บทที่ 2

ระเบียบวิธีที่ใช้ในการวิจัย และผลงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ต้องการศึกษาเกี่ยวกับการวิเคราะห์ข้อมูลเชิงคุณภาพแบบแบ่งชั้นภูมิในแผนแบบสุ่มผสมสำหรับการทดลองทางคลินิค (Qualitative Analysis of Stratification in Composite Randomization Designs for Clinical Trials) และวิเคราะห์ข้อมูลโดยใช้ตัวแบบลอกการิทึมเชิงเส้นตรง (Log-Linear Model) พร้อมทั้งทดสอบตัวแบบด้วยตัวสถิติไลลิฮูด G^2

โดยทั่วไปการทดลองทางคลินิคจะเริ่มตั้งแต่ศึกษาคุณสมบัติของยาด้วยการทดลองในสัตว์ทดลองก่อน จากนั้นนำยาดังกล่าวมาทดลองใช้กับอาสาสมัครที่มีสุขภาพแข็งแรง และขั้นต่อไปคือศึกษาทดลองในคลินิค ซึ่งขั้นตอนการทดลองทางคลินิคสรุปได้ดังนี้

ก. ทดลองขั้นแรก (Initial Trials) ทดลองในผู้ป่วย 2-3 รายเท่านั้น โดยมีวัตถุประสงค์ เพื่อดูความปลอดภัยของยา หาขนาดของยาที่ไม่เป็นอันตราย และประเมินผลการรักษา

ข. การทดลองเพื่อความแน่ใจ (Comfirmation Trials) เป็นการทดลองชนิดเปรียบเทียบ เพื่อดูผลของการรักษาในผู้ป่วยจำนวนมากขึ้น

การทดลองทางคลินิคนั้น ได้นำเอาสถิติเข้ามาใช้ในขั้นตอนต่างๆ ดังนี้

ก. การออกแบบการทดลอง

ข. วิธีขจัดความคลาดเคลื่อน หรืออคติ (Bias) ที่อาจจะเกิดขึ้น เช่น

1) วิธีขจัดความคลาดเคลื่อนจากการเลือกตัวอย่าง เช่น การสุ่มเลือกตัวอย่าง

2) วิธีขจัดความคลาดเคลื่อนในการประเมินผลการศึกษา อาจจะทำได้ดังนี้

(ก) ไม่ทราบทรีทเมนต์แบบชั้นเดียว (Single Blind) ผู้ป่วยจะไม่ทราบว่าได้รับทรีทเมนต์ใด ซึ่งยาที่ใช้ควรจะมีลักษณะเหมือนกันทุกประการ

(ข) ไม่ทราบทรีทเมนต์แบบสองชั้น (Double Blind) ทั้งผู้ป่วยและแพทย์จะไม่ทราบว่า ผู้ป่วยได้รับทรีทเมนต์ใด ส่วนยาที่ใช้นั้นควรมีลักษณะเหมือนกัน กรณีนี้สามารถลดได้ทั้งความคลาดเคลื่อนในการประเมินผลของการรักษาทั้งทางผู้ป่วยและทางแพทย์ที่ประเมินผลด้วย

(ค) ยาหลอก (Placebo) เป็นแป้งที่ทำขึ้นเพื่อใช้เปรียบเทียบผลของการได้รับยา กับไม่ได้รับยา

3) ลดความคลาดเคลื่อนของตัวอย่างโดยการทำซ้ำ หรือวัดค่าซ้ำ

ค. สถิติช่วยในการพิจารณา และชี้แนวการปรับมาตรฐานของวิธีวัด เครื่องมือที่ใช้วัด และผู้วัดค่าต่างๆ ที่ต้องการศึกษา

ง. ขั้นตอนในการวิเคราะห์ข้อมูลนั้น ต้องใช้สถิติโดยตรง

จ. ประสิทธิภาพของการทดลองทางคลินิค สถิติช่วยทำให้การศึกษามีประสิทธิภาพ (Efficiency) มากขึ้น เช่น ไม่จำเป็นต้องนำผู้ป่วยทั้งหมดที่มีลักษณะตามที่กำหนดไว้มาศึกษาทดลอง แต่จะนำมาเพียงบางส่วนด้วยวิธีการสุ่มตัวอย่าง โดยทำการคำนวณหาขนาดตัวอย่างที่เหมาะสมก่อนที่จะทำการทดลอง ซึ่งขนาดตัวอย่างที่คำนวณได้นั้นสามารถนำมาอธิบาย และสรุปผลได้ถูกต้องตามหลักสถิติ นอกจากนั้นยังสามารถอธิบายถึงอิทธิพล หรือปัจจัย (Factor) ต่างๆ ที่เกี่ยวข้องได้

การรับผู้ป่วยเข้าการทดลองอาจจะแบ่งได้เป็น 2 กรณี คือ

ก. กรณีทราบตัวผู้ป่วยแล้ว คือผู้ป่วยอยู่ที่คลินิค หรืออยู่ที่หอผู้ป่วยแล้ว

ข. กรณีไม่ทราบตัวผู้ป่วยมาก่อน

การพิจารณาขนาดตัวอย่างของข้อมูลที่ได้จากการนับ (Quantal Data or Counted Data) กรณีเลือกตัวอย่างสุ่มอย่างง่าย (Simple Random Sampling : SRS) โดยประมาณค่าสัดส่วนความชุกของโรคที่จะทำการวิจัย (Prevalence

