บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ศึกษาเกี่ยวกับอำนาจการทดสอบของอัตราส่วนไลค์ลิฮูดไค-สแควร์ในการวิเคราะห์โมเดลล็อกลิเนียร์ ผู้วิจัยขอนำเสนอผลการศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 3 ตอน ดังนี้ ตอนที่ 1 ความรู้ทั่วไปเกี่ยวกับโมเดลล็อกลิเนียร์ ตอนที่ 2 อำนาจการทดสอบ และตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับอัตราส่วนไลค์ลิฮูดไค-สแควร์ โมเดลล็อกลิเนียร์ และอำนาจการทดสอบ

ตอนที่ 1 ความรู้ทั่วไปเกี่ยวกับโมเดลล็อกลิเนียร์

โมเดลล็อกลิเนียร์เป็นการประเมินความสัมพันธ์และอิทธิพลของตัวแปรที่ไม่ต่อเนื่องซึ่งวิเคราะห์โดยใช้ลอการิทึมธรรมชาติฐาน e การวิเคราะห์ล็อกลิเนียร์เป็นวิธีวิทยาที่เหมาะสมในการวิเคราะห์ข้อมูลในตารางการณ์จรพหุมิติ (Le, 1998; Baglivo, Oliver และ Pagano, 1988) การวิเคราะห์ล็อกลิเนียร์เป็นวิธีที่ผสมผสานระหว่างสถิติวิเคราะห์ 3 ตัวคือ การวิเคราะห์การถดถอย การวิเคราะห์ความแปรปรวน การทดสอบภาวะสารูปสนิทแบบไค-สแควร์ โมเดลอิ่มตัวของโมเดลล็อกลิเนียร์มีลักษณะเหมือนกับโมเดลในการวิเคราะห์ความแปรปรวน (Bonett และ Bentler, 1983) ในโมเดลล็อกลิเนียร์จะมีการวิเคราะห์ความสัมพันธ์แบบสมมาตรเป็นการศึกษาความสัมพันธ์โดยไม่สนใจพิจารณาตัวแปรตาม ส่วนการวิเคราะห์ความสัมพันธ์แบบอสมมาตรเป็นการศึกษาความสัมพันธ์โดยพิจารณาตัวแปรตาม 1 ตัว ประเภทหลังนี้สามารถเทียบเคียงได้กับการวิเคราะห์ความแปรปรวน โมเดลการวิเคราะห์ล็อกลิเนียร์เป็นโมเดลที่ทำนายโอกาสในการเกิดค่าความถี่ที่คาดหวังของตารางการณ์จรด้วยชุดอิทธิพลของตัวแปรอิสระโดยที่โมเดลเป็นแบบคูณ (ทวีพร บุญวานิช, 2541)

1. ลักษณะทั่วไปของโมเดลล็อกลิเนียร์

กำหนดตัวแปร 2 ตัวคือ A และ B ตัวแปรแต่ละตัวมี 2 ระดับคือ i = 1, 2 และ j = 1, 2 ให้ p_{ij} คือ ความน่าจะเป็นของ i สำหรับตัวแปร A และ j สำหรับตัวแปร B กำหนด $v_{ij} = \ln p_{ij}$ ในตารางการณ์จร 2 ทางขนาด 2×2 สามารถเขียนแทนได้ด้วย

$$F_{ij} = \eta \tau_i^A \tau_j^B \tau_{ij}^{AB}$$

ในการวิเคราะห์โมเดลล็อกลิเนียร์ทั่วไปรูปแบบของโมเดลจะเป็นแบบคูณ แต่ถ้านำลอการิทึมมาใช้รูปแบบของสมการจะเปลี่ยนเป็นสมการเส้นตรง ดังนั้นจึงเป็นสมการเชิงเส้นในรูปของลอการิทึม

จากสมการ $$F_{ij} = \eta\tau_i^A\tau_j^B\tau_{ij}^{AB}$$

เปลี่ยนรูปโดยใช้ลอการิทึมจะได้สมการ

$$Ln(F_{ij}) = Ln(\eta\tau_i^A\tau_j^B\tau_{ij}^{AB}) = Ln(\eta) + Ln(\tau^A) + Ln(\tau_j^B) + Ln(\tau_{ij}^{AB})$$

กำหนด λ แทนค่าลอการิทึมของ τ

μ แทนค่าลอการิทึมของ η

v_{ij} แทนค่าลอการิทึมของ F_{ij}

จะได้โมเดลล็อกลิเนียร์ สำหรับตาราง 2 ทาง ในรูปการบวกดังนี้

$$v_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

โดยที่
$$v_{i.} = \frac{1}{J}\sum_j v_{ij}$$
$$v_{.j} = \frac{1}{I}\sum_i v_{ij}$$
$$v_{..} = \frac{1}{IJ}\sum_i\sum_j v_{ij}$$

จะได้
$$\lambda_i^A = v_{i.} - v_{..}$$
$$\lambda_j^B = v_{.j} - v_{..}$$

และ
$$\lambda_{ij}^{AB} = v_{ij} - v_{i.} - v_{.j} + v_{..}$$
$$= (v_{ij} - v_{..}) - [(v_{i.} - v_{..}) + (v_{.j} - v_{..})]$$

จะเรียก λ_i^A, λ_j^B และ λ_{ij}^{AB} ว่า อิทธิพลหลักของ A (main effect of A) อิทธิพลหลักของ B (main effect of B) และ ปฏิสัมพันธ์สองปัจจัยของ A และ B (two-factor interaction) ตามลำดับ และจะกำหนดนิยามของอิทธิพลต่างๆ ดังนี้

$$\sum_i \lambda_i^A = \sum_j \lambda_j^B = \sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} = \sum_i\sum_j \lambda_{ij}^{AB} = 0$$

เรียกโมเดล $v_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$ ว่า โมเดลอิ่มตัว (saturated model) คือ โมเดลที่สามารถอธิบายอิทธิพลที่เป็นไปได้ทั้งหมดที่มีต่อพารามิเตอร์ และเรียกโมเดลที่ $\lambda^{AB} = 0$ ว่า โมเดลอิสระต่อกัน (mutual independence model) คือ โมเดลที่ได้เกิดอิทธิพลหลักจากตัวแปร A และตัวแปร B ซึ่งจะได้โมเดลล็อกลิเนียร์ที่มีลักษณะดังนี้ $v_{ijk} = \mu + \lambda_i^A + \lambda_j^B$ ส่วนโมเดลที่ได้รับอิทธิพลจากตัวแปร A หรือตัวแปร B เพียงตัวเดียวเรียกว่า โมเดลความน่าจะเป็นเท่าอย่างมีเงื่อน

ไข (conditional equiprobability model) จะได้โมเดลล็อกลิเนียร์ดังนี้ $v_{ijk} = \mu + \lambda_i^A$ (กรณีได้รับอิทธิพลจากตัวแปร A) และสุดท้ายโมเดลที่ไม่ได้รับอิทธิพลจากตัวแปรใด ๆ เลย เรียกโมเดลนี้ว่าโมเดลความน่าจะเป็นเท่า (mutual equiprobability model) มีลักษณะ $v_{ijk} = \mu$

ในทำนองเดียวกันสามารถอธิบายตาราง 3 ทางขนาด $2 \times 2 \times 2$ ได้ดังนี้

กำหนดตัวแปร 3 ตัวคือ A, B และ C ตัวแปรแต่ละตัวมี 2 ระดับคือ i = 1, 2, j = 1, 2 และ k = 1, 2 ให้ p_{ijk} คือความน่าจะเป็นของตัวแปร A, B และ C กำหนด $v_{ijk} = \ln p_{ijk}$ ในตารางการณ์จร 3 ทางขนาด $2 \times 2 \times 2$ สามารถเขียนโมเดลล็อกลิเนียร์ในรูปการคูณได้ดังนี้

$$F_{ijk} = \eta \tau_i^A \tau_j^B \tau_k^C \tau_{ij}^{AB} \tau_{ik}^{AC} \tau_{jk}^{BC} \tau_{ijk}^{ABC}$$

ในลักษณะเดียวกับตาราง 2 ทาง เมื่อนำค่าลอการิทึมมาใช้รูปแบบสมการจะเปลี่ยนเป็นเส้นตรง

ดังสมการ $$v_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

เมื่อ μ คือ ค่าคงที่ โดยที่ $\sum_{ijk} p_{ijk} = 1$

$\lambda_i^A, \lambda_j^B, \lambda_k^C$ คือ อิทธิพลหลัก (main effects) ในระดับที่ i ของตัวแปร A, ระดับที่ j ของตัวแปร B และระดับที่ k ของตัวแปร C ตามลำดับ

$\lambda_{ij}^{AB}, \lambda_{ik}^{AC}, \lambda_{jk}^{BC}$ คือ อิทธิพลปฏิสัมพันธ์ 2 ปัจจัย (two-factor interactions) ของ A_i และ B_j, A_i และ C_k, และ B_j และ C_k ตามลำดับ

λ_{ijk}^{ABC} คือ อิทธิพลปฏิสัมพันธ์ 3 ปัจจัย (three-factor interactions) ของ A_i, B_j และ C_k

ถ้า $\lambda^{AB} = 0$ หมายความว่า $\lambda_{ij}^{AB} = 0$ ทุกระดับของ i และ j ลักษณะเดียวกันถ้า $\lambda^{AC} = 0$ หมายความว่า $\lambda_{ik}^{AC} = 0$ ทุกระดับของ i และ k และถ้า $\lambda^{BC} = 0$ หมายความว่า $\lambda_{jk}^{BC} = 0$ ทุกระดับของ j และ k

จากโมเดลอิ่มตัว สำหรับตาราง 3 ทางจะได้

$$v_{...} = \frac{\sum_i \sum_j \sum_k v_{ijk}}{IJK}$$

$$v_{i..} = \frac{\sum_j \sum_k v_{ijk}}{JK}$$

$$v_{.j.} = \frac{\sum_i \sum_k v_{ijk}}{IK}$$

$$v_{..k} = \frac{\sum_i \sum_j v_{ijk}}{IJ}$$

$$v_{ij.} = \frac{\sum_k v_{ijk}}{K}$$

$$v_{i.k} = \frac{\sum_j v_{ijk}}{J}$$

$$v_{.jk} = \frac{\sum_i v_{ijk}}{I}$$

และจะได้ อิทธิพลหลัก (main effects) ของ A, B และ C ดังนี้

$$\lambda_i^A = v_{i..} - v_{...}, (i = 1,\ldots,I)$$
$$\lambda_j^A = v_{.j.} - v_{...}, (j = 1,\ldots,J)$$
$$\lambda_k^A = v_{..k} - v_{...}, (k = 1,\ldots,K)$$

นอกจากนี้จะได้ ปฏิสัมพันธ์ 2 ปัจจัย (two-factor interactions) ดังนี้

$$\lambda_{ij}^{AB} = v_{ij.} - v_{i..} - v_{.j.} + v_{...}, (i = 1,\ldots,I; j = 1,\ldots,J)$$
$$\lambda_{ik}^{AC} = v_{i.k} - v_{i..} - v_{..k} + v_{...}, (i = 1,\ldots,I; k = 1,\ldots,K)$$
$$\lambda_{jk}^{BC} = v_{.jk} - v_{.j.} - v_{..k} + v_{...}, (j = 1,\ldots,J; k = 1,\ldots,K)$$

และจะได้ ปฏิสัมพันธ์ 3 ปัจจัย (three-factor interaction) ดังนี้

$$\lambda_{ijk}^{ABC} = v_{ijk} - v_{ij.} - v_{i.k} - v_{.jk} + v_{i..} + v_{.j.} + v_{..k} - v_{...};$$
$$(i = 1,\ldots,I, j = 1,\ldots,J, k = 1,\ldots,K)$$

