จะเป็นอิสระต่อกันแบบมีเงื่อนไข นั่นคือ ตัวแปรแต่ละคู่เป็นอิสระต่อกันแบบมีเงื่อนไขและเป็นอิสระต่อกันแบบชายขอบด้วย เพราะฉะนั้น ความเป็นอิสระต่อ

กันแบบชายขอบ จึงหมายความว่ามีความเป็นอิสระต่อกันแบบมีเงื่อนไข ตัวอย่างเช่น ความสัมพันธ์แบบชายขอบ X–Y เป็นอันเดียวกัน (Identical) กับความสัมพันธ์บางส่วน (เมื่อกำหนด Z) ซึ่งเป็นเพราะว่า Z เป็นอิสระต่อกันแบบมีเงื่อนไขกับ X (เมื่อกำหนด Y) หรือเพราะว่า Z เป็นอิสระต่อกันแบบมีเงื่อนไขกับ Y (เมื่อกำหนด X) ด้วยก็ได้

สมมติว่า Y เป็นตัวแปรตามแบบ Binary ดังนั้นตัวแบบ (3.5) จะมีโลจิทบน Y เท่ากับ

$$\ln\left(\frac{m_{i1k}}{m_{i2k}}\right) = u_1^Y - u_2^Y$$

สำหรับเงื่อนไขของผลรวมเป็นศูนย์ จะได้ว่า

$$\ln\left(\frac{m_{i1k}}{m_{i2k}}\right) = 2u_1^Y$$

และ $\exp\left(2u_1^Y\right)$ คือ ออดส์ $\left(\frac{m_{i1k}}{m_{i2k}}\right)$ สำหรับแต่ละหมวดหมู่ของ X และ Z

ส่วนในกรณีที่ เมื่อ Y มี $c > 2$ ระดับจะได้ว่า

$$\ln\left(\frac{m_{iak}}{m_{ibk}}\right) = u_a^Y - u_b^Y$$

หมายความว่ายิ่งค่าของ u_j^Y มีค่าสูงขึ้น จะมีหน่วยที่ถูกจัดอยู่ในระดับที่ j ของ Y มากขึ้น

3. ตัวแบบความเป็นอิสระร่วม [Z][XY], [Y][XZ], [X][YZ]

เป็นตัวแบบที่มีตัวแปรเพียงคู่เดียวที่ไม่เป็นอิสระแบบมีเงื่อนไข เช่นจากตัวแบบ (3.7) ซึ่งมีสัญลักษณ์คือ [Y][XZ] หมายความว่า Y เป็นอิสระร่วมกับ X และ Z ดังนั้นจะได้ว่า Y และ X เป็นอิสระต่อกันแบบมีเงื่อนไข (เมื่อกำหนด Z) ในทำนองเดียวกัน Y และ Z เป็นอิสระต่อกันแบบมีเงื่อนไข (เมื่อกำหนด X) ดังนั้น สัญลักษณ์ของตัวแบบ (3.7) จึงแสดงความสัมพันธ์แบบมีเงื่อนไขของ X และ Z โดยพิจารณาจากเทอม u_{ik}^{XZ}

สำหรับตัวแบบ (3.7) จะหมายรวมถึงว่า Y เป็นอิสระต่อกันกับ X และ Z ในตารางแบบชายขอบ X–Y และตารางแบบชายขอบ Y–Z ด้วย นอกจากนี้ด้วยเงื่อนไขของการยุบได้ ยังแสดงถึง อัตราส่วนออดส์แบบชายขอบ X–Z (The X–Z Marginal Odds Ratios) เท่ากับอัตราส่วนออดส์แบบชายขอบ X–Z (The X–Z Partial Odds Ratios) เพราะว่า Y เป็นอิสระกันกับ X (เมื่อกำหนด Z) หรือ Y เป็นอิสระกันกับ Z (เมื่อกำหนด X)

โลจิทของ Y สำหรับตัวแบบ [Y][XZ] จะเท่ากันกับโลจิทของ Y สำหรับตัวแบบ [X][Y][Z] และความเป็นอิสระร่วมของ Y กับ (X และ Z) จะหมายถึงโลจิทของ Y เท่ากันในทุกหมวดหมู่ของ X และ Z ด้วย

4. ตัวแบบความเป็นอิสระอย่างมีเงื่อนไข [XY][YZ], [XY][XZ], [XZ][YZ]

เป็นตัวแบบที่มีตัวแปรเพียงคู่เดียว ที่เป็นอิสระแบบมีเงื่อนไข เช่น จากตัวแบบ (3.9) ซึ่งมีสัญลักษณ์คือ [XZ][YZ] หมายความว่าตัวแปร X และ Y เป็นอิสระต่อกันแบบมีเงื่อนไข (เมื่อกำหนด Z) โดยมีพารามิเตอร์ u_{ij}^{XY} และ u_{jk}^{YZ} แทนความสัมพันธ์บางส่วนของ X – Z และ Y – Z ตามลำดับ

ในหัวข้อ 3.2 มีข้อสังเกตว่า X และ Y อาจจะขึ้นอยู่ต่อกันแบบชายขอบ ทั้งๆ ที่เป็นอิสระต่อกันแบบมีเงื่อนไข ซึ่งเป็นผลมาจากเงื่อนไขการยุบตัวแปร เพราะว่า Z ไม่เป็นอิสระแบบมีเงื่อนไขกับทั้ง X และ Y อย่างไรก็ตาม ตารางแบบชายขอบ X – Z และตารางแบบชายขอบ Y – Z ต่างมีอัตราส่วนออดส์เท่ากันกับตารางบางส่วน เพราะว่า X และ Y เป็นอิสระต่อกันแบบมีเงื่อนไข

สำหรับตัวแบบ (3.9) ถ้า Y เป็น Binary และมีโลจิทเท่ากับ

$$\ln\left(\frac{m_{i1k}}{m_{i2k}}\right) = \left(u_1^{Y} - u_2^{Y}\right) + \left(u_{1k}^{YZ} + u_{2k}^{YZ}\right)$$

ภายใต้เงื่อนไขของผลรวมเป็นศูนย์จะได้โลจิทเท่ากับ $u_{1k}^{YZ} + u_{2k}^{YZ}$ และรูปแบบของโลจิทคือ $\alpha + \beta_k^{Z}$ โดยที่ $2u_{1k}^{YZ}$ แทนอิทธิพลที่ k ของ Z ของโลจิทของ Y (นั่นคือ $2u_{1k}^{YZ} = \beta_k^{Z}$) และ $2u_1^{Y} = \alpha$ ดังนั้น โลจิทข้างต้นจะขึ้นอยู่กับระดับที่ k ของ Z (ด้วยเทอมความสัมพันธ์ของ Y และ Z) แต่จะไม่ขึ้นอยู่กับระดับที่ i ของ X