Proportion : P) ด้วย p ด้วยการกำหนดความแม่นยำ (Precision) ของการศึกษาว่าจะยอมให้มีความผิดพลาดของโอกาสที่จะพบโรคเป็นเท่าใด (d) และความน่าจะเป็นที่ความผิดพลาดนี้จะอยู่ในขอบเขตที่ต้องการ (1-a) กล่าวคือ ต้องการให้

$$Pr(|p-P| < d) = 1-a$$

โดยที่ $$d = Z_{a/2}\sqrt{Var(p)}$$

เมื่อ Ignore fpc. $$Var(p) = d^2 / Z^2_{a/2}$$
$$\cong NPQ/(N-1)N_0$$

โดย $Q = 1-P$

N_0 = ขนาดตัวอย่าง

เมื่อ N คือ ขนาดของประชากรที่จะเลือกตัวอย่าง และถ้า N มีค่ามาก แล้ว

$$Var(p) \cong PQ/N_0$$

<u>ตัวอย่าง</u> จากการสำรวจความชุกของโรคๆ หนึ่งในชุมชนหนึ่งพบว่า มีอัตราความชุกของโรค = 40 % ยอมให้มีความผิดพลาดของโอกาสที่จะพบโรค เท่ากับ 0.05 ด้วยความเชื่อมั่น 95 % ($Z_{.025}$ = 1.96) จงหาขนาดตัวอย่างที่เหมาะสม

$$
\begin{aligned}
N_0 &= PQ/Var(p) \\
&= PQ(Z_{.025})^2/d^2 \\
&= (.4)(.6)(1.96)^2/(.05)^2 \\
&= 369
\end{aligned}
$$

2.1 แผนการสุ่มผสมสำหรับการทดลองทางคลินิค

ไซมอน (1981: 723-731) ได้ทำการศึกษาแผนแบบดังกล่าว โดยมีแผนแบบการสุ่มตลอดจนการวิเคราะห์ข้อมูล ดังต่อไปนี้

2.1.1 แผนแบบการสุ่มตัวอย่างและการกำหนดทรีทเมนต์

แผนแบบดังกล่าวสนใจศึกษาทรีทเมนต์ 2 ชนิด คือ A_1 และ A_2 ซึ่งเป็นการศึกษาร่วมกันระหว่างโรงพยาบาลที่เป็นกลุ่มหลัก หรือศูนย์กลางการวิจัย (Main Group) กับ แพทย์จากโรงพยาบาลอื่น (Private Physicians) จำนวนผู้ป่วยทั้งหมด เท่ากับ $N + nk$ โดยมีวิธีการดังนี้

2.1.1.1 โรงพยาบาลที่เป็นกลุ่มหลัก หรือศูนย์กลางการวิจัย (Main Group) มีผู้ป่วยจำนวน N คน ผู้ป่วยแต่ละคนได้รับทรีทเมนต์อย่างสุ่ม โดยที่แต่ละทรีทเมนต์จะมีผู้ป่วยจำนวนเท่ากัน คือเท่ากับ $N/2$ คน

2.1.1.2 แพทย์จากโรงพยาบาลอื่น (Private Physicians) แพทย์จำนวน K คน แต่ละคนมีผู้ป่วยอยู่ในความดูแลจำนวน n คน การกำหนดทรีทเมนต์นั้นทำโดย แบ่งแพทย์ออกเป็น 2 กลุ่ม ๆ ละ $K/2$ คน อย่างสุ่มจากนั้นกำหนดทรีทเมนต์ที่แพทย์แต่ละกลุ่มจะต้องใช้รักษาผู้ป่วยที่อยู่ในความดูแลของตนทั้ง n คน อย่างสุ่ม ฉนั้นแต่ละทรีทเมนต์จะมีผู้ป่วยจำนวนเท่ากัน คือ เท่ากับ $nK/2$ คน

2.1.2 การเก็บรวบรวมข้อมูล

พิจารณาจากการตอบสนองของผู้ป่วยที่ได้รับทรีทเมนต์ โดยแบ่งการตอบสนองออกเป็น 2 อย่าง คือ อาการเป็นที่น่าพอใจ (Success) และอาการไม่เป็นที่น่าพอใจ (Failure) โดยมีหลักการดังนี้

2.1.2.1 โรงพยาบาลที่เป็นกลุ่มหลัก หรือศูนย์กลางการวิจัย (Main Group) : กำหนดให้

G_{A1} = จำนวนผู้ป่วยที่มีอาการเป็นที่น่าพอใจ เมื่อได้รับทรีทเมนต์ A_1

G_{A2} = จำนวนผู้ป่วยที่มีอาการเป็นที่น่าพอใจ เมื่อได้รับทรีทเมนต์ A_2

2.1.2.2 แพทย์จากโรงพยาบาลอื่น (Private Physician) : กำหนดให้

S_{A1} = จำนวนผู้ป่วยที่มีอาการเป็นที่น่าพอใจ เมื่อได้รับทรีทเมนต์ A_1

S_{A2} = จำนวนผู้ป่วยที่มีอาการเป็นที่น่าพอใจ เมื่อได้รับทรีทเมนต์ A_2

2.1.3 การวิเคราะห์ข้อมูล

สำหรับการทดสอบสมมติฐานว่างว่า ทรีทเมนต์ A_1 มีประสิทธิภาพในการรักษาไม่แตกต่างจาก ทรีทเมนต์ A_2 ตัวสถิติที่ใช้ในการทดสอบ คือ

$$T_{A1} = G_{A1} + S_{A1}$$

โดย T_{A1} นี้มีการแจกแจงเข้าสู่การแจกแจงแบบปกติ ซึ่งมีค่าพารามิเตอร์ต่างๆ ดังนี้