และกำหนดนิยามของอิทธิพลต่างๆ ดังนี้

$$\sum_i \lambda_i^A = \sum_j \lambda_j^B = \sum_k \lambda_k^C = 0$$
$$\sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} = \sum_i \lambda_{ik}^{AC} = \sum_j \lambda_{ik}^{AC} = \sum_j \lambda_{jk}^{BC} = \sum_k \lambda_{jk}^{BC} = 0$$
$$\sum_i \lambda_{ijk}^{ABC} = \sum_j \lambda_{ijk}^{ABC} = \sum_k \lambda_{ijk}^{ABC} = 0$$

2. โมเดลระดับลดหลั่น (Hierarchical Model)

โมเดลระดับลดหลั่น หมายถึง โมเดลที่มีอิทธิพลเนื่องจากตัวแปรที่มีอันดับสูงกว่า (higher order effect) และมีโมเดลที่มีอิทธิพลเนื่องจากตัวแปรที่มีอันดับต่ำกว่า (lower order effect) รวมอยู่ด้วย เช่น ในโมเดลอิ่มตัวของ 2 ตัวแปรมีลักษณะดังนี้

$$v_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

ซึ่งในโมเดลนี้ก็จะรวมอิทธิพลเนื่องจากตัวแปรที่มีอันดับต่ำกว่าไว้ด้วย คือ

$$v_{ijk} = \mu + \lambda_i^A + \lambda_j^B$$

และ

$$v_{ijk} = \mu + \lambda_i^A$$

ในทางตรงกันข้ามถ้าโมเดลมีลักษณะดังนี้

$$v_{ijk} = \mu + \lambda_i^A + \lambda_{ij}^{AB}$$

คือ ไม่ได้รวมอิทธิพลหลักจากตัวแปร B ไว้ด้วย โมเดลลักษณะนี้ก็จะไม่เป็น โมเดลระดับลดหลั่น

3. ขั้นตอนการวิเคราะห์โมเดลล็อกลิเนียร์

การวิเคราะห์ล็อกลิเนียร์มี 6 ขั้นตอน ขั้นตอนแรก สร้างตารางการณ์จรและคำนวณความถี่ที่คาดหวังในแต่ละเซลล์ของตาราง ขั้นที่สอง สร้างโมเดลล็อกลิเนียร์ ขั้นที่สาม ประมาณค่าพารามิเตอร์ ขั้นที่สี่ คัดเลือกโมเดล ขั้นที่ห้า ตรวจสอบความเหมาะสมของโมเดล และขั้นตอนสุดท้าย ขั้นตอนที่หก ประมาณค่าพารามิเตอร์ มีรายละเอียด ดังนี้

ขั้นตอนที่ 1 การสร้างตารางการณ์จร และการคำนวณความถี่ที่คาดหวัง

1.1 การสร้างตารางการณ์จร การวิเคราะห์ล็อกลิเนียร์เป็นการวิเคราะห์ที่ข้อมูลอยู่ในรูปตารางการณ์จร ซึ่งเป็นไปตามข้อตกลงเบื้องต้นของไค-สแควร์ที่ว่าหน่วยตัวอย่างแต่ละหน่วยจะต้องถูกแจงนับอย่างเป็นอิสระและจำแนกลงได้เพียงเซลล์ใดเซลล์หนึ่งเท่านั้น และค่าความถี่ที่คาดหวังต้องมากกว่า 5 ส่วนขนาดของตารางการณ์จรจะเป็นเท่าใดนั้นขึ้นอยู่กับจำนวนตัวแปรและค่าของตัวแปรเหล่านั้น เช่น กรณีมี 2 ตัวแปร ตัวแปร A มี i ค่า ตัวแปร B มี j ค่าตารางการณ์จรจะมีขนาด i X j การพิจารณาอิทธิพลหลักจะพิจารณาจากผลรวมของแถวหรือหลักที่เป็นผลรวมความถี่ของตัวแปรนั้น ส่วนอิทธิพลปฏิสัมพันธ์ของตัวแปร A และตัวแปร B จะพิจารณาจากความถี่ในเซลล์ ij

1.2 การคำนวณค่าความถี่ที่คาดหวัง การวิเคราะห์ข้อมูลโดยการใช้สถิติทดสอบไค-สแควร์โดยทั่วไปนักวิจัยจะนำค่าความถี่ที่สังเกตได้มาคำนวณความถี่ที่คาดหวัง แล้วนำไปหาค่าสถิติไค-สแควร์เพื่อทดสอบสมมติฐานต่อไป การวิเคราะห์ล็อกลิเนียร์ใช้หลักเดียวกันต่างกันตรงที่การวิเคราะห์ล็อกลิเนียร์มีโมเดลในการวิเคราะห์หลายโมเดล การคำนวณค่าความถี่ที่คาดหวังสามารถพิจารณาได้ดังนี้

ความถี่ในเซลล์ที่คาดหวังของตารางการณ์จร (e_{ij}) เป็นฟังก์ชันของพารามิเตอร์ที่เป็นตัวแทนของคุณลักษณะของตัวแปรกลุ่มและเป็นการแสดงความสัมพันธ์ซึ่งกันและกัน ในการประเมินว่าโมเดลสามารถอธิบายความเหมาะสมของข้อมูลซึ่งต้องพิจารณาความถี่ที่คาดหวัง (e_{ij}) กับความถี่ที่สังเกตได้ (o_{ij}) ประกอบกัน

โมเดลล็อกลิเนียร์ภายใต้ตารางการณ์จรมี 2 วิธีใหญ่ๆ (two major approaches) คือ (1) โมเดลล็อกลิเนียร์ทั่วไป (general loglinear model) เป็นการวิเคราะห์ที่ไม่ต้องการพิจารณาตัวแปรตาม (2) โมเดลโลจิทล็อกลิเนียร์ (logit loglinear model) เป็นการวิเคราะห์ที่สนใจตัวแปรตาม โดยจะเลือกตัวแปรตาม 1 ตัว

โมเดลที่สามารถอธิบายอิทธิพลที่เป็นไปได้ทั้งหมดที่มีต่อพารามิเตอร์ เรียกโมเดลนั้นว่าโมเดลอิ่มตัว (saturated models) แสดงได้ดังนี้

$$F_{ij} = \eta\tau_i^A\tau_j^B\tau_{ij}^{AB}$$

เมื่อ F_{ij} แทนจำนวนหรือความถี่ของแต่ละกรณีในเซลล์ที่ i หลักที่ j (คือความถี่ที่คาดหวังที่ต้องการถ้าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์)

η (eta) แทนค่าเฉลี่ยเลขคณิตของความถี่ของแต่ละเซลล์ (ซึ่งจะคล้ายกับจุดตัด (intercept) ในสมการถดถอย)

τ (tau) แทนอิทธิพลหลัก

τ_i^A แทนอิทธิพลหลักที่เกิดจากแถวนอนที่ i

τ_j^B แทนอิทธิพลหลักที่เกิดจากแถวตั้งที่ j

τ_{ij}^{AB} แทนอิทธิพลปฏิสัมพันธ์ระหว่างแถวนอนที่ i แถวตั้งที่ j

ในการวิเคราะห์โมเดลล็อกลิเนียร์ทั่วไปรูปแบบของโมเดลจะเป็นแบบคูณ แต่ถ้านำลอการิทึมมาใช้ รูปแบบของสมการจะเปลี่ยนเป็นสมการเส้นตรง ดังนั้นจึงเป็นสมการเชิงเส้นในรูปของลอการิทึม ดังสมการ $F_{ij} = \eta\tau_i^A\tau_j^B\tau_{ij}^{AB}$

เปลี่ยนรูปโดยใช้ลอการิทึมจะได้สมการ

$$Ln(F_{ij}) = Ln(\eta\tau_i^A\tau_j^B\tau_{ij}^{AB}) = Ln(\eta) + Ln(\tau_i^A) + Ln(\tau_j^B) + Ln(\tau_{ij}^{AB})$$

หรือ

$$v_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

เมื่อ λ แทนค่าลอการิทึมของ τ

μ แทนค่าลอการิทึมของ η

v_{ij} แทนค่าลอการิทึมของ F_{ij}

โมเดลของล็อกลิเนียร์แบ่งออกเป็นโมเดลตามการคำนวณค่าความถี่ที่คาดหวัง ในกรณีมี 2 ตัวแปรแต่ละตัวมี 2 ค่า การคำนวณค่าความถี่ที่คาดหวังของโมเดลจะมีทั้งหมด 4 ลักษณะซึ่งจะได้โมเดลที่ใช้อธิบายความถี่ที่คาดหวัง 4 รูปแบบ คือ โมเดลความน่าจะเป็นเท่า (mutual equiprobability model) โม เดลความน่าจะเป็นเท่าอย่างมีเงื่อนไข (conditional equiprobability model) โมเดลอิสระต่อกัน (mutual independence model) และโมเดลอิ่มตัว (saturated model) ดังนี้

1.2.1 <u>โมเดลความน่าจะเป็นเท่า (Mutual Equiprobability Model)</u> คือ โมเดลที่ทำนายความถี่ที่คาดหวังโดยไม่มีอิทธิพลจากตัวแปรใดๆเลย ค่าความถี่ที่คาดหวังในแต่ละเซลล์มีค่าเท่ากันทั้ง 4 เซลล์ การคำนวณหาความถี่ที่คาดหวังใช้จำนวนกลุ่มตัวอย่างทั้งหมดเป็นฐานในการคำนวณ แล้วหารด้วยจำนวนเซลล์ จึงได้ค่าความถี่ที่คาดหวังของแต่ละเซลล์ ซึ่งจะได้โมเดลล็อกลิเนียร์ที่มีลักษณะดังนี้

$$F_{ij} = \eta \;\; = [\;]$$

ลักษณะการคำนวณค่าความถี่ที่คาดหวัง แสดงได้ดังตัวอย่างต่อไปนี้

ตารางที่ 1 ตัวอย่างการคำนวณค่าความถี่ที่คาดหวัง ตามลักษณะโมเดลความน่าจะเป็นเท่า

A\B	B_1	B_2	TOTAL
A_1	n/4	n/4	$n_{1.}$
A_2	n/4	n/4	$n_{2.}$
TOTAL	$n_{.1}$	$n_{.2}$	N

1.2.2 <u>โมเดลความน่าจะเป็นเท่าอย่างมีเงื่อนไข (Conditional Equiprobability Model)</u> โมเดลนี้จะมี 2 ลักษณะคือโมเดลที่ทำนายความถี่ที่คาดหวังโดยได้อิทธิพลหลักจากตัวแปร A หรือ B ตัวแปรใดตัวแปรหนึ่ง ถ้าค่าความถี่ที่คาดหวังค่าของ n_{11} เท่ากับ n_{12} และ n_{21} เท่ากับ n_{22} แสดงว่าได้รับอิทธิพลหลักจากตัวแปร A แต่ถ้าค่าความถี่ที่คาดหวังค่าของ n_{11} เท่ากับ n_{21}