อนึ่ง ตัวแบบ [XZ][YZ] เป็นตัวแบบที่สำคัญตรงที่ ถ้าตัวแปร Z เป็นตัวแปรที่สามารถควบคุมได้ ความสัมพันธ์ระหว่าง X และ Y จะไม่จริงนั่นเอง

5. ตัวแบบที่ไม่มีตัวแปรคู่ใดเป็นอิสระต่อกัน [XY][XZ][YZ]

เป็นตัวแบบที่เทอมความสัมพันธ์บางส่วนปรากฏอยู่ในแต่ละคู่ของตัวแปร ดังนั้น จะไม่มีตัวแปรคู่ใดที่เป็นอิสระต่อกันแบบมีเงื่อนไข และไม่มีเทอมอิทธิพลร่วมของ 3 ตัวประกอบ หรือ 3 ตัวแปร ไม่มีเงื่อนไขการยุบตัวแปร ดังนั้น อัตราส่วนออดส์แบบชายขอบ อาจมีค่าแตกต่างจากค่าของอัตราส่วนออดส์แบบบางส่วนได้ในแต่ละคู่ของตัวแปร

จากตัวแบบ (3.12) แทนตัวแบบ ด้วยนิพจน์สำหรับ $\ln\left(\theta_{ij(k)}\right)$ ดังนั้น

$$\ln\left(\theta_{ij(k)}\right) = u_{ij}^{XY} + u_{i+1,j+1}^{XY} - u_{i,j+1}^{XY} - u_{i+1,j}^{XY}$$

(3.13)

เนื่องจากเทอมทางขวาของสมการ (3.13) ไม่มีอิทธิพลร่วม 3 ตัวประกอบ และมีค่าเท่าเดิมทุกๆ k ดังนั้น

$$\theta_{ij(1)} = \theta_{ij(2)} = ... = \theta_{ij(K)} \qquad \text{ทุกๆ } i \text{ และ } j$$

ทำนองเดียวกัน สำหรับอัตราส่วนออดส์แบบบางส่วนอื่นๆ จะได้ว่า

$$\theta_{i(1)k} = \theta_{i(2)k} = ... = \theta_{i(j)k} \qquad \text{ทุกๆ } i \text{ และ } k$$

และ

$$\theta_{(1)jk} = \theta_{(2)jk} = ... = \theta_{(I)jk} \qquad \text{ทุกๆ } j \text{ และ } k$$

ซึ่งหมายความว่า เมื่อตัวแบบไม่มีอิทธิพลร่วมของ 3 ตัวประกอบ ความสัมพันธ์ระหว่าง 2 ตัวแปร จะเท่ากันในแต่ละระดับของตัวแปรตัวที่ 3

สำหรับ ตัวแบบ [XY][XZ][YZ] เมื่อ Y เป็น Binary และภายใต้เงื่อนไขของผลรวมเท่ากับศูนย์ จะได้โลจิทเท่ากับ

$$\ln\left(\frac{m_{i1k}}{m_{21k}}\right) = 2u_1^{Y} + 2u_{i1}^{XY} + 2u_{1k}^{YZ}$$

โดยอยู่ในรูปแบบ $\alpha + \beta_i^X + \beta_k^Z$ ซึ่งขึ้นอยู่กับระดับของ X และ Z แต่อยู่ในลักษณะที่เป็นผลมากกว่าศูนย์ (Additive Manner) และหมายความว่าอิทธิพลของ X ในโลจิทจะเท่ากันในแต่ละระดับของ Z และอิทธิพลของ Z ในโลจิทจะเท่ากันในแต่ละระดับของ X

3.2.2 ตัวแบบล็อกลิเนียร์สำหรับตารางการณ์จรที่มากกว่าสามทาง

ถ้าเรามีความเข้าใจเกี่ยวกับตัวแบบล็อกลิเนียร์สำหรับตารางการณ์จรสองทาง (2 มิติ) และสามทาง (3 มิติ) แล้ว ตัวแบบล็อกลิเนียร์สำหรับตารางการณ์จรที่มากกว่าสามทาง จะมีหลักเกณฑ์ทำนองเดียวกัน โดยขยายจากตารางการณ์จรสามทาง (Three – Way Contingency Table) ให้เป็นตารางการณ์จรหลายทาง (Multi – Way Contingency Tables)

ตัวแบบล็อกลิเนียร์แบบมากกว่า 3 มิตินี้ อาจมีปัญหาเกี่ยวกับอิทธิพลร่วมของตัวประกอบ (หรือตัวแปรของมิติต่างๆ) ซึ่งเพิ่มจำนวนจากเดิม 3 ตัวประกอบเป็นหลายตัวประกอบ นอกจากนี้ การที่จำนวนช่องซึ่งเพิ่มตามจำนวนตัวประกอบและระดับของตัวประกอบเหล่านั้น อาจมีผลทำให้

บางช่องหรือหลายช่องมีจำนวนนับเป็นศูนย์ได้ อย่างไรก็ตามปัญหาหลังจะลดน้อยลงได้เมื่อใช้ขนาดตัวอย่างให้มากพอ แต่ถ้ายังคงมีปัญหาเดิมอยู่อีก ให้ตรวจสอบค่าศูนย์ดังกล่าวว่ามีสาเหตุเนื่องมาจากตัวอย่าง (Sampling Zeroes) หรือเนื่องมาจากโครงสร้าง (Structural Zeroes)

3.2.2.1 ตัวแบบล็อกลิเนียร์สำหรับตารางการณ์จรสี่ทาง (4 มิติ)

กำหนดให้มีตัวแปร 4 ตัว คือ H X Y และ Z โดยที่ตัวแบบของความเป็นอิสระอย่างสมบูรณ์ แทนด้วย [HXYZ]

ถ้าตัวแบบไม่มีอิทธิพลร่วมของตัวแปร 3 ตัว ซึ่งจะทำให้ไม่มีอิทธิพลร่วมของตัวแปร 4 ตัวด้วย จะมีตัวแบบเป็น

$$\ln(m_{hijk}) = u + u_h^H + u_i^X + u_j^Y + u_k^Z + u_{hi}^{HK} + u_{hj}^{HY} + u_{hk}^{HZ} + u_{ij}^{XY} + u_{ik}^{XZ} + u_{jk}^{YZ}$$