ค่าเฉลี่ย (Mean) = $\mu_0 = T/2$

ค่าความแปรปรวน (Variance) = V_0

$$= 1/4 \left[\frac{K}{K-1} \sum_j (S_j - S_0/K)^2 + G_0(N-G_0)/(N-1) \right]$$

เมื่อ $T = T_{A1} + T_{A2}$

$S_0 = S_{A1} + S_{A2}$

$G_0 = G_{A1} + G_{A2}$

โดยที่ การทดลองนี้จะปฏิเสธสมมติฐานว่างเมื่อ $\dfrac{T_{A1} - \mu_0}{\sqrt{V_0}} \geq Z_{\alpha/2}$

ไซมอนได้ให้ข้อคิดเกี่ยวกับการศึกษาครั้งนี้ว่า วิธีการดังกล่าวจะมีอำนาจการทดสอบสูง ถ้ามีแพทย์จำนวนมาก (K มาก) และผู้ป่วยต่อแพทย์จำนวนน้อย (n น้อย) เพราะเมื่อแพทย์แต่ละคนมีผู้ป่วยในความดูแลน้อย ทำให้สามารถอธิบายขั้นตอนและวิธีการรักษา ให้กับผู้ป่วยได้อย่างชัดเจน ทำให้ผู้ป่วยเกือบทั้งหมดเต็มใจที่จะเข้าร่วมในการศึกษา ความเอนเอียงที่เนื่องมาจากการละเมิดข้อตกลงจะน้อยลง โดยปกติแล้วถ้ามีผู้ป่วยละเมิดข้อตกลงจะต้องตัดผู้ป่วยเหล่านั้นออกจากการวิเคราะห์ ซึ่งมีผลทำให้การวิเคราะห์ดังกล่าวคลาดเคลื่อนได้มาก

2.2 การวิเคราะห์ข้อมูลเชิงคุณภาพแบบแบ่งชั้นภูมิในแผนแบบสุ่มผสมสำหรับการทดลองทางคลีนิค

สำหรับการวิเคราะห์ข้อมูล และแผนแบบการสุ่มดังกล่าว เป็นแผนแบบที่ผู้วิจัยสนใจ และทำการศึกษา เนื่องจากการศึกษาของ ไซมอน นั้นเป็นวิธีการที่น่าสนใจ เพราะสามารถแก้ปัญหาเกี่ยวกับขนาดตัวอย่างได้ แต่ไม่สามารถทดสอบความเป็นอิสระระหว่างปัจจัยหรือตัวแปรได้ เมื่อสนใจตัวแปรอิสระเพิ่ม ซึ่งในการทดลองส่วนใหญ่แล้วมักจะมีตัวแปรอิสระที่ไม่ใช่ทรีทเมนต์เข้ามามีอิทธิพลต่อการทดลองเสมอ เช่น เพศ อายุ และ อาชีพ เป็นต้น หากไม่หาวิธีกำจัดหรือแยกอิทธิพลดังกล่าวออกจากการทดลองก่อน จะทำให้ผลการทดลองที่ได้คลาดเคลื่อน (Error) มีผลทำให้สรุปผลการทดลองผิดพลาด เมื่อมีการแยกตัวแปรอิสระที่ไม่ใช่ทรีทเมนต์ออกมา แล้วควรจะมีการทดสอบความสัมพันธ์ของตัวแปรอิสระดังกล่าว กับตัวแปรอื่นๆ ของการทดลองนั้นด้วย

ดังนั้น ผู้วิจัยจึงนำเอา การศึกษาของ ไซมอน การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ และตัวแบบลอกการิทึมเชิงเส้นตรง มาประยุกต์เข้าด้วยกัน ดังนี้

2.2.1 แผนแบบการสุ่มตัวอย่าง

การศึกษาครั้งนี้ กำหนดให้ปัจจัยที่ไม่ใช่ทรีทเมนต์ หรือตัวแปรอิสระ B เป็นตัวแบ่งชั้นภูมิออกเป็น L ชั้นภูมิ ซึ่งมีจำนวนผู้ป่วยทั้งหมด เท่ากับ $N_0 = M+N$ คน โดยสุ่มผู้ป่วยจากแต่ละชั้นภูมิ เท่ากับ $N_0/L = N_1 + Km_1$ คน ขั้นตอนการสุ่มมีดังต่อไปนี้ (ดังที่แสดงไว้ในรูปที่ 2.1)

2.2.1.1 โรงพยาบาลที่เป็นกลุ่มหลัก หรือศูนย์กลางการวิจัย (Main Group) มีผู้ป่วยทั้งหมดจำนวน $N = \Sigma N_1$ คน ซึ่งมีลักษณะตามเกณฑ์ที่ตั้งไว้ และผู้ป่วยเหล่านั้นเต็มใจที่จะเข้าร่วมการทดลองดังกล่าว โดยชั้นภูมิที่ 1,2,3,...,L มีผู้ป่วยจำนวน $N_1, N_2, N_3, \ldots, N_L$ คน ตามลำดับ ซึ่ง N คนนี้เป็นจำนวนผู้ป่วยทั้งหมดที่อยู่ในโรงพยาบาลที่เป็นกลุ่มหลัก หรือศูนย์กลางการวิจัย