และ n_{12} เท่ากับ n_{22} แสดงว่าได้รับอิทธิพลหลักจากตัวแปร B การคำนวณค่าความถี่ที่คาดหวัง นำค่าความถี่ที่สังเกตได้ของแถวตั้งที่ B_1 หารด้วยจำนวนเซลล์ของแถวตั้งนั้น และให้นำค่าความถี่ที่สังเกตได้ของแถวตั้งที่ B_2 หารด้วยจำนวนเซลล์ของแถวตั้งนั้น จึงได้ค่าความถี่ที่คาดหวังของแถวตั้งที่ *i* แถวนอนที่ *j* จะได้โมเดลล็อกลิเนียร์ที่มีลักษณะดังนี้

$$F_{ij} = \eta \tau_j^B = [B]$$

ลักษณะการคำนวณค่าความถี่ที่คาดหวัง แสดงได้ดังตัวอย่างข้างล่าง

ตารางที่ 2 ตัวอย่างการคำนวณค่าความถี่ที่คาดหวัง ตามลักษณะโมเดลความน่าจะเป็นเท่าอย่างมีเงื่อนไข

A\B	B_1	B_2	TOTAL
A_1	$n_{.1}/2$	$n_{.2}/2$	$n_{1.}$
A_2	$n_{.1}/2$	$n_{.2}/2$	$n_{2.}$
TOTAL	$n_{.1}$	$n_{.2}$	N

1.2.3 <u>โมเดลอิสระต่อกัน (Mutual Independence Model)</u> คือ โมเดลที่ทำนายความถี่ที่คาดหวังโดยได้อิทธิพลหลักจากตัวแปร A และจากตัวแปร B

การคำนวณค่าความถี่คาดหวังของเซลล์ A_1B_1 คำนวณจาก ผลคูณของค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวนอนที่ 1 ($n_{1.}$) กับค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวตั้งที่ 1 ($n_{.1}$) หารด้วยจำนวนกลุ่มตัวอย่างทั้งหมด (N)

การคำนวณค่าความถี่คาดหวังของเซลล์ A_1B_2 คำนวณจาก ผลคูณของค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวนอนที่ 1 ($n_{1.}$) กับค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวตั้งที่ 2 ($n_{.2}$) หารด้วยจำนวนกลุ่มตัวอย่างทั้งหมด (N)

การคำนวณค่าความถี่คาดหวังของเซลล์ A_2B_1 คำนวณจาก ผลคูณของค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวนอนที่ 2 ($n_{2.}$) กับค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวตั้งที่ 1 ($n_{.1}$) หารด้วยจำนวนกลุ่มตัวอย่างทั้งหมด (N)

การคำนวณค่าความถี่คาดหวังของเซลล์ A_2B_2 คำนวณจาก ผลคูณของค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวนอนที่ 2 ($n_{2.}$) กับค่าความถี่ที่สังเกตได้ของค่าสัดส่วนส่วนริมของแถวตั้งที่ 2 ($n_{.2}$) หารด้วยจำนวนกลุ่มตัวอย่างทั้งหมด (N)

จะได้โมเดลล็อกลิเนียร์ที่มีลักษณะดังนี้

$$F_{ij} = \eta\tau_i^A\tau_j^B = [A][B]$$

ลักษณะการคำนวณค่าความถี่ที่คาดหวัง แสดงได้ดังตัวอย่างข้างล่าง

ตารางที่ 3 ตัวอย่างการคำนวณค่าความถี่ที่คาดหวัง ตามลักษณะโมเดลอิสระต่อกัน

A\B	B_1	B_2	TOTAL
A_1	$n_{1.}n_{.1}/n$	$n_{1.}n_{.2}/n$	$n_{1.}$
A_2	$n_{2.}n_{.1}/n$	$n_{2.}n_{.2}/n$	$n_{2.}$
TOTAL	$n_{.1}$	$n_{.2}$	N

1.2.4 <u>โมเดลอิ่มตัว (Saturated Model)</u> คือโมเดลที่ทำนายความถี่ที่คาดหวังโดยได้อิทธิพลหลักจากตัวแปร A, อิทธิพลหลักจากตัวแปร B และอิทธิพลปฏิสัมพันธ์ระหว่าง A กับ B ในกรณีนี้ค่าความถี่ที่คาดหวังจะเท่ากับค่าความถี่ที่สังเกตได้ในแต่ละเซลล์นั่นเอง จะได้โมเดลล็อกลิเนียร์ที่มีลักษณะดังนี้

$$F_{ij} = \eta\tau_i^A\tau_j^B\tau_{ij}^{AB} = [AB]$$

ลักษณะการคำนวณค่าความถี่ แสดงได้ดังตัวอย่างข้างล่าง

ตารางที่ 4 ตัวอย่างการคำนวณค่าความถี่ที่คาดหวัง ตามลักษณะโมเดลอิ่มตัว

A\B	B_1	B_2	TOTAL
A_1	O_{11}	O_{12}	$n_{1.}$
A_2	O_{21}	O_{22}	$n_{2.}$
TOTAL	$n_{.1}$	$n_{.2}$	N

เมื่อสร้างตารางการณ์จรและจำแนกความถี่ที่คาดหวังลงตาราง โดยพิจารณาตามเทอมอิทธิพลต่างๆ ขั้นต่อไปเป็นการสร้างสมการเพื่อเป็นสัญลักษณ์แทนโมเดลต่างๆ

สำหรับตาราง 3 ทางขนาด 2 × 2 × 2 การคำนวณค่าความถี่ที่คาดหวังสามารถคำนวณได้ 2 วิธีคือ การคำนวณทางตรง คือใช้ ค่าสถิติที่เพียงพอต่ำที่สุด (minimal sufficient statistics) และการคำนวณแบบทวนซ้ำ (iterative method)

ตารางที่ 5 การคำนวณค่าความถี่ที่คาดหวังโดยใช้ค่าสถิติที่ที่เพียงพอต่ำที่สุด (minimal sufficient statistics) สำหรับตาราง 3 ทาง

model	expected frequency
[ ]	N/8
[A]	$O_{ijk}/4$
[A][B]	$(O_{i..})(O_{.j.})/2N$
[A][B][C]	$(O_{i..})(O_{.j.})(O_{..k})/N^2$
[A][B][C][AB]	$(O_{ij.})(O_{..k})/N$
[A][B][C][AB][AC]	$(O_{ij.})(O_{i.k})/O_{i..}$
[A][B][C][AB][AC][BC]	iterative method

โมเดลที่ 1 : โมเดลที่ไม่มีอิทธิพลจากตัวแปรใดๆ (no effect model: [ ]) ค่าความถี่ที่คาดหวังคำนวณจากกลุ่มตัวอย่างทั้งหมดหารด้วยจำนวนเซลล์ ($\frac{n}{8}$) กล่าวคือหากกลุ่มตัวอย่างเท่ากับ 100 ตาราง 3 ทางมีจำนวนเซลล์เท่ากับ 8 ดังนั้นค่าความถี่ที่คาดหวังในแต่ละเซลล์มีค่าประมาณ 12 ถึง 13

โมเดลที่ 2 : โมเดลที่เกิดอิทธิพลหลัก 1 ตัว (single main effect model: [A]) ค่าความที่คาดหวังของอิทธิพลหลัก 1 ตัว คำนวณจากค่าความถี่ที่สังเกตได้ของแต่ละเซลล์ (o_{ijk})หารด้วย 4 เช่น เซลล์ $A_1B_1C_1$ มีค่าความถี่ที่สังเกตได้เท่ากับ 12 เมื่อหารด้วย 4 จะได้ค่าความถี่ที่คาดหวังของเซลล์ $A_1B_1C_1$ เท่ากับ 3

โมเดลที่ 3 : โมเดลที่เกิดอิทธิพลหลัก 2 ตัว (two main effects model: [A][B]) การคำนวณค่าความถี่ที่คาดหวังของโมเดลอิทธิพลหลัก 2 ตัว กรณีที่ต้องการค่าความถี่ที่คาดหวังของเซลล์ $A_iB_jC_k$ ให้นำผลบวกของค่าความถี่ที่สังเกตได้ของเซลล์ A_i ทั้งหมด ($o_{i..}$) คูณกับผลบวกกับค่าความถี่ที่สังเกตได้ของเซลล์ B_j ทั้งหมด ($o_{.j.}$) แล้วหารด้วย 2 เท่าของกลุ่มตัวอย่าง ($2n$) เช่น กรณีที่ต้องการค่าความถี่ที่คาดหวังของเซลล์ $A_1B_1C_1$ให้นำผลบวกของค่าความถี่ที่สังเกตได้ของเซลล์ A_1 ทั้งหมดในที่นี้มี 4 เซลล์คือ $A_1B_1C_1$, $A_1B_1C_2$, $A_1B_2C_1$ และ $A_1B_2C_2$ คูณกับ ผลรวมค่าความถี่ที่สังเกตได้ของเซลล์ B_1 ทั้งหมด ในที่นี้มี 4 เซลล์คือ $A_1B_1C_1$, $A_1B_1C_2$, $A_2B_1C_1$ และ $A_2B_1C_2$ จากนั้นจึงหารด้วย 2เท่าของขนาดกลุ่มตัวอย่าง ถ้ากลุ่มตัวอย่าง 100 จะหารด้วย 200

โมเดลที่ 4 : โมเดลอิทธิพลหลักทั้งหมด (all main effects model: [A][B][C]) (model if independence of factors) คือ โมเดลที่เกิดอิทธิพลหลักจากตัวแปร A, ตัวแปร B และตัวแปร C กรณีที่ต้องการค่าความถี่ที่คาดหวังของเซลล์ $A_iB_jC_k$ ให้นำ ผลบวกของค่าความถี่ที่สังเกตได้ของเซลล์ A_i ทั้งหมด ($o_{i..}$) คูณกับผลบวกกับค่าความถี่ที่สังเกตได้ของเซลล์ B_j ทั้งหมด ($o_{.j.}$) คูณกับผลบวกกับค่าความถี่ที่สังเกตได้ของเซลล์ C_k ทั้งหมด ($o_{..k}$) แล้วนำไปหารด้วย ขนาดกลุ่มตัวอย่างยกกำลังสอง ถ้ากลุ่มตัวอย่างเท่ากับ 100 จะหารด้วย 10,000 ยกตัวอย่างเช่น กรณีต้องการหาความถี่ที่สังเกตได้ของเซลล์ $A_1B_1C_1$ ให้หาผลบวกของเซลล์ A_1 ทั้งหมดคือ $A_1B_1C_1$, $A_1B_1C_2$, $A_1B_2C_1$ และ $A_1B_2C_2$ คูณกับผลบวกของเซลล์ B_1 ทั้งหมดคือ $A_1B_1C_1$, $A_1B_1C_2$, $A_2B_1C_1$ และ $A_2B_1C_2$ คูณกับผลบวกของเซลล์ C_1 ทั้งหมดคือ $A_1B_1C_1$, $A_1B_2C_1$, $A_2B_1C_1$ และ $A_2B_2C_1$ แล้วนำผลที่ได้ไปด้วย ขนาดกลุ่มตัวอย่างยกกำลังสอง

โมเดลที่ 5 : โมเดลอิทธิพลหลักทั้งหมดและอิทธิพลปฏิสัมพันธ์ 2 ปัจจัย 1 ตัว (all main effects and single interaction effect model: [A][B][C][AB]) การคำนวณค่าความถี่ที่คาดหวังของเซลล์ $A_iB_jC_k$ ให้นำ ผลบวกของค่าความถี่ที่สังเกตได้ของเซลล์ A_iB_j ทั้งหมด ($o_{ij.}$) คูณกับผลบวกค่าความถี่ที่สังเกตได้ของเซลล์ C_k ทั้งหมด ($o_{..k}$) หารด้วยขนาดกลุ่มตัวอย่าง เช่น กรณีต้องการค่าความถี่ที่คาดหวังของเซลล์ $A_1B_1C_1$ ให้นำผลบวกของความถี่ที่สังเกตได้ของเซลล์ $A_1B_1C_1$ และ$A_1B_1C_2$ คูณกับ ค่าความถี่ที่สังเกตได้ของเซลล์ C_1 ในที่นี้มี 4 เซลล์คือ $A_1B_1C_1$, $A_1B_2C_1$, $A_2B_1C_1$ และ $A_2B_2C_1$ แล้วนำผลลัพธ์ไปหารด้วย ขนาดกลุ่มตัวอย่าง