ซึ่งแทนด้วย [HX][HY][HZ][XY][XZ][YZ] โดยตัวแปรแต่ละคู่จะเป็นอิสระกันแบบมีเงื่อนไข (เมื่อกำหนดตัวแปร 2 ตัวอื่น) เช่น ถ้า $u_{jk}^{YZ} = 0$ สำหรับทุกๆ j และ k แล้ว Y และ Z จะเป็นอิสระต่อกันแบบมีเงื่อนไข ณ แต่ละหมวดหมู่ของระดับของ H และ X เป็นต้น ดังนั้นการเป็นอิสระต่อกันแบบมีเงื่อนไขของตัวแปรคู่ใด หมายความว่า จะไม่มีเทอมอิทธิพลร่วมของตัวแปรคู่นั้นๆ เพราะมีค่าเป็นศูนย์ไป

ตัวอย่างของตัวแบบที่มีอิทธิพลร่วมของ 3 ตัวประกอบ เช่น [HXY][HZ][XZ][YZ] แต่ละคู่ของตัวแปรไม่เป็นอิสระต่อกันแบบมีเงื่อนไข ณ แต่ละระดับของ Z ความสัมพันธ์ระหว่าง H กับ X หรือ H กับ Y หรือ X กับ Y จะเปลี่ยนไปตามระดับของตัวแปรที่เหลือ (พิจารณา HXY) และความสัมพันธ์บางส่วนระหว่าง Z และตัวแปรอื่นจะคงที่หรือเท่ากันในแต่ละหมวดหมู่ของระดับต่างๆ ของตัวแปรอื่นๆ อีก 2 ตัว (พิจารณา HZ, XZ, YZ)

3.2.2.2 การยุบมิติ

การวิเคราะห์ตารางการณ์จรหลายทางจะเป็นไปโดยง่ายขึ้น เมื่อสามารถยุบมิติหนึ่งหรือหลายมิติของตัวแปร ดังที่กล่าวในกรณี 3 มิตินั้น การยุบมิติต้องอยู่ภายใต้เงื่อนไขความพอเพียงต่างๆ ด้วย ซึ่งกรณีของตารางหลายมิติก็เช่นกัน โดยอาศัยเงื่อนไขตามที่ Bishop et, al. (1975) (อ้างถึงใน Agresti: 2002) แสดงไว้ กล่าวคือ

สมมติให้ตัวแปรในตาราง S มิติถูกแบ่งออกเป็น 3 กลุ่ม แบบไม่มีส่วนร่วมระหว่างกลุ่ม (Three Mutually Exclusive Groups) ถ้าทุกๆ เทอมของตัวแบบที่เชื่อม (Link) ตัวแปรจากกลุ่มที่หนึ่งกับตัวแปรจากกลุ่มที่สองมีค่าเท่ากับศูนย์ ดังนั้นเทอมของตัวแบบระหว่างตัวแปรสำหรับกลุ่มที่หนึ่งจะไม่เปลี่ยนแปลง เมื่อมีการยุบตารางภายใต้ตัวแปรกลุ่มที่สอง

ตัวอย่างเช่น ตารางที่ 4 มิติ ให้ตัวแปร A และ B ประกอบกันเป็นกลุ่มของตัวแปรกลุ่มที่ 1 และ C เป็นกลุ่มที่ 2 ส่วน D เป็นกลุ่มที่ 3 ถ้าเทอม u_{hj}^{AC} และ u_{ij}^{BC} มีค่าเป็นศูนย์ ความสัมพันธ์บางส่วนของ AB (AB Partial Association) จะยังคงเท่าเดิมเมื่อมีการยุบตารางภายใต้ตัวแปร C ซึ่งเกิดขึ้นได้สำหรับตัวแบบ [ABD][CD] ในกรณีอื่น เช่นสมมติให้ A และ B อยู่ในกลุ่มที่ 1 C และ D อยู่ในกลุ่มที่ 2 ส่วนกลุ่มที่ 3 เป็นกลุ่มว่างเปล่า สำหรับตัวแบบ [AB][CD] ความสัมพันธ์บางส่วนของ AB ยังคงเท่ากับความสัมพันธ์แบบชายขอบของ AB

ผลดังกล่าวข้างต้น สามารถใช้ได้กับกรณีของตัวแปรหนึ่งๆ ที่เป็นอิสระต่อกันกับตัวแปรอื่นๆ ที่เหลือก็สามารถยุบตารางภายใต้ตัวแปรตัวนั้น โดยไม่กระทบกระเทือนต่อเทอมของตัวแบบเทอมอื่นๆ ด้วย เช่น ความสัมพันธ์ของ A B และ C ในตัวแบบ [AB][AC][BC][D] จะยังคงเท่าเดิมกับความสัมพันธ์ที่อยู่ในตัวแบบ [AB][AC][BC จึงสามารถยุบมิติของ D ได้

เป็นที่น่าสังเกตว่า ถ้าปรากฏเทอมอิทธิพลระหว่างตัวประกอบ 2 ตัวทุกคู่ขึ้นในตัวแบบหนึ่งแล้ว การยุบมิติภายใต้ตัวแปรใดตัวแปรหนึ่งอาจจะมีผลทำให้อิทธิพลดังกล่าวเปลี่ยนแปลงได้ เช่น ตัวแบบ [AB][AC][AD][BC][BD][CD] แต่ละคู่ของตัวแปรอาจจะมีความสัมพันธ์บางส่วน และความสัมพันธ์แบบชายขอบที่แตกต่างกันนั่นเอง จึงไม่ควรมีการยุบมิติในกรณีเช่นนี้

3.2.2.3 ความสอดคล้องของตัวแบบล็อกลิเนียร์กับตัวแบบโลจิท

ในการประยุกต์กับงานต่างๆ นั้น ตัวแบบล็อกลิเนียร์ที่นำไปใช้ยังไม่กว้างขวาง มีเพียงเซตย่อยของตัวแบบล็อกลิเนียร์เท่านั้นที่ผู้วิจัยใช้งานกัน เช่นการที่งานวิจัยส่วนใหญ่เน้นที่ความแตกต่างระหว่างตัวแปรอิสระและตัวแปรตาม การสร้างตัวแบบที่แสดงอิทธิพลระหว่างตัวแปรอิสระและตัวแปรตามจึงมีความสำคัญ และมีการใช้มากกว่าการสร้างตัวแบบที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระต่างๆ เซตย่อยของตัวแบบล็อกลิเนียร์ที่สอดคล้องกับวัตถุประสงค์ดังกล่าวนี้ คือ ตัวแบบที่สมมูลกับตัวแบบโลจิทสำหรับตัวแปรตาม เพราะโครงสร้างของตัวแบบล็อกลิเนียร์นั้น ได้รวมโครงสร้างของตัวแบบโลจิทสำหรับความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระไว้ด้วย เช่น ในกรณีของตารางข้อมูลขนาด $r \times c \times 2$ จะสามารถสร้างตัวแบบล็อกลิเนียร์ที่สอดคล้องกับตัวแบบโลจิทได้ดังนี้

$$\ln\left(\frac{m_{ij1}}{m_{ij2}}\right) = \alpha + \beta_i^A + \beta_j^B \tag{3.14}$$