2.2.1.2 แพทย์จากโรงพยาบาลอื่น (Private Physician) มีแพทย์จำนวน K คน ซึ่งมีผู้ป่วยอยู่ในความดูแลจำนวน $M/K = m = \Sigma m_1$ คน ต่อแพทย์ 1 คน โดยชั้นภูมิที่ 1,2,3,...,L มีผู้ป่วยต่อแพทย์ 1 คนจำนวน $m_1, m_2, m_3, \ldots, m_L$ คน ตามลำดับ

2.2.2 การกำหนดทรีทเมนต์

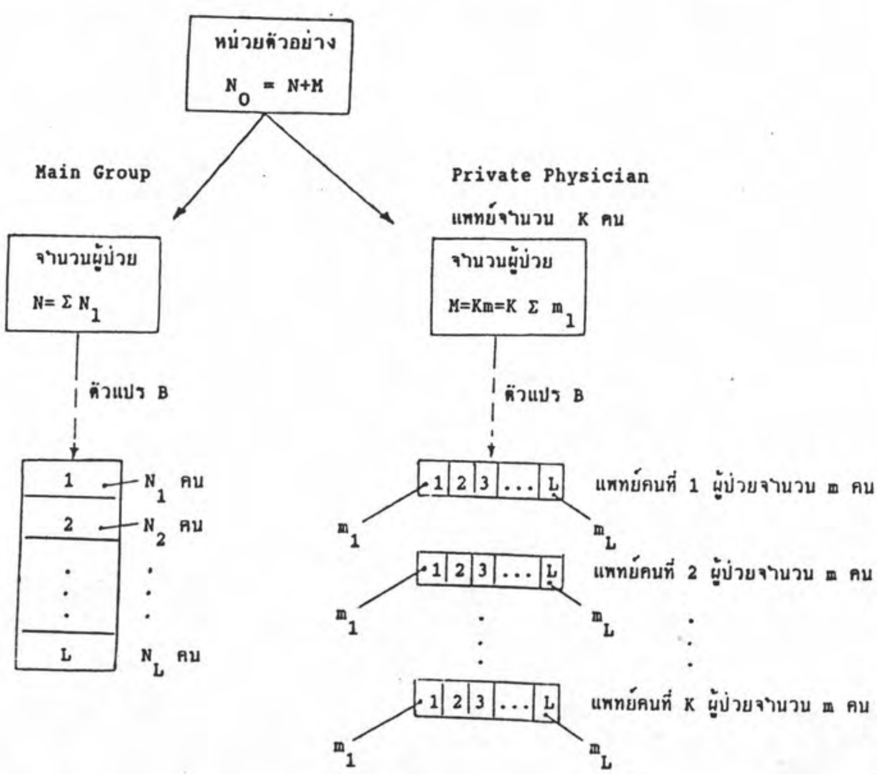
ทรีทเมนต์ที่สนใจมี 2 ทรีทเมนต์ คือ A_1 และ A_2 การกำหนดทรีทเมนต์มีขั้นตอนดังนี้ (ดังที่แสดงไว้ในรูปที่ 2.2)

2.2.2.1 โรงพยาบาลที่เป็นกลุ่มหลัก หรือศูนย์กลางการวิจัย (Main Group) : ในแต่ละชั้นภูมิแบ่งผู้ป่วยออกเป็น 2 กลุ่มโดยสุ่ม กลุ่มละ $N_1/2$ คน จากนั้นกำหนดทรีทเมนต์ให้กับแต่ละกลุ่มโดยสุ่ม ฉะนั้นแต่ละทรีทเมนต์จะมีผู้ป่วยจำนวน $N/2 = \Sigma N_1/2$ คน

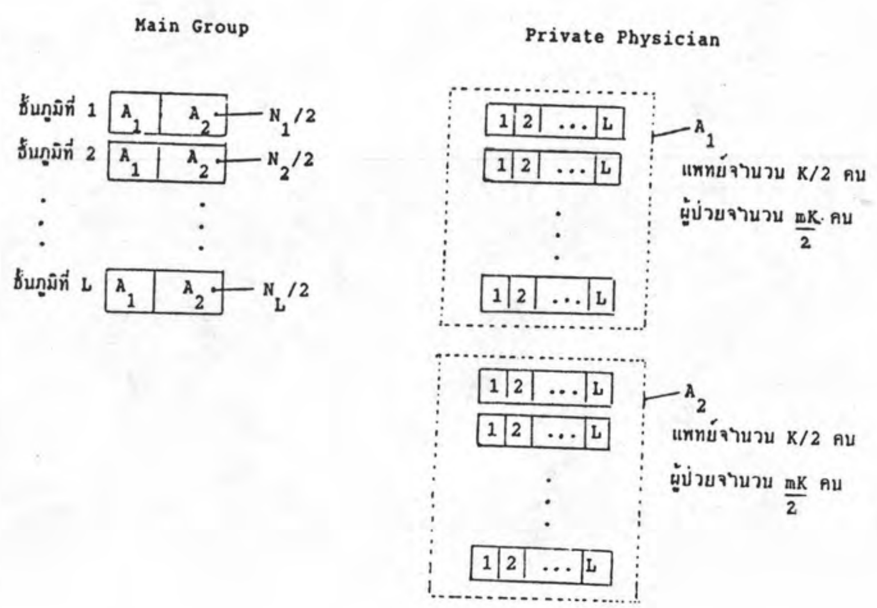
2.2.2.2 แพทย์จากโรงพยาบาลอื่น (Private Physician) : แบ่งแพทย์ออกเป็น 2 กลุ่มๆ ละ K/2 คนโดยสุ่มจากนั้นกำหนดทรีทเมนต์ที่แพทย์แต่ละ