โมเดลที่ 6 : โมเดลอิทธิพลหลักทั้งหมดและมีอิทธิพลปฏิสัมพันธ์ 2 ปัจจัย 2 ตัว (all main effects and two interaction effects model: [A][B][C][AB][AC]) การคำนวณค่าความถี่ที่คาดหวังของเซลล์ $A_iB_jC_k$ ให้นำผลบวกของค่าความถี่ที่สังเกตได้ของเซลล์ A_iB_j ($o_{ij.}$) คูณด้วยผลบวกของค่าความถี่ที่สังเกตได้ของเซลล์ A_iC_k ($o_{i.k}$) แล้วนำไปหารด้วย ค่าความถี่ที่สังเกตได้ของเซลล์ A_i ทั้งหมด ($o_{i..}$) เช่น ต้องการหาค่าความถี่ที่คาดหวังของเซลล์ $A_1B_1C_1$ ให้หาผลบวกของเซลล์ $A_1B_1C_1$ และ $A_1B_1C_2$ นำไปคูณกับผลบวกของเซลล์ $A_1B_1C_1$ และ $A_1B_2C_1$ จากนั้นนำผลลัพธ์ที่ได้ไปหารด้วย ผลบวกระหว่างเซลล์ $A_1B_1C_1$, $A_1B_1C_2$, $A_1B_2C_1$ และ $A_1B_2C_2$

โมเดลที่ 7 : โมเดลอิทธิพลหลักทั้งหมดและมีอิทธิพลปฏิสัมพันธ์ 2 ปัจจัย 3 ตัว (all main effects and three interaction effects model: [A][B][C][AB][AC][BC]) การคำนวณค่าความถี่ที่คาดหวังของโมเดลใช้วิธีการคำนวณค่าความถี่ที่คาดหวัง โดยใช้การทวนซ้ำ

การคำนวณด้วยวิธี IPF

ขั้นที่ 1 กำหนดค่าตั้งต้นทุกๆ ค่าของ i,j และ k ให้มีค่าเท่ากับ 1 นั่นคือ $e_{ijk}^{(0)} = 1$

กำหนด v รอบในการวนซ้ำ เริ่มต้นที่ 0

ขั้นที่ 2 คำนวณจากค่าสัดส่วนส่วนริมของ AB ด้วย $e_{ijk}^{3v+1} = \dfrac{o_{ij.}e_{ijk}^{3v}}{e_{ij.}^{3v}}$

ขั้นที่ 3 นำผลที่ได้จากขั้นที่ 2 มาคำนวณกับค่าสัดส่วนส่วนริมของ AC ด้วย

$$e_{ijk}^{3v+2} = \frac{o_{i.k}e_{ijk}^{3v+1}}{e_{i.k}^{3v+1}}$$

ขั้นที่ 4 นำผลที่ได้จากขั้นที่ 3 มาคำนวณกับค่าสัดส่วนริมของ BC ด้วย

$$e_{ijk}^{3(v+1)} = \frac{o_{.jk}e_{ijk}^{3v+2}}{e_{.jk}^{3v+2}}$$

เมื่อจบขั้นที่ 1 จะได้รอบที่ 1 ของการวนซ้ำ (first cycle of the iteration) ของการคำนวณความถี่ที่คาดหวัง ทำซ้ำขั้นที่ 2 – 4 อีกครั้ง จะได้ v = 1,2,... จนกระทั่งได้ค่า $| e_{ijk}^{(s)} - e_{ijk}^{(s-1)} | \leq 0.1$

ขั้นตอนที่ 2 การสร้างโมเดลล็อกลิเนียร์

โมเดลล็อกลิเนียร์สามารถเขียนได้ในรูปสมการคณิตศาสตร์ เพื่อประมาณค่าความถี่ที่คาดหวังในรูปผลคูณของเทอมอิทธิพลของตัวแปร เนื่องด้วยโมเดลเป็นแบบคูณจะแสดงความสัมพันธ์ที่ไม่ใช่เส้นตรง ซึ่งสามารถเปลี่ยนให้โมเดลอยู่ในรูปการบวกได้โดยใช้ลอการิทึมฐาน e (ทวีพร บุญวานิช, 2541) โมเดลล็อกลิเนียร์ที่ได้จากการคำนวณความถี่ที่คาดหวัง 4 แบบที่ได้เสนอไว้ในหัวข้อ 1.2 สามารถทำให้เป็นโมเดลเชิงเส้นได้ดังนี้

กำหนดให้ $\mu = \ln\eta, \lambda_i^A = \ln\tau_i^A, \lambda_j^B = \ln\tau_j^B \lambda_{ij}^{AB} = \ln\tau_{ij}^{AB}$

$F_{ij} = \eta$ จะได้ $\ln F_{ij} = \mu$ เขียนแทนด้วย [ ]

$F_{ij} = \eta\tau_j^B$ $\ln F_{ij} = \mu + \lambda_i^B$ [B]

$F_{ij} = \eta\tau_i^A\tau_j^B$ $\ln F_{ij} = \mu + \lambda_i^A + \lambda_j^B$ [A][B]

$F_{ij} = \eta\tau_i^A\tau_j^B\tau_{ij}^{AB}$ $\ln F_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$ [AB]

สำหรับตาราง 3 ทาง มี 3 ตัวแปร แต่ละตัวมี 2 ค่า จะได้โมเดลเชิงเส้นดังนี้

กำหนดให้ $\mu = \ln\eta, \lambda_i^A = \ln\tau_i^A, \lambda_j^B = \ln\tau_j^B, \lambda_k^C = \ln\tau_k^C$

$$\lambda_{ij}^{AB} = \ln \tau_{ij}^{AB}, \lambda_{ik}^{AC} = \ln \tau_{ik}^{AC}, \lambda_{jk}^{BC} = \ln \tau_{jk}^{BC} \text{ และ } \lambda_{ijk}^{ABC} = \ln \tau_{ijk}^{ABC}$$

$F_{ijk} = \eta$ จะได้ $\ln F_{ijk} = \mu$ เขียนแทนด้วย [ ]

$F_{ijk} = \eta\tau_i^A$ $\ln F_{ijk} = \mu + \lambda_i^A$ [A]

$F_{ijk} = \eta\tau_i^A\tau_j^B$ $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B$ [A][B]

$F_{ijk} = \eta\tau_i^A\tau_j^B\tau_k^C$ $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C$ [A][B][C]

$F_{ijk} = \eta\tau_i^A\tau_j^B\tau_k^C\tau_{ij}^{AB}$ $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$ [A][B][C][AB]

$F_{ijk} = \eta\tau_i^A\tau_j^B\tau_k^C\tau_{ij}^{AB}\tau_{ik}^{AC}$ $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$ [A][B][C][AB][AC]

$F_{ijk} = \eta\tau_i^A\tau_j^B\tau_k^C\tau_{ij}^{AB}\tau_{ik}^{AC}\tau_{jk}^{BC}$ $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$
[A][B][C][AB][AC][BC]

$F_{ijk} = \eta\tau_i^A\tau_j^B\tau_k^C\tau_{ij}^{AB}\tau_{ik}^{AC}\tau_{jk}^{BC}\tau_{ijk}^{ABC}$ $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$
[A][B][C][AB][AC][BC][ABC]

ความหมายของโมเดลสำหรับตาราง 3 ทาง

กรณีที่ 1 โมเดล $\ln F_{ijk} = \mu$ เขียนแทนด้วย [ ] คือโมเดลที่การคำนวณความถี่ที่คาดหวังโดยไม่ได้รับอิทธิพลหลักหรืออิทธิพลปฏิสัมพันธ์จากตัวแปรใด ๆ เลย ค่าความถี่ที่คาดหวังแต่ละเซลล์เท่ากันหมด

กรณีที่ 2 โมเดล $\ln F_{ijk} = \mu + \lambda_i^A$ เขียนแทนด้วย [A] คือโมเดลที่การคำนวณความถี่ที่คาดหวังโดยได้รับอิทธิพลหลักจากตัวแปร A เท่านั้น

กรณีที่ 3 โมเดล $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B$ เขียนแทนด้วย [A][B] คือโมเดลที่การคำนวณความถี่ที่คาดหวังได้รับอิทธิพลหลักจากตัวแปร A และตัวแปร B ไม่มีอิทธิพลปฏิสัมพันธ์ AB

กรณีที่ 4 โมเดล $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C$ เขียนแทนด้วย [A][B][C] คือโมเดลที่การคำนวณความถี่ที่คาดหวังได้รับอิทธิพลหลักจากตัวแปร A ตัวแปร B และตัวแปร C ตัวแปรทั้ง 3 ไม่มีอิทธิพลปฏิสัมพันธ์กัน

กรณีที่ 5 โมเดล $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$ เขียนแทนด้วย [A][B][C][AB] คือโมเดลที่การคำนวณความถี่ที่คาดหวังได้รับอิทธิพลหลักจากตัวแปร A ตัวแปร B ตัวแปร C และอิทธิพลปฏิสัมพันธ์ AB

กรณีที่ 6 โมเดล $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$ เขียนแทนด้วย [A][B][C][AB][AC] คือโมเดลที่การคำนวณความถี่ที่คาดหวังได้รับอิทธิพลหลักจากตัวแปร A ตัวแปร B ตัวแปร C และอิทธิพลปฏิสัมพันธ์ AB และอิทธิพลปฏิสัมพันธ์ AC

กรณีที่ 7 โมเดล $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$ เขียนแทนด้วย [A][B][C][AB][AC][BC] คือ โมเดลที่การคำนวณความถี่ที่คาดหวังได้รับอิทธิพลหลักจากตัวแปร A ตัวแปร B ตัวแปร C อิทธิพลปฏิสัมพันธ์ AB อิทธิพลปฏิสัมพันธ์ AC และอิทธิพลปฏิสัมพันธ์ BC

กรณีที่ 8 โมเดล $\ln F_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$ เขียนแทนด้วย [A][B][C][AB][AC][BC][ABC] คือ โมเดลที่การคำนวณความถี่ที่คาดหวังได้รับอิทธิพลหลักจากตัวแปร A ตัวแปร B ตัวแปร C อิทธิพลปฏิสัมพันธ์ AB อิทธิพลปฏิสัมพันธ์ AC อิทธิพลปฏิสัมพันธ์ BC และอิทธิพลปฏิสัมพันธ์ ABC

ขั้นตอนที่ 3 การประมาณค่าพารามิเตอร์

การประมาณค่าพารามิเตอร์ในการวิเคราะห์ล็อกลิเนียร์มีวิธีที่นิยมใช้คือวิธีความเป็นไปได้สูงสุด (maximum likelihood: ML) โดยอาศัยหลักการคำนวณทวนซ้ำ (iteration) (ทวีพร บุญวานิช, 2541; Hagenaars, 1990) ซึ่ง Deming และ Stephen (1940 อ้างใน Agresti, 1990) เป็นผู้ ริเริ่ม "*Iterative proportional fitting (IPF) algorithm*" นอกจากการคำนวณทวนซ้ำแล้วยังมีอีกวิธีคือ Newton-Raphson