ซึ่งเป็นตัวแบบที่มีตัวแปรตาม Y มีความสัมพันธ์กับตัวประกอบ A และ B โดยอิทธิพลของแต่ละตัวประกอบจะเท่ากันในทุกระดับของอีกตัวประกอบหนึ่ง ส่วนตัวแบบล็อกลิเนียร์ของ

ตารางขนาด $r\times c\times 2$ จะรวมเทอมความสัมพันธ์ต่างๆ คือ u_{ik}^{AY} และ u_{jk}^{BY} หรือเขียนใหม่เป็นตัวแบบ [AB][AY][BY] ซึ่งเป็นตัวแบบล็อกลิเนียร์ที่มีความสอดคล้องกับตัวแบบ (3.14) โดยที่ Y มี 2 ระดับ ส่วน A และ B มี r และ c ระดับ ตามลำดับ

การพิสูจน์ ตัวแบบ (3.14)

$$\ln\left(\frac{m_{ij1}}{m_{ij2}}\right) = \ln(m_{ij1}) - \ln(m_{ij2})$$

$$= [u + u_i^A + u_j^B + u_1^Y + u_{ij}^{AB} + u_{i1}^{AY} + u_{j1}^{BY}]$$

$$- [u + u_i^A + u_j^B + u_2^Y + u_{ij}^{AB} + u_{i2}^{AY} + u_{j2}^{BY}]$$

$$= (u_1^Y - u_2^Y) + (u_{i1}^{AY} - u_{i2}^{AY}) + (u_{j1}^{BY} - u_{j2}^{BY})$$

เนื่องจากเงื่อนไข $\sum_k u_k^Y = \sum_k u_{ik}^{XY} = \sum_k u_{ik}^{BY} = 0$ และ Y มี 2 ระดับ

$$\therefore \qquad u_1^Y = -u_2^Y, \quad u_{i1}^{AY} = -u_{i2}^{AY} \text{ และ } u_{j1}^{BY} = -u_{j2}^{BY}$$

ดังนั้น รูปแบบของโลจิทจึงสามารถเขียนให้อยู่ในรูปของ

$$\ln\left(\frac{m_{ij1}}{m_{ij2}}\right) = 2u_i^Y + 2u_{i1}^{AY} + 2u_{j1}^{BY}$$

ซึ่งมีรูปแบบเดียวกันกับตัวแบบ (3.14) คือ

$$\ln\left(\frac{m_{ij1}}{m_{ij2}}\right) = \alpha + \beta_i^A + \beta_j^B$$

โดยที่ เทอม $2u_{i1}^{AY}$ แทนอิทธิพลที่ i ของ A ในตัวแบบโลจิทของ Y (นั่นคือ β_i^A)
เทอม $2u_{j1}^{BY}$ แทนอิทธิพลที่ j ของ B ในตัวแบบโลจิทของ Y (นั่นคือ β_j^A)
เทอม $2u_1^Y$ แทน α

สำหรับเทอม u_{ij}^{AB} ซึ่งแทนความสัมพันธ์ระหว่างตัวแปรอิสระนั้น ได้ตัดกันหมดในเทอมผลต่างของลอการิทึมที่กำหนดโดยโลจิท ตัวแบบโลจิทจึงไม่รวมสารสนเทศเกี่ยวกับความสัมพันธ์ระหว่างสองตัวประกอบดังกล่าว

ในกรณีของตารางที่ 4 มิติ ที่ Y เป็นตัวแปรตามแบบ Binary และมีตัวประกอบ 3 ตัว คือ A B และ C ตัวแบบโลจิทของกรณี 4 มิตินี้จะมีรูปแบบทำนองเดียวกันกับ (3.14) คือ

$$\ln\left(\frac{m_{ij1}}{m_{ij2}}\right) = \alpha + \beta_h^A + \beta_j^B + \beta_j^C \tag{3.15}$$

เป็นตัวแบบที่รวมเทอมอิทธิพลหลักของตัวแปรต่างๆ แต่ไม่รวมเทอมอิทธิพลร่วม ตัวแบบ (3.15) จึงสอดคล้องกับตัวแบบล็อกลิเนียร์ที่ประกอบด้วยเทอมอิทธิพลร่วมของตัวแปรอิสระทั้งหมดและเทอมความสัมพันธ์ระหว่างแต่ละตัวแปรอิสระกับตัวแปรตาม นั่นคือ ตัวแบบ [ABC][AY][BY][CY] ส่วนตัวแบบทั่วไปของตัวแบบ (3.15) ที่รวมเทอม β_{hi}^{AB} หรือเทอมความสัมพันธ์ระหว่าง 2 ตัวแปรอิสระ และตัวแปรตามจะสอดคล้องกับตัวแบบล็อกลิเนียร์ [ABC][ABY][CY] ที่มีอิทธิพลร่วมระหว่าง A และ B บนตัวแปร Y นอกจากนี้ กรณีพิเศษของตัวแบบ (3.15) ซึ่งมี $\beta_j^C = 0$ ทุกๆ j ก็สอดคล้องกับตัวแบบล็อกลิเนียร์ [ABC][AY][BY] ที่มีความเป็นอิสระต่อกันแบบมีเงื่อนไขระหว่าง Y และ C เมื่อกำหนด A และ B

สรุป ถ้าพิจารณาตัวแปรตาม ตัวแบบล็อกลิเนียร์จะสอดคล้องกับตัวแบบโลจิทสำหรับตัวแปรตามนั้นๆ ถ้าตัวแปรตามมีมากกว่า 2 กลุ่ม ตัวแบบล็อกลิเนียร์ที่เกี่ยวข้องจะสอดคล้องกับ ตัวแบบโลจิทที่วางนัยทั่วไป (Generalized Logit Models) ส่วนตัวแบบล็อกลิเนียร์โดยทั่วไปไม่จำเป็นต้องคำนึงถึงตัวแปรใดตัวแปรหนึ่งเป็นตัวแปรตาม โดยที่ลักษณะที่ศึกษา คือ ตัวแปรหรือตัวประกอบที่กำลังวิเคราะห์อยู่นั้น มีลักษณะเชิงกลุ่มที่เป็นความสำคัญสำหรับกรณีตัวแบบโลจิทที่วางนัยทั่วไป เมื่อมีการกำหนดตัวแปรเชิงกลุ่มตัวใดเป็นตัวแปรตาม

3.3 ตัวแบบล็อกลิเนียร์มาตรเรียงลำดับ (Ordinal Log-linear Model)