กลุ่มจะต้องใช้รักษาผู้ป่วยที่อยู่ในความดูแลของตนทั้ง $M/K = m$ คน ฉนั้นแต่ละทรีทเมนต์จะมีผู้ป่วย จำนวน $M/2 = mK/2$ คน

รูปที่ 2.1 แสดงการสุ่มตัวอย่างของการแบ่งชั้นภูมิในแผนแบบการสุ่มผสมสำหรับการทดลองทางคลินิค



รูปที่ 2.2 แสดงการกำหนดทรีทเมนต์ A_1 และ A_2 ให้กับผู้ป่วยของการแบ่งชั้นภูมิในแผนแบบสุ่มผสมสำหรับการทดลองทางคลินิค



2.2.3 การเก็บรวบรวมข้อมูล

พิจารณาจากการตอบสนองของผู้ป่วยที่ได้รับทรีทเมนต์ โดยแบ่งการตอบสนองออกเป็น 2 อย่าง คือ อาการเป็นที่น่าพอใจ (Success) และอาการไม่เป็นที่น่าพอใจ (Failure) สำหรับ $L = 2$ สามารถเขียนเป็นตารางค่าสังเกตได้ดังตารางที่ 2.1

โดย ตัวแปรตัวที่ 1 คือ ตัวแปรตอบสนอง Y หรือตัวแปรตาม

ตัวแปรตัวที่ 2 คือ ตัวแปรอิสระ A หรือทรีทเมนต์

ตัวแปรตัวที่ 3 คือ ตัวแปรอิสระ B หรือปัจจัยที่ไม่ใช่ทรีทเมนต์

x_{ijl} คือ ความถี่ที่ได้จากการทดลอง ; $i=1,2$

$j=1,2$

$l=1,2$

ตารางที่ 2.1 แสดงตารางการเก็บรวบรวมข้อมูล เมื่อตัวแปรอิสระ B แบ่งประชากรออกเป็น 2 ชั้นภูมิ

ตัวแปรอิสระ / ตัวแปรตาม	A_1		A_2		รวม
	B_1	B_2	B_1	B_2	
อาการเป็นที่น่าพอใจ	x_{111}	x_{112}	x_{121}	x_{122}	$x_{1..}$
อาการไม่เป็นที่น่าพอใจ	x_{211}	x_{212}	x_{221}	x_{222}	$x_{2..}$
รวม	$x_{.11}$	$x_{.12}$	$x_{.21}$	$x_{.22}$	N_0

2.2.4 การวิเคราะห์ข้อมูล

สำหรับการทดสอบสมมติฐานเกี่ยวกับความเป็นอิสระระหว่างตัวแปร และความสัมพันธ์ในระดับย่อยของตัวแปรต่างๆนั้น ศึกษาเฉพาะตัวแบบลอกการิทึมเชิงเส้นตรง โดยใช้ตัวสถิติไลลิฮูด G^2

$$G^2 = 2 \quad \Sigma\Sigma\Sigma \quad [x_{ijl} \; (\ln x_{ijl} - \ln \hat{m}_{ijl})]$$

โดย $\hat{m}_{ijl}$ เป็นตัวประมาณโดยวิธีแมกซิมัมไลลิฮูด (Maximum Likelihood Estimation) ของ m_{ijl} ซึ่ง G^2 นั้นมีการแจกแจงเข้าสู่การแจกแจงแบบไคสแควร์ โดยมีองศาแห่งความเป็นอิสระที่ a^* สำหรับการหาค่า $\hat{m}_{ijl}$ และองศาแห่งความเป็นอิสระที่ a^* นั้นจะกล่าวในหัวข้อต่อไป

2.3 <u>ตัวแบบลอกการิทึมเชิงเส้นตรง</u>

สำหรับการศึกษาตัวแปร 3 ตัวในตารางการณ์จร 3 มิติ ขนาด IxJxL สัญลักษณ์ที่ใช้มีดังต่อไปนี้

p_{ijl} = ความน่าจะเป็นที่หน่วยหนึ่งๆ จะตกอยู่ในเซล (i,j,l)

x_{ijl} = ความถี่ที่ได้จากการทดลอง (Observed Frequency or Observed Counts)

m_{ijl} = ค่าคาดหวังของความถี่ (Expected Frequency or Expected Counts) ของเซล (i,j,l) = $E(x_{ijl}) = N_0 P_{ijl}$

$\hat{m}_{ijl}$ = ตัวประมาณค่าคาดหวังของความถี่ (Expected Frequency Estimate) ของ m_{ijl} โดยวิธี MLE

U = ค่าเฉลี่ยทั้งหมดของลอกการิทึมของความถี่ (Overall Mean Effect)

$U_{1(i)}$ = อิทธิพลสำคัญของระดับที่ i (Main Effect of i^{th} Category) ของตัวแปรที่ 1 อาจจะเขียนแทนด้วย U_1

$U_{2(j)}$ = อิทธิพลสำคัญของระดับที่ .j (Main Effect of j^{th} Category) ของตัวแปรที่ 2 อาจจะเขียนแทนด้วย U_2