จะเห็นว่าการวิเคราะห์ด้วยโมเดลล็อกลิเนียร์กรณี 2 ตัวแปรจะมีโมเดลที่แตกต่างกัน 4 โมเดล แต่ถ้าจำนวนตัวแปรเพิ่มขึ้นโมเดลก็จะยิ่งเพิ่มความซับซ้อนขึ้น เพื่อให้ได้โมเดลที่ประหยัดและเป็นตัวแทนที่ดีที่สุด จึงต้องมีการทดสอบภาวะสารูปสนิทเพื่อหาโมเดลที่ดีที่สุดเพียงหนึ่งโมเดล

ขั้นตอนที่ 4 การคัดเลือกโมเดล

การคัดเลือกโมเดลเป็นขั้นตอนที่นักวิจัยจะต้องคัดเลือกโมเดลล็อกลิเนียร์ให้เหลือโมเดลที่เหมาะสมกับการอธิบายความถี่ที่คาดหวังโดยการทดสอบความสอดคล้องระหว่างความถี่ที่คาดหวังกับข้อมูลเชิงประจักษ์หรือความถี่ที่สังเกตได้ เพื่อให้ได้โมเดลที่ดีที่สุดควรจะเป็นโมเดลที่ประหยัดและสอดคล้องกับข้อมูลอย่างเพียงพอ กระบวนการคัดเลือกมีขั้นตอนการทดสอบภาวะสารูปสนิทของโมเดล 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 4.1 การทดสอบความสอดคล้องความถี่ที่คาดหวังของโมเดลแต่ละโมเดลกับความถี่ที่สังเกตได้ โดยตั้งสมมติฐานหลักว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์โดยใช้อัตราส่วนไลค์ลิฮูดไค-สแควร์เป็นสถิติทดสอบเนื่องจากสามารถแบ่งได้หลายค่าโดยที่แต่ละค่าแสดงอิทธิพลของตัวแปรแต่ละเทอมของโมเดลนั้น (ทวีพร บุญวานิช, 2541; ประยง มิหากิตติคุณ, 2538; Knoke และ Burke, 1980) มีสูตรในการคำนวณดังนี้

$$G^2 = 2\sum\sum o_{ij} \ln\left(\frac{O_{ij}}{e_{ij}}\right)$$

Knoke และ Burke (1980) แนะนำว่าควรใช้ G^2 มากกว่า χ^2 เพราะ (1) ความถี่ที่คาดหวังประมาณค่าด้วยวิธีความเป็นไปได้สูงสุด (maximun likelihood) และ (2) การแบ่งค่า (partition) ของ G^2 มีสมรรถภาพ (powerful) ดีกว่าไค-สแควร์ องศาอิสระของ G^2 เท่ากับผลรวมของจำนวนเทอมอิทธิพลในแต่ละค่าของตัวแปรจากโมเดลอิ่มตัวที่ประมาณค่าได้อย่างอิสระลบออกด้วยผลรวมของจำนวนเทอมในแต่ละค่าของตัวแปรจากโมเดลที่ต้องการทดสอบและประมาณค่าได้อย่างอิสระ (ทวีพร บุญวานิช, 2541) (ดังตัวอย่างที่จะอธิบายถัดไป) ในการทดสอบภาวะสารูปสนิทนักวิจัยต้องการที่จะคงสมมติฐานหลักซึ่งตรงกันข้ามกับการทดสอบความเป็นอิสระของ χ^2 ที่ต้องการปฏิเสธสมมติฐานหลัก

ตารางที่ 6 ตัวอย่างการหาองศาอิสระในตาราง 2 x 4

A\B	B_1	B_2	B_3	B_4	TOTAL
A_1	O_{11}	O_{12}	O_{13}	O_{14}	$n_{1.}$
A_2	O_{21}	O_{22}	O_{23}	O_{24}	$n_{2.}$
TOTAL	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.4}$	N

ต้องการประมาณค่าองศาอิสระของโมเดลความน่าจะเป็นเท่า ($F_{ij} = \eta$) และโมเดลความน่าจะเป็นเท่าแบบมีเงื่อนไข ($F_{ij} = \eta\tau_i^A$) จากตารางตัวแปร A มี 2 ค่า ตัวแปร B มี 4 ค่า การพิจารณาจำนวนเทอมอิทธิพลที่ประมาณค่าได้อย่างอิสระทำได้ดังนี้

η คือ เทอมอิทธิพลคงที่ประมาณค่าได้เพียง 1 ค่าเท่านั้น ดังนั้นจำนวนเทอมที่ประมาณได้อย่างอิสระมีค่าเท่ากับ 1

τ_i^A คือ เทอมของอิทธิพลหลัก A ซึ่งมีเทอมอิทธิพลทั้งหมด 2 เทอม นั่นคือ τ_1^A หรือ τ_2^A แต่จะประมาณค่า τ_i^A ได้อิสระ 1 เทอมเท่านั้น

τ_j^B คือ เทอมอิทธิพลหลัก B ซึ่งมีเทอมอิทธิพลทั้งหมด 4 เทอมนั่นคือ $\tau_1^B, \tau_2^B, \tau_3^B$ และ τ_4^B และจะประมาณได้อย่างอิสระ 3 เทอมเท่านั้น

τ_{ij}^{AB} คือ เทอมอิทธิพลปฏิสัมพันธ์ AB มี 8 เทอมคือ $\tau_{11}^{AB}, \tau_{12}^{AB}, \tau_{13}^{AB}, \tau_{14}^{AB}, \tau_{21}^{AB}, \tau_{22}^{AB}$, τ_{23}^{AB} และ τ_{24}^{AB} ซึ่งอิทธิพลจาก A จะประมาณได้อย่างอิสระ 1 เทอม อิทธิพลจาก B จะประมาณได้อย่างอิสระ 3 เทอมดังนั้นจะประมาณค่าได้อย่างอิสระ 1 X 3 = 3 เทอม

จากรายละเอียดข้างต้นสามารถคำนวณผลรวมของจำนวนเทอมอิทธิพลในแต่ละค่าของตัวแปรในโมเดลไม่อิ่มตัวได้ทั้งหมด 1 + 1 + 3 + 3 = 8 เทอม และคำนวณผลรวมของจำนวนเทอมอิทธิพลในแต่ละค่าของตัวแปรในโมเดลความน่าจะเป็นเท่าที่ประมาณค่าได้อย่างอิสระเท่ากับ 1 องศาอิสระของโมเดลความน่าจะเป็นเท่า ($F_{ij} = \eta$) มีค่าเท่ากับ 8 - 1 = 7 ส่วนผลรวมของจำนวนเทอมอิทธิพลในแต่ละค่าของตัวแปรในโมเดลความน่าจะเป็นเท่าแบบมีเงื่อนไข ($F_{ij} = \eta\tau_i^A$) ที่ประมาณได้อย่างอิสระเท่ากับ 1 + 1 = 2 ดังนั้น องศาอิสระของโมเดลที่ 2 มีค่าเท่ากับ 8 – 2 = 6

ขั้นตอนที่ 4.2 ทดสอบนัยสำคัญของเทอมอิทธิพลที่เพิ่มเข้าไปในโมเดลที่ผ่านการคัดเลือกจากขั้นแรก โดยตั้งสมมติฐานหลักว่า อิทธิพลที่เพิ่มเข้าไปใหม่มีค่าเท่ากับศูนย์ แล้วหาผลต่างของค่าสถิติ G^2 ยกตัวอย่างเช่น

ตารางที่ 7 ตัวอย่างการทดสอบภาวะสารูปสนิทของโมเดลล็อกลิเนียร์ด้วยอัตราส่วนไลค์ลิฮูดไค-สแควร์และผลต่างของค่าสถิติ กรณีมี 2 ตัวแปร แต่ละตัวแปรมี 2 ค่า

Model	Model	G^2	d.f.	p	ΔG^2	d.f.	p
1	[ ]	77.53	3	.000			
2	[A]	77.50	2	.000	.030	1	.068
3	[A][B]	5.26	2	.088	72.240	0	.010
4	[A][B][AB]	5.04	1	.300	.220	1	.080

หมายเหตุ α = .05

จากตารางจะพบว่า มีโมเดลต้องการทดสอบ 4 โมเดล ได้ดำเนินการทดสอบสองขั้นตอน ขั้นตอนแรกเป็นการคัดเลือกโมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์หรือความถี่ที่คาดหวังมากที่สุดซึ่งในขั้นนี้จะได้โมเดลที่ 3 และ 4 ขั้นที่สอง พิจารณาจากค่าความแตกต่างของ G^2 และความแตกต่างขององศาอิสระ แล้วนำค่า G^2 เปรียบเทียบกับค่าสถิติไค-สแควร์ ถ้าโมเดลใดปฏิเสธสมมติฐานหลักแสดงว่าเทอมอิทธิพลมีค่าเท่ากับศูนย์ จึงจะได้ว่าโมเดลนั้นมีอิทธิพลเพียงพอที่ทำให้โมเดลมีสอดคล้องกับข้อมูลเชิงประจักษ์ จากตารางเมื่อพิจารณา ΔG^2 พบความแตกต่างอย่างมีนัยสำคัญระหว่างโมเดลที่ 2 กับ 3 ซึ่งโมเดลที่ 3 เป็นโมเดลที่ผ่านการพิจารณาในขั้นแรกมาแล้ว จึงได้ว่าโมเดลที่ 3 เป็นโมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์และประหยัดที่สุด

เนื่องจากตัวแปรในตารางที่ 7 มีเพียง 2 ตัวแปร ตารางจึงมีขนาด 2 X 2 เพื่อให้เข้าใจยิ่งขึ้นในการคัดเลือกโมเดลมากขึ้น ผู้วิจัยขอยกตัวอย่างในกรณีมี 3 ตัวแปร ดังรายละเอียด

ตารางที่ 8 ตัวอย่างการคัดเลือกโมเดล กรณีตาราง 3 ทาง ตัวแปร 3 ตัวแปร

Model	Model	G^2	d.f.	P	ΔG^2	d.f.	P
1	[ ]	137.93	7	0.00			
					6.25	1	≤ .05
2	[A]	131.68	6	0.00			
					6.03	1	≤ .05
3	[A][B]	137.71	5	0.00			
					129.58	1	≤ .05
4	[A][B][C]	8.13	4	0.04			
					123.33	1	NS
5	[A][B][C][AB]	131.46	3	0.00			
					123.55	1	NS
6	[A][B][C][AB][AC]	7.91	2	0.02			
					6.03	1	NS
7	[A][B][C][AB][AC][BC]	1.88	1	0.39			
					1.18	1	≤ .05
8	[A][B][C][AC][BC][ABC]	0.70	0	0.40			

หมายเหตุ $\alpha = .05$

จากตารางจะพบว่าในขั้นที่หนึ่งมีโมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์ ซึ่งคงสมมติฐานหลักที่ระดับนัยสำคัญ .05 คือโมเดลที่ 7 และ 8 จึงนำมาทดสอบในขั้นที่สอง ปรากฏว่าโมเดลที่ 8 เมื่อเพิ่มอิทธิพลปฏิสัมพันธ์ ABC พบว่าปฏิเสธสมมติฐานหลัก แสดงว่าอิทธิพลปฏิสัมพันธ์ ABC ที่เพิ่มเข้าไป มีนัยสำคัญทางสถิติเพียงพอที่ทำให้โมเดลที่ 8 มีความสอดคล้องกับมูลเชิงประจักษ์ และเมื่อพิจารณาโมเดลที่ 7 ซึ่งเพิ่มอิทธิพลปฏิสัมพันธ์ BC พิจารณาร่วมกับอิทธิพลปฏิสัมพันธ์ AB และ AC พบว่าคงสมมติฐานหลัก จึงสรุปได้ว่าโมเดลที่ 8 เป็นโมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์และประหยัดที่สุด