ตัวแบบล็อกลิเนียร์ที่กล่าวมาข้างต้นนั้น เป็นตัวแบบที่ได้จากการวิเคราะห์ตัวแปรเชิงกลุ่ม โดยที่ไม่ได้ให้ความสำคัญกับอันดับ (Order) ของตัวแปร แต่ถ้าผู้วิจัยให้ความสนใจกับอันดับของตัวแปรเชิงกลุ่มนั้น การวิเคราะห์ข้อมูลที่ตัวแปรถูกจัดเป็นกลุ่มแบบมีลำดับในตารางกาณ์จรก็จะแตกต่างไปจากการวิเคราะห์ที่ผ่านมา ซึ่งไม่ได้ให้ความสำคัญกับอันดับของตัวแปรเชิงกลุ่ม

การวิเคราะห์ล็อกลิเนียร์มาตรเรียงลำดับ สามารถวิเคราะห์ได้ในกรณีที่ตัวแปรตั้งแต่หนึ่งตัวขึ้นไปเป็นมาตรเรียงลำดับ โดยข้อมูลของตัวแปรต้องมีการจัดเรียงกันตามธรรมชาติ เช่น อายุ

ความดันเลือด หรือระดับอาการของโรค เป็นต้น ดังนั้นจึงอาจนำตัวแปรมาตรเรียงอันดับเข้ามาศึกษากับตัวแปรมาตรนามบัญญัติหรือ เป็นการศึกษาในตัวแปรที่เป็นมาตรเรียงอันดับทั้งหมดก็ได้ ซึ่งตัวแบบล็อกลิเนียร์มาตรเรียงลำดับจะแตกต่างจากตัวแบบล็อกลิเนียร์สำหรับตัวแปรเชิงกลุ่มมาตรนามบัญญัติ ตรงที่มีการศึกษาถึงตัวแบบเกี่ยวเนื่องเพิ่มขึ้น

สำหรับกรณีตารางการณ์จรสองทาง แทนที่จะวิเคราะห์ข้อมูลด้วยตัวสถิติไคสแควร์ เราสามารถใช้การวิเคราะห์การถดถอย เข้ามาช่วยในการวิเคราะห์ข้อมูลในตารางการณ์จรที่มีตัวแปรเชิงกลุ่มแบบเรียงอันดับโดยการให้คะแนน (Score) แก่กลุ่มของตัวแปรเชิงกลุ่ม โดยคำนึงถึงอันดับก่อนหลัง ส่วนกรณีตารางการณ์จรหลายทาง อาจใช้การวิเคราะห์โดยอาศัย ตัวแบบความสัมพันธ์แบบ Linear-by-Linear ตัวแบบ Row Effect และตัวแบบ Column Effect เป็นต้น

3.3.1 ตัวแบบความสัมพันธ์แบบ Linear-by-Linear

คือตัวแบบที่สร้างขึ้นเพื่อใช้แสดงความสัมพันธ์ระหว่างตัวแปรเชิงกลุ่มแบบเรียงลำดับสองตัว สำหรับตารางการณ์จรสองทาง ซึ่งตัวแปรที่ศึกษาทั้งสองตัวเป็นตัวแปรมาตรเรียงลำดับ โดยจะกำหนดให้ v_i แทนคะแนนของระดับการแปรค่าที่ให้แก่กลุ่มในแถวที่ i เมื่อ $i = 1, 2, ..., r$ และ w_j แทนคะแนนของระดับการแปรค่าที่ให้แก่กลุ่มในสดมภ์ที่ j เมื่อ $j = 1, 2, ..., c$ ในการกำหนดคะแนนในแถวแต่ละแถวและสดมภ์แต่ละสดมภ์นั้น จะกำหนดให้คะแนนของแถวระดับ (กลุ่ม) ที่ 1 น้อยกว่าคะแนนในแถวระดับที่ 2 น้อยกว่าคะแนนในแถวที่ 3 ไปจนกระทั่งน้อยกว่าคะแนนในแถวที่ r นั่นคือ $v_1 \le v_2 \le ... \le v_r$ และ คะแนนในสดมภ์ที่ 1 น้อยกว่าคะแนนในสดมภ์ที่ 2 น้อยกว่าคะแนนในสดมภ์ที่ 3 ไปจนกระทั่งน้อยกว่าคะแนนในสดมภ์ที่ c นั่นคือ $w_1 \le w_2 \le ... \le w_c$ ตัวแบบที่ได้จะมีรูปแบบคล้ายกับตัวแบบล็อกลิเนียร์ทั่วๆ ไป แต่จะมีเทอมของ β, v_i และ w_j เพิ่มขึ้นมาโดยมีตัวแบบดังนี้

$$\ln(m_{ij}) = u + u_i^X + u_j^Y + \beta v_i w_j \tag{3.15}$$

ถ้า $\beta = 0$ ตัวแบบ (3.15) จะเป็นตัวแบบที่เป็นอิสระต่อกัน และเนื่องจากค่า v_i และ w_j แทนอิทธิพลที่คงที่ของตัวแบบนี้ จึงมีพารามิเตอร์ β ซึ่งเป็นเทอมเกี่ยวเนื่อง (Association term) เพิ่มจากตัวแบบที่เป็นอิสระต่อกัน ดังนั้นภายใต้เงื่อนไข $\sum_{i=1}^{r} u_i^X = \sum_{j=1}^{c} u_j^Y = 0$ จะได้จำนวนองศาความเป็นอิสระลดลงไป 1 ตัวจากตัวแบบที่เป็นอิสระต่อกัน จึงทำให้จำนวนองศาความเป็นอิสระเท่ากับ $(r-1)(c-1)-1 = rc - r - c$ และจะพบว่าในตัวแบบ (3.15) จะมีพารามิเตอร์ β เพียงตัวเดียวที่ใช้อธิบายความสัมพันธ์ของตัวแปรเชิงกลุ่มแบบเรียงลำดับทั้งทางแถวและทางสดมภ์ โดยไม่คำนึงถึงจำนวนแถวและจำนวนสดมภ์ของตารางการณ์จร สำหรับเทอม $\beta v_i w_j$ ซึ่งเป็นส่วนที่

แตกต่างกันระหว่างตัวแปรตัวแบบที่เป็นอิสระต่อกันกับตัวแบบ (3.15) นั้น จะอยู่ในรูปแบบเชิงเส้นในแต่ละตัวแปร กล่าวคือ เป็น Linear in X for Fixed Y และเป็น Linear in Y for Fixed X จึงเรียกตัวแบบ (3.15) ว่า ตัวแบบความสัมพันธ์แบบ Linear-by-Linear