$U_{3(l)}$ = อิทธิพลสำคัญของระดับที่ l (Main Effect of l^{th} Category) ของตัวแปรที่ 3 อาจจะเขียนแทนด้วย U_3

$U_{12(ij)}$ = อิทธิพลร่วม (Interaction) ระหว่างระดับที่ i และ j ของตัวแปรที่ 1 และ 2

$U_{13(il)}$ = อิทธิพลร่วม (Interaction) ระหว่างระดับที่ i และ l ของตัวแปรที่ 1 และ 3

$U_{23(jl)}$ = อิทธิพลร่วม (Interaction) ระหว่างระดับที่ j และ l ของตัวแปรที่ 2 และ 3

$U_{123(ijl)}$ = อิทธิพลร่วมของ 3 ตัวแปร (Second Order Efect Between the Three Variables or Three Factor Effect)

ข้อจำกัดของ U-term มีดังนี้

$$\sum_i U_1 = 0 \quad , \quad \sum_j U_2 = 0 \quad , \quad \sum_l U_3 = 0$$
$$\sum_i U_{12} = \sum_j U_{12} = 0$$
$$\sum_i U_{13} = \sum_l U_{13} = 0$$
$$\sum_j U_{23} = \sum_l U_{23} = 0$$
$$\sum_i U_{123} = \sum_j U_{123} = \sum_l U_{123} = 0$$

ในการทดสอบความเป็นอิสระระหว่างตัวแปร 3 ตัว คือการทดสอบว่า

$$p_{ijl} = p_{i..}\, p_{.j.}\, p_{..l} \quad ; \; i = 1,2,3,\ldots,I \qquad \ldots\ldots(1)$$
$$j = 1,2,3,\ldots,J$$
$$l = 1,2,3,\ldots,L$$

โดย $p_{i..} = \sum_{jl}\sum p_{ijl}$, $p_{.j.} = \sum_{il}\sum p_{ijl}$ และ $p_{..l} = \sum_{ij}\sum p_{ijl}$

เนื่องจากผลรวมของความน่าจะเป็น (P_{ijl}) ของทุกเซล ในตารางการณ์จร นั้นต้องมีค่าเท่ากับ 1 ซึ่งตามความเป็นจริงแล้ว ไม่สามารถที่จะทราบค่าความน่าจะเป็นของประชากรที่แท้จริงได้ ดังนั้น จึงใช้ความถี่จากการทดลอง (x_{ijl}) ในแต่ละเซลมาหาค่าประมาณความน่าจะเป็นดังกล่าว การหาความสัมพันธ์ระหว่างความเป็นอิสระของตัวแปรต่างๆ กับตัวแบบลอกการิทึมเชิงเส้นตรงนั้น สามารถทำได้โดยการจัดรูปสมการที่ (1) ให้อยู่ในรูปของผลบวกโดยการใส่ลอกการิทึมธรรมชาติ (Natural Logarithm) ลงไปทั้งสองข้างของสมการ ซึ่งได้ตัวแบบใหม่ดังนี้

$$\ln p_{ijl} = \ln p_{i..} + \ln p_{.j.} + \ln p_{..l} \qquad \text{.....} \ (2)$$

หรือ $v_{ijl} = \ln p_{ijl} = U + U_{1(i)} + U_{2(j)} + U_{3(l)}$ ..... (3)

เนื่องจาก $m_{ijl} = N_0\, p_{ijl}$

ฉนั้น $\ln m_{ijl} = (U + \ln N_o) + U_{1(i)} + U_{2(j)} + U_{3(l)}$ ..... (4)

$$= U' + U_{1(i)} + U_{2(j)} + U_{3(l)} \qquad \text{.....} \ (5)$$

โดยที่ $U = v_{...}/IJL$; $U' = U + \ln N_0$

$U_{1(i)} = (v_{i..}/JL) - U$

$U_{2(j)} = (v_{.j.}/IL) - U$

$U_{3(l)} = (v_{..l}/IJ) - U$

ถ้าความสัมพันธ์ระหว่าง 3 ตัวแปรไม่เป็นอิสระกัน นั่นคือตัวแบบในสมการข้างต้นจะมีอิทธิพลร่วมระหว่างตัวแปรต่างๆ เข้ามาเกี่ยวข้องด้วย ซึ่งสามารถเขียนเป็นตัวแบบอิ่มตัว (Saturated Model) ได้ดังนี้

$$v_{ijl} = U + U_{1(i)} + U_{2(j)} + U_{3(l)} + U_{12(ij)} + U_{13(il)} + U_{23(jl)} + U_{123(ijl)} \qquad \text{.....} \ (6)$$

อาจจะเขียนสั้น ๆ ได้ดังนี้

$$v_{ijl} = U + U_1 + U_2 + U_3 + U_{12} + U_{13} + U_{23} U_{123}$$

$$m_{ijl} = U' + U_1 + U_2 + U_3 + U_{12} + U_{13} + {}_{23} + U_{123} \qquad \ldots\ldots (7)$$