ขั้นตอนที่ 5 การตรวจสอบความเหมาะสมของโมเดล

เมื่อคัดเลือกโมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์และประหยัดได้แล้ว ขั้นต่อไปคือการตรวจสอบเพื่อให้มั่นใจว่าโมเดลที่คัดเลือกมาเป็นตัวแทนที่ดีที่ใช้ทำนายความถี่ที่คาดหวังและนำไปสู่การสรุปที่ถูกต้องโดยตรวจสอบความเหมาะสมของโมเดลทุกๆ เซลล์ด้วยค่าเศษเหลือมาตรฐาน (standardized residual)

จากสูตร
$$e_i = \frac{o_{ij} - e_{ij}}{\sqrt{e_{ij}}}$$

ซึ่งกำลังสองของเศษเหลือมาตรฐานจะมีค่าเท่ากับไค-สแควร์ $\sum e_i^2 = \chi^2$

จากนั้นจึงมีการปรับเศษเหลือ (adjusted residual)

ด้วยสูตร
$$r_{ij} = \frac{o_{ij} - e_{ij}}{\left[e_{ij}(1 - p_{i.})(1 - p_{.j})\right]^{\frac{1}{2}}}$$

(หมายเหตุ สูตรนี้ใช้กับตาราง 2 x 2 เท่านั้น)

เมื่อ p_{ij} คือ โอกาสที่ความถี่ที่คาดหวังจะถูกจำแนกลงในเซลล์ที่ ij

e_{ij} คือ $np_{i.}p_{.j}$

สำหรับตาราง 3 ทาง (three-way tables) สูตรในการปรับเศษเหลือจะแตกต่างกันตามโมเดล ยกตัวอย่างเช่น

Model	Adjusted residual
[A][B][C]	$r_{ijk} = \frac{o_{ijk} - e_{ijk}}{\left[e_{ij}(1 - p_{i..}p_{.j.} - p_{i..}p_{..k} - p_{.j.}p_{..k} + 2p_{i..}p_{.j.}p_{..k})\right]^{\frac{1}{2}}}$
[ ][A][B][C][AC]	$r_{ijk} = \frac{o_{ijk} - e_{ijk}}{\left[e_{ij}(1 - p_{ij.})(1 - p_{..k})\right]^{\frac{1}{2}}}$
[ ][A][B][C][AB][AC]	$r_{ijk} = \frac{o_{ijk} - e_{ijk}}{\left[e_{ij}(1 - \frac{o_{ij.}}{o_{i..}})(1 - \frac{o_{i.k}}{o_{i..}})\right]^{\frac{1}{2}}}$

ในการศึกษาความสัมพันธ์แบบสมมาตรเมื่อผ่านการคัดเลือกโมเดลและผ่านการตรวจสอบความเหมาะสมของโมเดลและสรุปได้ว่าโมเดลมีความสัมพันธ์กัน นักวิจัยต้องศึกษาว่าความสัมพันธ์มีขนาดและทิศทางอย่างไร โดยใช้วิธีการวัดความสัมพันธ์ (measure of association) เช่น สถิติวีของคาร์เมอร์ (Cramer's V), สถิติ Q ของยูล (Yule's Q) หรืออัตราส่วนแต้มต่อตามแต่กรณี

ของสถิติเหล่านั้นมีเงื่อนไข แต่ถ้านักวิจัยศึกษาความสัมพันธ์อสมมาตรและพบว่าตัวแปรอิสระตัวใดส่งผลต่อตัวแปรตาม นักวิจัยจะต้องประมาณค่าพารามิเตอร์ลอการิทึมของอัตราส่วนแต้มต่อ (log odds ratio) หากนักวิจัยศึกษาความสัมพันธ์สมมาตรจะต้องประมาณค่าพารามิเตอร์ด้วยอัตราส่วนแต้มต่อ (odds ratio) ซึ่งจะอธิบายได้ในขั้นตอนที่ 6 ต่อไป

ขั้นตอนที่ 6 การประมาณค่าพารามิเตอร์อัตราส่วนแต้มต่อ

แต้มต่อ คือ สัดส่วนของความถี่ที่นักวิจัยสนใจโอกาสในการเกิดเหตุการณ์สองเหตุการณ์ดังรายละเอียดต่อไปนี้

ตารางที่ 9 ตัวอย่างการคำนวณอัตราส่วนแต้มต่อในตาราง 3 ทาง

A	B	C		TOTAL
		C_1	C_2	
A_1	B_1	15	9	24
	B_2	26	17	43
A_2	B_1	13	22	35
	B_2	24	23	47
TOTAL		78	71	149

ส่วนอัตราส่วนแต้มต่อคือ สัดส่วนของแต้มต่อนั่นเอง ในที่นี้ก็จะได้

พิจารณาในตัวแปร A_1 $$\theta_1 = \frac{o_{11}o_{22}}{o_{12}o_{21}} = \frac{(15)(17)}{(9)(26)} = 1.09$$

พิจารณาในตัวแปร A_2 $$\theta_2 = \frac{o_{11}o_{22}}{o_{12}o_{21}} = \frac{(13)(23)}{(22)(24)} = 0.57$$

เมื่อได้ค่าอัตราส่วนแต้มต่อมาแล้วจะนำมาทดสอบนัยสำคัญของเทอมอิทธิพล ถ้าเป็นตาราง 3 ทางการทดสอบนัยสำคัญจะพิจารณาจาก

$$z = \frac{\ln\theta_1 - \ln\theta_2}{\sqrt{s_{\theta_1}^{\ 2} + s_{\theta_2}^{\ 2}}}$$

เมื่อ $s_{\theta_1}^{\ 2}$ คือ ความแปรปรวนของ $\ln\theta_1$ ซึ่งหาได้จาก $1/o_{11} + 1/o_{12} + 1/o_{21} + 1/o_{22}$
เมื่อ Z เป็นการแจกแจงปกติที่มีค่าเฉลี่ยเลขคณิตเท่ากับ 0 ความแปรปรวนเท่ากับ 1 ถ้าเป็นตัวอย่างดังตารางที่ 9 จะทดสอบนัยสำคัญได้ดังนี้

$$z = \frac{\ln\theta_1 - \ln\theta_2}{\sqrt{s_{\theta_1}^{\ 2} + s_{\theta_2}^{\ 2}}}$$

$$s_{\theta_1}^{\ 2} = 1/o_{11} + 1/o_{12} + 1/o_{21} + 1/o_{22}$$

$$= 1/15 + 1/9 + 1/26 + 1/17$$

$$= 0.07 + 0.1 + 0.04 + 0.06 = 0.27$$

$$s_{\theta_2}^{\ 2} = 1/o_{11} + 1/o_{12} + 1/o_{21} + 1/o_{22}$$

$$= 1/13 + 1/22 + 1/24 + 1/23$$

$$= 0.08 + 0.05 + 0.04 + 0.04 = 0.21$$

จะได้

$$z = \frac{\ln\theta_1 - \ln\theta_2}{\sqrt{s_{\theta_1}^{\ 2} + s_{\theta_2}^{\ 2}}} = 0.94$$

นำค่า Z ที่คำนวณได้ไปเปรียบเทียบกับค่า z จากตาราง ($\pm$ 1.96) จึงสรุปได้ว่า ยอมรับสมมติฐานหลัก สรุปได้ว่าไม่มีนัยสำคัญทางสถิติของปฏิสัมพันธ์ 3 ทาง (no significant three - way interaction)

ค่าลอการิทึมของอัตราส่วนแต้มต่อ คือ สัดส่วนค่าลอการิทึมผลต่างของโอกาสในการเกิดเหตุการณ์สองเหตุการณ์ที่ใช้อธิบายความแตกต่างของความถี่ 2 ระดับของตัวแปรตาม ในที่นี้เป็นค่าที่บ่งถึงค่าของอิทธิพลของตัวแปรหนึ่งที่มีต่ออีกตัวแปรหนึ่ง เช่น มีโมเดล $F_{ij} = \eta\tau_i^A\tau_j^B\tau_{ij}^{AB}$ เมื่อเป็นโมเดลโลจิทล็อกลิเนียร์คือโมเดลที่พิจารณาความสัมพันธ์โดยมีตัวแปรตาม 1 ตัว โมเดลข้างต้นจะพัฒนามาเป็นโมเดลโลจิทล็อกลิเนียร์ ดังนี้ $F_{ij} = \tau_i^A\tau_{ij}^{AB}$ เมื่อเป็นอัตราส่วนแต้มต่อ จะได้

$$\frac{F_{i1}}{F_{i2}} = \frac{\tau_1^A\tau_{i1}^{AB}}{\tau_2^A\tau_{i2}^{AB}}$$

$$\frac{F_{i1}}{F_{i2}} = (\tau^A)^2(\tau^{AB})^2$$

$$\ln(\frac{F_{i1}}{F_{i2}}) = 2\ln(\tau^A) + 2\ln(\tau^{AB})$$

แทน $\ln\tau^a$ ด้วย λ จะได้

$$\ln(\frac{F_{i1}}{F_{i2}}) = 2\lambda^A + 2\lambda^{AB}$$

เขียนรูปใหม่ได้เป็น

$$\phi_{ij}^{AB} = \beta_i^{\ A} + \beta_{ij}^{AB}$$

เมื่อ ϕ_{ij}^{AB} คือ ลอการิทึมของแต้มต่อที่คาดหวัง

β_i^{A} คือ $2\lambda^{A}$ (ค่าลอการิทึมของอิทธิพลหลัก A)

β_{ij}^{AB} คือ $2\lambda^{AB}$ (ค่าลอการิทึมของอิทธิพลปฏิสัมพันธ์ AB)

ข้อจำกัดของโมเดลล็อกลิเนียร์

แม้ว่าการวิเคราะห์ด้วยโมเดลล็อกลิเนียร์จะมีประสิทธิภาพดังที่กล่าวมาแล้ว แต่การวิเคราะห์ด้วยโมเดลล็อกลิเนียร์ยังมีข้อจำกัดหลายประการดังนี้ (1) หากจำนวนตัวแปรในการวิเคราะห์มีมากจะทำให้การแปรผลทำได้ยาก (Jeansonne, 2001; สุวิมล ว่องวาณิช, 2535) (2) ความถี่ของแต่ละเซลล์ต้องเป็นอิสระจากเซลล์อื่นๆ (3) ความถี่ของแต่ละเซลล์ในโมเดลล็อกลิเนียร์จะต้องไม่ต่ำกว่า 5 เช่น ในตาราง 2 X 2 X 3 ความถี่ที่ต้องการทั้งหมดคือ 60 ด้วยเหตุผลนี้ทำให้นักวิจัยต้องเพิ่มขนาดกลุ่มตัวอย่างหรือลดจำนวนตัวแปร (4) ความถี่ที่คาดหวังต้องมากกว่า 1 และมีเซลล์ที่มีความถี่ต่ำกว่า 5 น้อยกว่า 20 เปอร์เซ็นต์ ถ้าไม่เป็นไปตามเงื่อนไขดังกล่าวจะทำให้อำนาจการทดสอบไม่เป็นที่น่าพอใจ และหากไม่เป็นไปตามเงื่อนไขดังกล่าว นักวิจัยสามารถเลือกทางใดทางหนึ่งจากข้อเสนอแนะต่อไปนี้ ทางเลือกแรกยอมรับอำนาจการทดสอบที่เกิดขึ้นจากความถี่ที่คาดหวังมีค่าน้อยๆ ทางเลือกที่สองรวมตัวแปรกลุ่ม 2 กลุ่มเข้าด้วยกันเพื่อให้เป็นตัวแปรใหม่ ซึ่งการกระทำตามเงื่อนไขนี้จะทำให้อำนาจการทดสอบลดลง ทางเลือกที่สามลดจำนวนตัวแปรเพื่อลดจำนวนเซลล์แต่ควรระวังว่าไม่ควรลดตัวแปรที่มีความสัมพันธ์กับตัวแปรอื่น และทางเลือกสุดท้ายบวกด้วยค่าคงที่ในแต่ละเซลล์ แต่ไม่แนะนำให้ใช้วิธีนี้เพราะอำนาจการทดสอบจะลดลง (Jeansonne, 2001)