การตีความหมายเกี่ยวกับขนาดของ β สามารถทำได้ในเทอมของอัตราส่วนออดส์ โดยพิจารณาในกรณีที่แถวที่ $a < b$ กับสดมภ์ที่ $c<d$ ในแต่ละคู่ของตารางการณ์จร ด้วยตัวแบบความสัมพันธ์แบบ Linear-by-Linear พบว่า

$$\ln\left(\frac{\hat{m}_{ac}\hat{m}_{bd}}{\hat{m}_{ad}\hat{m}_{bc}}\right) = \beta(v_b - v_a)(w_d - w_c)$$

สำหรับการตีความหมายของ β สามารถสรุปได้ดังนี้

1. ถ้า β มีค่าเข้าใกล้ 0 ตัวแปรกลุ่มทั้งสองเป็นอิสระต่อกันเชิงเส้น (แต่ตัวแปรทั้งสองตัวนั้นอาจมีความสัมพันธ์กันแบบ Nonlinear)
2. ถ้า $\beta > 0$ คาดได้ว่า ควรมีค่าสังเกตที่เพิ่มขึ้นในช่องที่ตัวแปรทั้งสองมีคะแนนเพิ่มมากขึ้นหรือในทางตรงกันข้ามค่าสังเกตที่ลดลงในช่องที่ตัวแปรทั้งสองมีคะแนนน้อยลง
3. ถ้า $\beta < 0$ คาดได้ว่า ควรมีค่าสังเกตที่เพิ่มขึ้นในช่องที่ตัวแปรหนึ่งมีคะแนนเพิ่มมากขึ้นในขณะที่อีกตัวแปรหนึ่งมีคะแนนลดลง

Birch, Goodman และ Habereman (อ้างใน Agresti, 1984) ได้กำหนดตัวแบบล็อกลิเนียร์สำหรับตัวแปรมาตรเรียงลำดับในเทอมของ $(r-1)(c-1)$ ในการคำนวณหาค่าอัตราส่วนออดส์ว่า

$$\theta_{ij} = \left(\frac{\hat{m}_{ij}\hat{m}_{i+1,j+1}}{\hat{m}_{i,j+1}\hat{m}_{i+1,j}}\right)$$

จะได้ว่า $$\ln(\theta_{ij}) = \beta(v_{i+1} - v_i)(w_{j+1} - w_j)$$

ในกรณีที่ใช้สำหรับแถวและสดมภ์ที่ติดกัน เมื่อ $1 \le i \le r-1$ และ $1 \le j \le c-1$ Goodman กล่าวว่าตัวแบบเกี่ยวเนื่องนี้จะได้ค่า θ_{ij} ทุกตัวที่มีค่าเท่ากันนั่นคือช่วงห่างของคะแนนที่เท่ากัน ($v_2 - v_1 = v_3 - v_2 = ... = v_r - v_{r-1}$ และ $w_2 - w_1 = w_3 - w_2 = ... = w_c - w_{c-1}$) มีผลทำให้ θ_{ij} ทุกตัวมีค่าเท่ากัน

3.3.2 ตัวแบบ Row Effect และ ตัวแบบ Column Effect

Goodman, Haberman และ Simon (อ้างใน Agresti, 1984) เป็นผู้พัฒนาตัวแบบ Row Effect และ ตัวแบบ Column Effect โดยการใช้ตัวแบบ Row Effect ในการวิเคราะห์ความสัมพันธ์นั้น เป็นการพิจารณาการจัดลำดับของตัวแปรทางสดมภ์เท่านั้น โดยไม่พิจารณาการจัดลำดับของตัวแปรทางแถว เริ่มต้นด้วยการให้คะแนน $w_1 \le w_2 \le \ldots \le w_c$ แก่กลุ่มของตัวแปรเชิงกลุ่มที่เรียงลำดับกันทางสดมภ์ ส่วนทางด้านแถวเป็นตัวแปรเชิงกลุ่มแบบนามบัญญัติ จึงไม่ต้องมีการให้คะแนนเหมือนทางสดมภ์ จากความสัมพันธ์แบบ Linear-by-linear ของตัวแบบ (3.15) จะแทนที่เทอม βv_i ด้วยพารามิเตอร์ที่ไม่มีลำดับ คือ υ_i ซึ่งเรียกว่า Row Effect ดังนั้นจะได้ตัวแบบ ดังนี้

$$\ln(m_{ij}) = u + u_i^X + u_j^Y + \upsilon_i w_j \tag{3.16}$$

ภายใต้เงื่อนไข $\sum_{i=1}^{r} u_i^X = \sum_{j=1}^{c} u_j^Y = \sum_{i=1}^{r} \upsilon_i = 0$ ตัวแบบ (3.16) จะมีจำนวนพารามิเตอร์มากกว่าตัวแบบที่เป็นอิสระกันอยู่ $r-1$ ตัว ดังนั้นตัวแบบความสัมพันธ์แบบ Linear-by-Linear จึงเป็นกรณีพิเศษของตัวแบบ Row Effect เมื่อ $\upsilon_i = \beta v_i$

เทอมของอัตราส่วนออดส์ สำหรับแถวที่ h กับ i และสดมภ์ที่ j กับ j+1 จะอยู่ในรูปแบบ

$$\ln\left(\frac{\hat{m}_{hj}\hat{m}_{i,j+1}}{\hat{m}_{h,j+1}\hat{m}_{ij}}\right) = \upsilon_i - \upsilon_h$$

โดยที่เทอมของอัตราส่วนออดส์จะมีค่าเท่ากันทั้งหมด $c-1$ คู่ของสดมภ์ที่อยู่ติดกัน ซึ่งเทอม $\upsilon_i - \upsilon_h$ จะอธิบายความแตกต่างในแถวที่ h กับ i

ส่วนตัวแบบ Column Effect ก็จะใช้หลักการเดียวกับการสร้างตัวแบบ Row Effect แต่กำหนดให้ตัวแปรทางแถวเป็นตัวแปรเชิงกลุ่มแบบเรียงลำดับ และให้คะแนน $v_1 \le v_2 \le \ldots \le v_r$ แก่กลุ่มของตัวแปรที่เรียงลำดับกันทางแถว ส่วนตัวแปรทางสดมภ์เป็นตัวแปรเชิงกลุ่มแบบนามบัญญัติ ส่วนพารามิเตอร์ที่ไม่มีการลำดับ แทนด้วย κ_j ซึ่งเรียกว่า Column Effect จะได้ตัวแบบดังนี้