2.3.1 ความหมายของตัวแบบสำหรับตาราง 3 มิติ

สำหรับการทดสอบตัวแบบที่สนใจศึกษา จะเหมือนกับการทดสอบในแฟกตอเรียล (Factorial) เพราะวัตถุประสงค์ของการทดลองโดยทั่วไปต้องการที่จะทดสอบอิทธิพลเนื่องจากทรีทเมนต์ หรือปัจจัยต่างๆ และอิทธิพลร่วมระหว่างปัจจัยต่างๆ ว่ามีผลต่อการทดลองหรือไม่ การทดสอบในแฟกตอเรียลนั้นมีพื้นฐานคล้ายกับ การทดสอบตัวแบบลอกการิทึมเชิงเส้นตรง (Log Linear Model) คือ สนใจปัจจัยตั้งแต่ 2 ตัวขึ้นไป โดยปัจจัยแต่ละตัวมีหลายระดับ และต้องการที่จะทดสอบอิทธิพลของแต่ละปัจจัย กับอิทธิพลร่วมต่างๆ ที่เป็นไปได้ เช่น สนใจทรีทเมนต์ 3 ชนิด คือ A , B และ C แต่ละชนิดมี 2 ระดับ การทดสอบที่ต้องการศึกษา คือ

- ไม่มีอิทธิพลเนื่องจากทรีทเมนต์ A
- ไม่มีอิทธิพลเนื่องจากทรีทเมนต์ B
- ไม่มีอิทธิพลเนื่องจากทรีทเมนต์ C
- ไม่มีอิทธิพลร่วมระหว่างทรีทเมนต์ A และ B
- ไม่มีอิทธิพลร่วมระหว่างทรีทเมนต์ A และ C
- ไม่มีอิทธิพลร่วมระหว่างทรีทเมนต์ B และ C
- ไม่มีอิทธิพลร่วมระหว่างทรีทเมนต์ A , B และ C

ดังนั้น การทดสอบตัวแบบที่นำมาเสนอ ณ.ที่นี้ จะมีหลักการเช่นเดียวกับการทดสอบตัวแบบในแฟกตอเรียล ซึ่งตัวแบบและความหมายของตัวแบบที่น่าสนใจมีดังต่อไปนี้

2.3.1.1 ตัวแบบที่ไม่มีอิทธิพลเนื่องจากตัวแปรที่ 1

นั่นคือ ความน่าจะเป็นของแต่ละระดับของตัวแปรที่ 1 เท่ากัน

สมมติฐานว่าง คือ $H_o : U_1=U_{12}=U_{13}=U_{123}=0$ หรือ $P_{ijl}=P_{jl}$; ทุกๆ i

ตัวแบบ คือ $\ln p_{ijl} = U+U_2+U_3+U_{23}$

สำหรับตัวแบบนี้เขียนแทนด้วย [23]

2.3.1.2 ตัวแบบที่ไม่มีอิทธิพลเนื่องจากตัวแปรที่ 2

นั่นคือ ความน่าจะเป็นของแต่ละระดับของตัวแปรที่ 2 เท่ากัน

สมมติฐานว่าง คือ $H_o : U_2=U_{12}=U_{23}=U_{123}=0$ หรือ $P_{ijl}=P_{il}$; ทุกๆj

ตัวแบบ คือ $\ln p_{ijl} = U+U_1+U_3+U_{13}$

สำหรับตัวแบบนี้เขียนแทนด้วย [13]

2.3.1.3 ตัวแบบที่ไม่มีอิทธิพลเนื่องจากตัวแปรที่ 3

นั่นคือ ความน่าจะเป็นของแต่ละระดับของตัวแปรที่ 3 เท่ากัน

สมมติฐานว่าง คือ $H_o : U_3=U_{13}=U_{23}=U_{123}=0$ หรือ $P_{ijl}=P_{ij}$; ทุกๆl

ตัวแบบ คือ $\ln p_{ijl} = U+U_1+U_2+U_{12}$

สำหรับตัวแบบนี้เขียนแทนด้วย [12]

2.3.1.4 ตัวแบบที่ไม่มีอิทธิพลร่วมระหว่างตัวแปรที่ 1 และ 2 สำหรับทุกระดับของตัวแปรที่ 3

นั่นคือ ตัวแปรที่ 1 และ 2 เป็นอิสระต่อกันเมื่อกำหนด ตัวแปรที่ 3 ให้คงที่ในระดับต่างๆ (Conditional Independence Model of Variable 1 and 2 Given the Level of Variable 3)

สมมติฐานว่าง คือ $H_o : U_{12}=U_{123}=0$

ตัวแบบ คือ $\ln p_{ijl} = U+U_1+U_2+U_3+U_{13}+U_{23}$

สำหรับตัวแบบนี้เขียนแทนด้วย [13][23]

2.3.1.5 ตัวแบบที่ไม่มีอิทธิพลร่วมระหว่างตัวแปรที่ 1 และ 3 สำหรับทุกระดับของตัวแปรที่ 2

นั่นคือ ตัวแปรที่ 1 และ 3 เป็นอิสระต่อกันเมื่อกำหนด ตัวแปรที่ 2 ให้คงที่ในระดับต่างๆ (Conditional Independence Model of Variable 1 and 3 Given the Level of Variable 2)

สมมติฐานว่าง คือ $H_o : U_{13}=U_{123}=0$

ตัวแบบ คือ $\ln p_{ijl} = U+U_1+U_2+U_3+U_{12}+U_{23}$

สำหรับตัวแบบนี้เขียนแทนด้วย [12][23]

2.3.1.6 ตัวแบบที่ไม่มีอิทธิพลร่วมระหว่างตัวแปรที่ 2 และ 3 สำหรับทุกระดับของตัวแปรที่ 1