<u>ตอนที่ 2 อำนาจการทดสอบ</u>

อำนาจการทดสอบ หมายถึง โอกาส (probability) ที่จะปฏิเสธสมมติฐานหลักที่ผิด (Glass และ Hopkins, 1996) เพื่อที่จะได้มั่นใจได้ว่านักวิจัยกำลังจะสรุปในสิ่งที่ถูกต้อง อำนาจการทดสอบจะขึ้นอยู่กับ (1) ระดับนัยสำคัญ (significance criterion) ถ้าระดับนัยสำคัญต่ำอำนาจการทดสอบก็จะต่ำเช่นกัน (2) ขนาดกลุ่มตัวอย่าง (sample size: n) ถ้าจำนวนกลุ่มตัวอย่างมากอำนาจการทดสอบก็จะยิ่งมากตามไปด้วย และ (3) ขนาดอิทธิพล (effect size) อำนาจการทดสอบจะสูงถ้าขนาดอิทธิพลสูง (Cohen, 1988; Murphy และ Myors, 1988)

ประเภทของการวิเคราะห์อำนาจการทดสอบ (Cohen, 1965 อ้างใน Cohen, 1988)

ถ้าจะกล่าวถึงอำนาจการทดสอบต้องกล่าวถึงค่าพารามิเตอร์อีก 3 ตัวคือ ระดับนัยสำคัญ (α) ขนาดกลุ่มตัวอย่าง (n) และขนาดอิทธิพล (effect size) พารามิเตอร์ทั้ง 4 ตัวมีความสัมพันธ์เกี่ยวข้องกัน ถ้ากำหนดให้อำนาจการทดสอบ ระดับนัยสำคัญและขนาดกลุ่มตัวอย่างคงที่ การหาค่าของพารามิเตอร์ตัวใดตัวหนึ่งจะต้องขึ้นอยู่กับค่าพารามิเตอร์ 3 ตัวที่เหลือ การวิเคราะห์อำนาจการทดสอบมี 4 ประเภทคือ

1. เมื่อกำหนดระดับนัยสำคัญ ขนาดกลุ่มตัวอย่างและขนาดอิทธิพลมีค่าคงที่ อำนาจการทดสอบควรจะเป็นเท่าใด
2. เมื่อกำหนดระดับนัยสำคัญ ขนาดอิทธิพลและอำนาจการทดสอบคงที่ ขนาดกลุ่มตัวอย่างควรจะเป็นเท่าใด
3. เมื่อกำหนดระดับนัยสำคัญ ขนาดกลุ่มตัวอย่างและอำนาจการทดสอบคงที่ ขนาดอิทธิพลควรจะเป็นเท่าใด
4. เมื่อกำหนดอำนาจการทดสอบ ขนาดกลุ่มตัวอย่างและขนาดอิทธิพลคงที่ ระดับของนัยสำคัญควรจะเป็นเท่าใด

Cohen (1988) กล่าวว่าขนาดอิทธิพลเป็นส่วนสำคัญในการตัดสินใจเพื่อหาอำนาจการทดสอบและขนาดกลุ่มตัวอย่างที่เหมาะสมจึงได้มีดัชนีขนาดอิทธิพล (effect size index) เพื่อบอกความแตกต่างของสัดส่วนของประชากร 2 กลุ่ม เช่น ในแต่ละเซลล์จะมีสัดส่วนของประชากร 2 กลุ่ม สัดส่วนแรกได้จากสมมติฐานหลักและอีกส่วนได้จากสมมติฐานทางเลือก ดัชนีขนาดอิทธิพลจึงเป็นการวัดความแตกต่าง (discrepency) ระหว่างสัดส่วนของเซลล์ สามารถหาได้จาก

$$W = \sqrt{\sum_{i=1}^{m} \frac{(P_{1i} - P_{0i})^2}{P_{0i}}}$$

เมื่อ P_{0i} คือ สัดส่วนในเซลล์ที่ i ในสมมติฐานหลัก

P_{1i} คือ สัดส่วนในเซลล์ที่ i ในสมมติฐานทางเลือก และเป็นผลสะท้อนจากอิทธิพล (reflect the effect) ของเซลล์

m คือ จำนวนของเซลล์

ค่าของ W ในตารางการณ์จรจะอยู่ระหว่าง 0 ถึง $\sqrt{m-1}$, df = m –1

นอกจากนี้ Cohen (1988) ยังได้แนะนำการพิจารณาขนาดอิทธิพล ซึ่งแบ่งได้เป็น 3 ขนาด คือ ขนาดอิทธิพลน้อย (W = .10) ขนาดอิทธิพลปานกลาง (W = .30) และขนาดอิทธิพลมาก (W = .50) และยังได้เสนอแนวทางในการแบ่งสัดส่วนขอบเขตเพื่อทดสอบภาวะสารูปสนิทในตารางการณ์จร ดังนี้

กรณีขนาดอิทธิพลน้อย W = .10 เมื่อ m คือจำนวนเซลล์

m = 2 H_0: .50 .50

H_1: .45 .55

m = 3 H_0: .333 .333 .333

H_1: .293 .333 .374

m = 4 H_0: .250 .250 .250 .250

H_1: .216 .239 .261 .284

m = 5 H_0: .200 .200 .200 .200 .200

H_1: .172 .186 .200 .214 .228

m = 10 H_0: .100 .100 .100 .100 .100 .100 .100 .100 .100 .100

H_1: .084 .088 .091 .095 .098 .102 .105 ..109 .112 .116

กรณีขนาดอิทธิพลปานกลาง W = .30 เมื่อ m คือจำนวนเซลล์

m = 2 H_0: .50 .50

H_1: .35 .65

m = 3 H_0: .333 .333 .333

H_1: .211 .333 .456

m = 4 H_0: .250 .250 .250 .250

H_1: .149 .216 .284 .351

m = 5 H_0: .200 .200 .200 .200 .200

H_1: .115 .158 .200 .242 .285

m = 10 H_0: .100 .100 .100 .100 .100 .100 .100 .100 .100 .100

H_1: .053 .063 .074 .084 .095 .105 .116 .126 .137 .147

กรณีขนาดอิทธิพลมาก W = .50 เมื่อ m คือจำนวนเซลล์

m = 2 H_0: .50 .50

H_1: .25 .75

m = 3 H_0: .333 .333 .333

H_1: .129 .333 .537

m = 4 H_0: .250 .250 .250 .250

H_1: .082 .194 .306 .418

m = 5 H_0: .200 .200 .200 .200 .200

H_1: .059 .129 .200 .271 .341

m = 10 H_0: .100 .100 .100 .100 .100 .100 .100 .100 .100 .100

H_1: .022 .039 .056 .074 .091 .109 .126 .143 .161 .178

หมายเหตุ H_1 นักวิจัยสามารถกำหนดเองได้ตามความเหมาะสม

ตอนที่ 3 งานวิจัยที่เกี่ยวข้องกับอัตราส่วนไลค์ลิฮูดไค-สแควร์ โมเดลล็อกลิเนียร์ และอำนาจการทดสอบ

Berry และ Mielke (1988) ได้ทำการเปรียบเทียบ asymptotic ของไค-สแควร์และอัตราส่วนไลค์ลิฮูดที่เป็น nonasymptotic สำหรับตาราง $r \times c$ ที่มีจำนวนความถี่น้อยๆ ด้วยการจำลองสถานการณ์เป็น 5 กรณีคือ (1) nonasymptotic chi- square test (2) asymptotic chi-square test with chi-square distribution (3) asymptotic chi-square test with normal distribution (4) asymptotic likelihood-ratio with chi-square distributionและ (5) asymptotic likelihood-ratio with normal distribution สำหรับทำการเปรียบเทียบในตาราง 2 X 2, 2 X 4 และ 3 X 4 โดยแบ่งค่าสัดส่วนส่วนริม (marginal) เป็น 3 กรณี คือ ค่าสัดส่วนส่วนริมเท่ากัน (equal marginal) 1 กรณีและค่าสัดส่วนส่วนริมไม่เท่ากัน (unequal marginal) 2 กรณี ในการทดสอบความเป็นอิสระ (test of independence) และการทดสอบความเป็นเอกพันธ์ (test of homogeneity) พบว่า asymptotic chi-square test with normal distribution และ asymptotic likelihood-ratio with normal distribution ไม่เหมาะจะนำไปใช้ เพราะการประมาณด้วยการแจกแจงปกติเหมาะสำหรับตารางการณ์จรที่มีองศาอิสระมากๆ ส่วน asymptotic likelihood-ratio with chi-square distribution ให้ผลดีกว่าการทดสอบด้วย 2 วิธีข้างต้นแต่ก็ยังไม่เหมาะจะนำมาใช้สำหรับ asymptotic chi-square test with chi-square distribution ให้ผลดีกว่าที่กล่าวมาแล้วทั้งหมด ส่วนสถิติทดสอบที่ดีที่สุดในที่นี้คือ nonasymptotic chi-square test

Parshall และ Kromrey (1996) ได้ทำการเปรียบเทียบอำนาจการทดสอบและการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของ Pearson chi-square test Continuity Correction test Likelihood Ratio chi-quare testและ Fisher Exact test ได้ทำการศึกษาโดยเทคนิคมอนติคาร์โลในการทดสอบความเป็นอิสระของตารางการณ์จรโดยที่กลุ่มตัวอย่างมีขนาดเล็กสุ่มกลุ่มตัวอย่างในตาราง 2 X 2 โดยแบ่งขอบเขตที่ศึกษาออกเป็น 60:40 และ 90:10 ส่วนตาราง 3 X 3 แบ่ง

ขอบเขตที่ศึกษาออกเป็น 30:30:40 และ 45:45:10 และพิจารณาค่าขนาดอิทธิพล (effect size) เป็น 0.10, 0.30 และ 0.50 ผลการวิจัยพบว่าสถิติทดสอบ Continuity Correction สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีที่สุด รองลงมาคือ Fisher Exact ส่วน Likelihood Ratio chi-square ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อยที่สุด นอกจากนี้ยังพบว่า Pearson chi-square เหมาะกับกลุ่มตัวอย่างที่มีขนาดเล็กและให้ใช้ Pearson chi-square แทน Likelihood Ratio chi-square ด้วยเหตุผล 2 ประการคือ Pearson chi-square สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีกว่า Likelihood Ratio chi-square ภายใต้กลุ่มตัวอย่างขนาดเล็กและ Pearson chi-square มีสมรรถภาพ (powerful) ดีกว่า Likelihood Ratio chi-square ภายใต้กลุ่มตัวอย่างขนาดเล็ก