$$\ln(m_{ij}) = u + u_i^X + u_j^Y + v_i \kappa_j$$

3.4 การประมาณค่าพารามิเตอร์

การประมาณค่าพารามิเตอร์ของตัวแบบเชิงเส้น โดยทั่วไปนิยมใช้วิธีภาวะน่าจะเป็นสูงและวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก (Weighted Least Squares) แต่ในบางครั้งก็ไม่สามารถใช้การประมาณดังกล่าวได้โดยตรง เนื่องจากสมการปกติที่พบอาจเป็นสมการไม่เชิงเส้น (Nonlinear Equation) จึงต้องใช้วิธีแก้สมการไม่เชิงเส้นแบบมีกระบวนการย้อนซ้ำ เรียกว่า การประมาณค่าพารามิเตอร์โดยวิธีภาวะน่าจะเป็นสูงสุดแบบย้อนซ้ำ (Iterative Maximum Likelihood Estimation) ซึ่งเป็นวิธีที่ใช้ในการประมาณค่าพารามิเตอร์ของตัวแบบในกลุ่มเอ็กโพเนนเชียล (Exponential Family Models) ที่ได้มีการประยุกต์ใช้เทคนิคการวิเคราะห์ตัวเลข (Numerical Analysis) มาช่วยในการหาตัวประมาณของพารามิเตอร์ที่ใช้กระบวนการย้อนซ้ำ และเมื่อนำมาใช้ร่วมกับวิธีภาวะน่าจะเป็นสูงสุดด้วย ทำให้เป็นวิธีที่มีประสิทธิภาพมากขึ้น

กระบวนการย้อนซ้ำมีหลายวิธี แต่ในการศึกษาครั้งนี้ จะกล่าวถึงวิธีนิวตัน – รัฟสัน (Newton – Raphson Method) เพียงวิธีเดียว โดยมีรายละเอียดดังนี้

กระบวนการย้อนซ้ำนิวตัน – รัฟสัน (Newton – Raphson)

สมการที่ใช้คือ สมการภาวะน่าจะเป็น (Likelihood Equation) ของการประมาณค่าพารามิเตอร์ที่ทำให้ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) หรือฟังก์ชันล็อกภาวะน่าจะเป็น (Log – Likelihood) มีค่าสูงสุด โดยเริ่มต้นจากการเดาค่าประมาณ (อาจเป็นค่าเฉลี่ยหรือค่ามัธยฐาน) ที่ทำให้ล็อกของฟังก์ชันภาวะน่าจะเป็นมีค่าสูงสุด และนำไปใช้ในกระบวนการย้อนซ้ำครั้งที่ 1 โดยมีหลักเกณฑ์ดังนี้

ให้ $$U = \left(\frac{\partial L}{\partial \beta_1}, \frac{\partial L}{\partial \beta_2}, \ldots \right)$$

ให้ H แทนเมตริกของอนุพันธ์อันดับสองของ L และมีสมาชิกเป็น

$$h_{ij} = \frac{\partial^2 L}{\partial \beta_i \partial \beta_j}$$

ให้ $L^{(s)}$ และ $H^{(s)}$ แทนเทอมของ L และ H ของการประมาณค่า β ครั้งที่ s หรือ $\underset{\sim}{\beta}^{(s)}$ ตามลำดับ เมื่อ s = 1, 2, … และในการย้อนซ้ำของขั้นตอนที่ s พิจารณาให้เทอม $L\left(\underset{\sim}{\beta}\right)$ ใกล้เทอม $L\left(\underset{\sim}{\beta}^{(s)}\right)$

ณ $\underset{\sim}{\beta}^{(s)}$ ในรูปแบบของเทเลอร์โพลิเนเมียล (Taylor polynomial) กำลังที่ 2 รอบ ค่า $\underset{\sim}{\beta} = \underset{\sim}{\beta}^{(s)}$ ซึ่งมีรูปแบบดังนี้

$$\begin{aligned} L\left(\underset{\sim}{\beta}\right) &= L\left(\underset{\sim}{\beta}^{(s)} + \left(\underset{\sim}{\beta} - \underset{\sim}{\beta}^{(s)}\right)\right) \\ &= L\left(\underset{\sim}{\beta}^{(s)}\right) + U^{(s)}\left(\underset{\sim}{\beta} - \underset{\sim}{\beta}^{(s)}\right) + \frac{1}{2}\left(\underset{\sim}{\beta} - \underset{\sim}{\beta}^{(s)}\right)' H^{(s)}\left(\underset{\sim}{\beta} - \underset{\sim}{\beta}^{(s)}\right) \end{aligned}$$

และแก้สมการปกติจากสมการ

$$\frac{\partial L\left(\underset{\sim}{\beta}\right)}{\partial \underset{\sim}{\beta}} = U^{(s)} + H^{(s)}\left(\underset{\sim}{\beta} - \underset{\sim}{\beta}^{(s)}\right) = 0$$

จะได้

$$\hat{\underset{\sim}{\beta}} = \underset{\sim}{\beta}^{(s)} - \left[H^{(s)}\right]^{-1} U^{(s)} \tag{3.17}$$

ให้ค่าเดาเริ่มต้น คือ $\underset{\sim}{\beta}^{0}$ ซึ่งอาจให้เท่ากับค่าประมาณ $\hat{\underset{\sim}{\beta}}$ ของการวิเคราะห์การถดถอยเชิงเส้นที่ใช้หลักของ Ordinary Least Square (OLS) แล้วนำไปแทน $\underset{\sim}{\beta}^{(s)}$ ในสมการ (3.17) และคำนวณค่าทางขวามือของสมการใหม่ เพื่อใช้สำหรับในกระบวนการย้อนซ้ำครั้งต่อไป นั่นคือการย้อนซ้ำครั้งต่อไปจะมีลักษณะเป็น

$$\underset{\sim}{\beta}^{(s+1)} = \underset{\sim}{\beta}^{(s)} - \left[H^{(s)}\right]^{-1} U^{(s)} \tag{3.18}$$

กระบวนการย้อนซ้ำจะทำเช่นนี้ไปเรื่อยๆ จนกว่าได้ค่าประมาณที่ลู่เข้า โดยมีเกณฑ์การหยุดกระบวนการย้อนซ้ำดังนี้

$$\left|\underset{\sim}{\beta}^{(s+1)} - \underset{\sim}{\beta}^{(s)}\right| = 0$$

จึงสรุปได้ว่าหลังจากการเดาค่าเริ่มต้น $\underset{\sim}{\beta}^{0}$ แล้ว แทนค่าลงไปในสมการที่ s = 0 เพื่อหา $\underset{\sim}{\beta}^{1}$ และทำต่อไปที่ s = 1, 2, … โดยค่าประมาณพารามิเตอร์ในครั้งที่ s แตกต่างจากครั้งที่ s + 1 น้อยมากหรือเกือบเท่ากับศูนย์ จะได้ค่า $\underset{\sim}{\beta}$ ครั้งที่ s + 1 คือค่าประมาณพารามิเตอร์ที่ต้องการ $\underset{\sim}{\beta}^{(s+1)}$

การประมาณค่าพารามิเตอร์ของตัวแบบล็อกลิเนียร์โดยวิธีกระบวนการย้อนซ้ำนิวตัน – รัฟสัน

จากตัวแบบ GLM สำหรับจำนวนนับของตารางการณ์จรที่มีการแจกแจงแบบปัวส์ซงหรือตัวแบบล็อกลิเนียร์ จะได้ฟังก์ชันภาวะน่าจะเป็น

$$L(\underset{\sim}{m}) = \sum_i n_i \log m_i - \sum_i m_i$$

$$= \sum_i n_i\left(\sum_j x_{ij}\beta_j\right) - \sum_i \exp\left(\sum_j x_{ij}\beta_j\right)$$

ให้ $$L\left(\underset{\sim}{\beta}\right) = \sum_i n_i\left(\sum_h x_{ih}\beta_h\right) - \sum_i \exp\left(\sum_h x_{ih}\beta_h\right)$$

จะได้ว่า $$u_j = \frac{\partial L\left(\underset{\sim}{\beta}\right)}{\partial \beta_j} = \sum_i n_i x_{ij} - \sum_i m_i x_{ij},$$

$$h_{ij} = \frac{\partial^2 L\left(\underset{\sim}{\beta}\right)}{\partial \beta_j \partial \beta_k} = -\sum_i m_i x_{ij} x_{ik},$$

ดังนั้นสำหรับการย้อนซ้ำครั้งที่ s ใดๆ จะได้ว่า

$$u_i^{(s)} = \sum_i \left(n_i - m_i^{(s)}\right)x_{ij} \quad \text{และ} \quad h_{jk}^{(s)} = -\sum_i m_i^{(s)} x_{ij} x_{ik}$$

ค่า $m_i^{(s)}$ จากครั้งที่ s สามารถหาได้จาก $\underset{\sim}{\beta}^{(s)}$ โดยคำนวณจาก $\underset{\sim}{m}^{(s)} = \exp\left(\underset{\sim}{X}\underset{\sim}{\beta}^{(s)}\right)$ และสามารถหา $\underset{\sim}{m}^{(s+1)}$ ได้จาก $\underset{\sim}{\beta}^{(s+1)}$ โดยการใช้กระบวนการย้อนซ้ำ (3.18) ดังนี้

$$\underset{\sim}{\beta}^{(s+1)} = \underset{\sim}{\beta}^{(s)} + \left[\underset{\sim}{X}'\,\mathrm{diag}\left(\underset{\sim}{m}^{(s)}\right)\underset{\sim}{X}\right]^{-1} \underset{\sim}{X}'\left(\underset{\sim}{n} - \underset{\sim}{m}^{(s)}\right) \tag{3.19}$$

โดยที่ $\mathrm{diag}\left(\underset{\sim}{m}^{(s)}\right)$ มีสมาชิกใน Main diagonal

สำหรับสมการ (3.19) สามารถคำนวณหา $\underset{\sim}{\beta}^{(s+1)}$ ได้จาก

$$\underset{\sim}{\beta}^{(s+1)} = -(H^{(s)})^{-1} r^{(s)} \tag{3.20}$$

โดยที่ $$r_j^{(s)} = \sum_i m_i^{(s)} x_{ij}\left[\log m_i^{(s)} + \frac{(n_i - m_i^{(s)})}{m_i^{(s)}}\right]$$

จะเห็นว่าใน [ ] คือเทอมแรกของเทเลอร์โพลิโนเมียลของ $\log(n_i)$ ที่ $\log(m_i^{(s)})$ ดังนั้น กระบวนการย้อนซ้ำจะเริ่มจากการเดาค่า $m_i^{(0)} = n_i$ หรือ $m_i^{(0)} = n_i + 0.5$ จากสมการ (3.20) จะให้ค่า $\underset{\sim}{\beta}^1$ และสำหรับ s > 0 กระบวนการย้อนซ้ำจะคำนวณต่อไปในทำนองเดียวกัน และขณะที่ s เพิ่มขึ้นค่าของ $\underset{\sim}{m}^{(s)}$ และ $\underset{\sim}{\beta}^{(s)}$ ก็จะลู่เข้าสู่ตัวประมาณ MLE, $\underset{\sim}{\hat{m}}$ และ $\underset{\sim}{\hat{\beta}}$ ตามลำดับ นอกจากนี้ เมตริก $H^{(s)}$ ก็จะลู่เข้าสู่ $\hat{H} = -\underset{\sim}{X}' \operatorname{diag}\left(\underset{\sim}{\hat{m}}\right)\underset{\sim}{X}$ ด้วย

ดังนั้นผลลัพธ์เพิ่มเติมจากวิธีนิวตัน-รัฟสัน คือการหาค่าของ $\operatorname{Cov}\left(\underset{\sim}{\hat{\beta}}\right) = \left[\underset{\sim}{X}' \operatorname{diag}\left(\underset{\sim}{\hat{m}}\right)\underset{\sim}{X}\right]$ โดยจะประมาณ $\operatorname{Cov}\left(\underset{\sim}{\hat{\beta}}\right)$ ด้วย $-\hat{H}^{-1}$ จากผลลัพธ์ของกระบวนการย้อนซ้ำ สามารถเขียนสมการ (3.20) ได้ใหม่ ดังนี้

$$\underset{\sim}{\beta}^{(s+1)} = \left(\underset{\sim}{X}' \hat{V}_s^{-1} \underset{\sim}{X}\right)^{-1} \underset{\sim}{X}' \hat{V}_s \underset{\sim}{Z}^{(s)}$$

เมื่อ $\underset{\sim}{Z}^{(s)}$ มีสมาชิก $\log m_i^{(s)} + \dfrac{(n_i - m_i^{(s)})}{m_i^{(s)}}$ และ $\hat{V}_s = \left[\operatorname{diag}\left(\underset{\sim}{m}^{(s)}\right)\right]^{-1}$

ดังนั้น $\underset{\sim}{\beta}^{(s+1)}$ คือตัวประมาณแบบถ่วงน้ำหนัก (WLS) สำหรับตัวแบบ $\underset{\sim}{Z}^{(s)} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ เมื่อ $\{\varepsilon_i\}$ มีความสัมพันธ์กับความแปรปรวน $\left[\dfrac{1}{m_i^{(s)}}\right]$ จะได้ว่า ถ้า $\{m_i^{(0)} = n_i\}$ แล้ว $\underset{\sim}{\beta}^{(1)}$ คือ ตัวประมาณแบบ WLS ของตัวแบบ $\log\left(\underset{\sim}{n}\right) = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$