ประยง มหากิตติคุณ (2538) ทำการทดสอบภาวะสารูปสนิทของโมเดลล็อกลิเนียร์ในตารางการณ์จรพหุมิติที่มีข้อมูลเบาบาง โดยพิจารณาจากความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบพบว่าเมื่อข้อมูลเบาบางตัวสถิติอัตราส่วนไลค์ลิฮูดไค-สแควร์สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี แต่อำนาจการทดสอบต่ำมาก และเมื่อค่าความถี่ที่คาดหวังมากกว่า 5 อำนาจการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วจนกระทั่งเข้าใกล้ 1 เมื่อเปรียบเทียบอำนาจการทดสอบพบว่าอัตราส่วนไลค์ลิฮูดไค-สแควร์มีอำนาจการทดสอบสูงกว่าตัวอื่นๆ กรณีที่ข้อมูลเบาบางอำนาจการทดสอบของอัตราส่วนไลค์ลิฮูดไค-สแควร์จะต่ำกว่าสถิติทดสอบไค-สแควร์เล็กน้อย และสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีกว่าอัตราส่วนไลค์ลิฮูดไค-สแควร์ ดังนั้นเมื่อข้อมูลบางเบาสถิติทดสอบไค-สแควร์จึงมีสมรรถนะดีกว่าอัตราส่วนไลค์ลิฮูดไค-สแควร์

วิชุดา ชุนชาติประเสริฐ (2531) ทำการวิเคราะห์ข้อมูลเชิงคุณภาพแบบแบ่งชั้นภูมิในแบบแผนการสุ่มผสมสำหรับการทดลองทางคลินิค โดยใช้การวิเคราะห์ตัวแบบลอการิทึมเชิงเส้นตรงและทดสอบตัวแบบด้วยตัวสถิติไลค์ลิฮูด G^2 โดยข้อมูลอยู่ในรูปตารางการณ์จรขนาด 2 X 2 X L (L = 2, 3 และ 4) ตัวแปรตัวที่หนึ่งคือ การตอบสนองของผู้ป่วย (ตัวแปรตาม Y) มี 2 ค่าคือ อาการเป็นที่น่าพอใจและอาการไม่เป็นที่น่าพอใจ โดยศึกษาความน่าจะเป็นของอาการพอใจ (P_{ijl}) 4 ระดับคือ 0.2, 0.4, 0.6 และ 0.8 ตัวแปรตัวที่สองคือ ทรีทเม้นต์ (ตัวแปรอิสระ A) มี 2 ค่า ตัวแปรตัวที่ 3 คือ ปัจจัยที่ไม่ใช่ทรีทเม้นต์แต่มีอิทธิพลต่อการทดลอง (ตัวแปรอิสระ B) หาอำนาจการทดสอบภายใต้ขนาดกลุ่มตัวอย่างเท่ากับ 24, 48, 72, 96, 120, 144, 192 และ 240 ระดับนัยสำคัญเท่ากับ .01 และ .05 จำลองข้อมูลด้วยเทคนิคมอนติคาร์โลและกระทำซ้ำกัน 1,000 ครั้ง ผลการศึกษาพบว่า เมื่อขนาดกลุ่มตัวอย่างสูงขึ้น อำนาจการทดสอบจะสูงขึ้น และถ้ามี P_{ijl} ชุดเดียวกันแม้ว่าอยู่ต่างชั้นภูมิอำนาจการทดสอบจะมีค่าใกล้เคียงกัน

สุวิมล มั่นมงคล (2526) ศึกษาเปรียบเทียบการทดสอบความเป็นอิสระของสถิติทดสอบโมเดลล็อกลิเนียร์กับสถิติทดสอบไค-สแควร์ โดยใช้เทคนิคมอนติคาร์โล แบ่งตารางการณ์จรเป็น 2 ทาง 3 ทาง และ 4 ทาง จำนวนความถี่แต่ละเซลล์มี 4 ลักษณะ คือ เท่ากัน ใกล้เคียงกันทุกๆ เซลล์ แตกต่างกันน้อยในแต่ละเซลล์ และแตกต่างกันมากในแต่ละเซลล์ กำหนดระดับนัยสำคัญเท่ากับ 0.2, 0.1, 0.05 และ 0.005 ผลการวิจัยสรุปว่า ณ ระดับนัยสำคัญเท่ากับ 0.2 สามารถใช้สถิติทดสอบทั้ง 2 ตัวแทนกันได้ไม่ว่าความถี่ในแต่ละเซลล์จะมากน้อยเพียงใด แต่จะใช้โมเดลล็อกลิเนียร์ไม่ได้เมื่อผลรวมของความถี่ในระดับใดระดับหนึ่งของตัวแปรใดตัวแปรหนึ่งเท่ากับ 0 แต่ถ้าผู้วิจัยต้องการทราบอิทธิพลหลักหรืออิทธิพลปฏิสัมพันธ์ของตัวแปร สถิติทดสอบไค-สแควร์ไม่สามารถให้คำตอบได้ แต่สถิติทดสอบโมเดลล็อกลิเนียร์สามารถประเมินอิทธิพลเหล่านั้นได้ แต่ถ้าต้องการหาความสัมพันธ์ระหว่างตัวแปรต่างๆ ควรใช้สถิติไค-สแควร์

ทวีพร บุญวานิช (2541) ศึกษาเรื่อง การประยุกต์โมเดลล็อกลิเนียร์ในการวิเคราะห์สาเหตุเพื่อการศึกษาปัจจัยที่ส่งผลต่อระยะเวลาที่ใช้ในการศึกษาของมหาบัณฑิตทางสังคมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ตัวแปรต้นเป็นนามบัญญัติ 6 ตัวแปร คือ อายุ ลักษณะการลาศึกษา คุณลักษณะอาจารย์ที่ปรึกษาวิทยานิพนธ์ การบริการของหน่วยงาน สมรรถภาพการทำวิทยานิพนธ์ และปัญหาการทำวิทยานิพนธ์ ตัวแปรตามคือ ระยะเวลาที่ใช้ในการศึกษาเป็นตัวแปรนามบัญญัติ กลุ่มตัวอย่างคือ มหาบัณฑิตทางสังคมศาสตร์ ปีการศึกษา 2535 จำนวน 346 คน และปีการศึกษา 2536 จำนวน 403 คน เก็บรวบรวมโดยส่งแบบสอบถามทางไปรษณีย์ วิเคราะห์ข้อมูลด้วยค่าสถิติเบื้องต้น การวิเคราะห์ล็อกลิเนียร์แบบอสมมาตร (โดยประยุกต์เข้ากับการวิเคราะห์เชิงสาเหตุ ด้วยเทคนิคการวิเคราะห์เส้นทาง) ผลการวิจัยสรุปว่า ปีการศึกษา 2535 ปัจจัยที่ส่งผลต่อระยะเวลาที่ใช้ในการศึกษา มีนัยสำคัญทางสถิติเฉพาะอิทธิพลทางตรงเท่านั้น ได้แก่ อิทธิพลหลักของลักษณะการลาศึกษา และปฏิสัมพันธ์อันดับ 1 ระหว่างอายุกับการบริการของหน่วยงาน ระหว่างลักษณะการลาศึกษากับปัญหาการทำวิทยานิพนธ์ และระหว่างคุณลักษณะของอาจารย์ที่ปรึกษากับปัญหาการทำวิทยานิพนธ์ ส่วนปีการศึกษา 2536 ปัจจัยที่ส่งผลต่อระยะเวลาที่ใช้ในการศึกษามีทั้งทางตรงและทางอ้อม ได้แก่ อิทธิพลหลักของลักษณะการลาศึกษา คุณลักษณะของอาจารย์ที่ปรึกษาวิทยานิพนธ์ และการบริการของหน่วยงาน และปฏิสัมพันธ์อันดับ 1 ระหว่างการบริการของหน่วยงานกับปัญหาการทำวิทยานิพนธ์ ระหว่างอายุกับลักษณะการลาศึกษา และระหว่างอายุกับคุณลักษณะของอาจารย์ที่ปรึกษาวิทยานิพนธ์

เบ็ญจมาศ แสงอนุเคราะห์ (2541) ศึกษาเรื่อง การพัฒนาโมเดลความคาดหวังในการศึกษาต่อของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนขยายโอกาสทางการศึกษาขั้นพื้นฐาน เขตการศึกษา 6 : การวิเคราะห์ล็อกลิเนียร์มาตรอันดับ (ordinal loglinear model) มีวัตถุประสงค์เพื่อ

ศึกษาความสัมพันธ์ของปัจจัยที่ส่งผลต่อความคาดหวังในการศึกษาต่อของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนขยายโอกาสทางการศึกษาขั้นพื้นฐาน เขตการศึกษา 6 เพื่อศึกษาอิทธิพลหลักและอิทธิพลปฏิสัมพันธ์ของปัจจัยที่ส่งผลต่อความคาดหวังในการศึกษาต่อของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนขยายโอกาสทางการศึกษาขั้นพื้นฐาน เขตการศึกษา 6 และเพื่อพัฒนาโมเดลล็อกลิเนียร์สำหรับตัวแปรมาตรอันดับของความคาดหวังในการศึกษาต่อของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนขยายโอกาสทางการศึกษาขั้นพื้นฐาน เขตการศึกษา 6 กลุ่มตัวอย่างมาจากนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนขยายโอกาสทางการศึกษาขั้นพื้นฐาน จำนวน 500 คน ตัวแปรตามคือ ความคาดหวังในการศึกษาต่อของนักเรียน ตัวแปรอิสระคือ การศึกษาของผู้ปกครอง อาชีพผู้ปกครอง รายได้ผู้ปกครอง จำนวนบุตรในครอบครัวของนักเรียน และผลสัมฤทธิ์ทางการเรียนของนักเรียน วิเคราะห์ข้อมูลด้วยค่าสถิติเชิงบรรยาย, Kendall's Tau, โมเดล ล็อกลิเนียร์ทั่วไป เพื่อหาความสัมพันธ์ระหว่างอิทธิพลหลักของตัวแปรอิสระกับความคาดหวังในการศึกษาต่อ นอกจากนี้ยังใช้โมเดลโลจิทล็อกลิเนียร์ เพื่อประมาณค่าพารามิเตอร์ลอการิทึมแต้มต่อที่แสดงอิทธิพลของตัวแปรอิสระที่มีผลต่อความคาดหวังในการศึกษาต่อ ผลพบว่า ปัจจัยที่มีความสัมพันธ์กับความคาดหวังในการศึกษาต่อของนักเรียน อย่างมีนัยสำคัญ ได้แก่ ผลสัมฤทธิ์ทางการเรียน อาชีพผู้ปกครอง รายได้ของผู้ปกครอง และการศึกษาของผู้ปกครอง ส่วนปัจจัยที่ส่งผลต่อความคาดหวังในการศึกษาต่อของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนขยายโอกาสทางการศึกษาขั้นพื้นฐาน อย่างมีนัยสำคัญ มีเฉพาะอิทธิพลหลัก ได้แก่ อิทธิพลหลักของผลสัมฤทธิ์ทางการเรียน อิทธิพลหลักของการศึกษาของผู้ปกครอง และอิทธิพลหลักของอาชีพผู้ปกครอง

จะเห็นว่าอัตราส่วนไลค์ลิฮูดไค-สแควร์จะมีอำนาจการทดสอบสูงกว่าไค-สแควร์เมื่อค่าความถี่ที่คาดหวังมากกว่า 5 แต่เมื่อน้อยกว่า 5 สถิติทดสอบไค-สแควร์จะมีสมรรถนะดีกว่าอัตราส่วนไลค์ลิฮูดไค-สแควร์ และยังพบว่าโมเดลล็อกลิเนียร์เหมาะกับการวิเคราะห์ข้อมูลที่ต้องการศึกษาอิทธิพลหลักและอิทธิพลปฏิสัมพันธ์ และเมื่อพิจารณาอำนาจการทดสอบ จากงานวิจัยที่เกี่ยวข้องพบว่าอำนาจการทดสอบจะสูงขึ้นเมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้น