นั่นคือ ตัวแปรที่ 2 และ 3 เป็นอิสระต่อกันเมื่อกำหนด ตัวแปรที่ 1 ให้คงที่ในระดับต่างๆ (Conditional Independence Model of Variable 2 and 3 Given the Level of Variable 1)

สมมติฐานว่าง คือ $H_o : U_{23}=U_{123}=0$

ตัวแบบ คือ $\ln p_{ijl} = U+U_1+U_2+U_3+U_{12}+U_{13}$

สำหรับตัวแบบนี้เขียนแทนด้วย [12][13]

2.3.1.7 ตัวแบบที่ทั้ง 3 ตัวแปรเป็นอิสระต่อกัน

นั่นคือ ไม่มีความสัมพันธ์ระหว่าง 3 ตัวแปร (Model of Complete Independence)

สมมติฐานว่าง คือ $H_o : U_{12}=U_{13}=U_{23}=U_{123}=0$

หรือ P_{ijl} เท่ากัน ทุกๆ i,j และ l

ตัวแบบ คือ $\ln p_{ijl} = U+U_1+U_2+U_3$

สำหรับตัวแบบนี้เขียนแทนด้วย [1][2][3]

สำหรับการวิจัยครั้งนี้ ตัวแปรที่ 1 คือ ตัวแปรตอบสนอง หรือตัวแปรตาม Y

ตัวแปรที่ 2 คือ ทรีทเมนต์ หรือตัวแปรอิสระ A

ตัวแปรที่ 3 คือ ปัจจัยที่ไม่ใช่ทรีทเมนต์ หรือตัวแปรอิสระ B

สมมติฐานว่าง หรือตัวแบบที่สนใจศึกษามี 3 ข้อ ดังนี้

- ทั้ง 3 ตัวแปรเป็นอิสระกันอย่างสมบูรณ์
- ไม่มีอิทธิพลเนื่องจากทรีทเมนต์ หรือตัวแปรที่ 2 (ตัวแปรอิสระ A)
- ไม่มีอิทธิพลเนื่องจากปัจจัยที่ไม่ใช่ทรีทเมนต์ หรือตัวแปรที่ 3 (ตัวแปรอิสระ B)

2.3.2 วิธีคำนวณหาค่าประมาณค่าคาดหวังของความถี่ (Expected Frequency Estimate ; $\hat{m}_{ijl}$)

ในที่นี้จะพิจารณาการประมาณแบบแมกซิมัมไลลิฮูด (MLE) ซึ่งสามารถหาได้ง่าย และให้ค่าประมาณค่าคาดหวังของความถี่โดยไม่ต้องหาพารามิเตอร์อีก สำหรับโครงสร้าง หรือลักษณะของตารางจะพิจารณาเฉพาะ วิธีการสุ่มแบบพหุนาม (Multinomial Sampling) กำหนดขนาดตัวอย่างเท่ากับ N_0 การแจกแจงแบบพหุนามมี Probability Density Functions (pdf) ดังนี้

$$f(\{x_{ijl}\}) = \frac{N_0!}{\pi_{ijl}\, x_{ijl}!} \pi_{ijl} (p_{ijl})^{x_{ijl}}$$

$$= \frac{N_0!}{\pi_{ijl}\, x_{ijl}!} \pi_{ijl} (m_{ijl}/N_0)^{x_{ijl}} \qquad \text{..... (8)}$$

$$= \exp\left[\ln \frac{N_0!}{\pi_{ijl}\, x_{ijl}!} + \Sigma\Sigma\Sigma\, x_{ijl} \ln m_{ijl} - N_0 \ln N_0 \right] \qquad \text{..... (9)}$$

พิจารณาเทอม $\Sigma\Sigma\Sigma\, x_{ijl} \ln m_{ijl}$ เมื่อแทน $\ln m_{ijl}$ ด้วยสมการที่ (7) ดังต่อไปนี้

$$\Sigma\Sigma\Sigma\, x_{ijl} \ln m_{ijl} = N_0 U' + \Sigma_i x_{i..} U_1 + \Sigma_j x_{.j.} U_2 + \Sigma_l x_{..l} U_3 + \Sigma_{ij} x_{ij.} U_{12} + \Sigma_{il} x_{i.l} U_{13} + \Sigma_{jl} x_{.jl} U_{23} + \Sigma_{ijl} x_{ijl} U_{123} \qquad \text{....(10)}$$

เนื่องจากการแจกแจงแบบพหุนามจัดอยู่ในกลุ่ม Exponential Class ดังนั้น ตัวสถิติที่เพียงพอ (Sufficient Statistics) สำหรับพารามิเตอร์ m_{ijl} คือเทอมของ x ที่อยู่ติดกับเทอม U สำหรับสมการที่ (10) ซึ่งได้มาจากตัวแบบอิ่มตัว จะมีตัวสถิติที่เพียงพอ ดังนี้

$$N_0, x_{i..}, x_{.j.}, x_{..l}, x_{ij.}, x_{i.l}, x_{.jl} \text{ และ } x_{ijl}$$

การศึกษาวิจัยของ Birch พบว่าเมื่อมี Minimal Set ของตัวสถิติที่เพียงพอ และ θ_i เป็นเซ็ทของ Subscripts ผลคือ

1) ถ้า x_{θ_i} เป็นสมาชิกของเซ็ทตัวสถิติที่เพียงพอแล้ว MLE ของ m_{θ_i} คือ x_{θ_i} เช่น ถ้า $x_{ij.}, x_{i.l}, x_{.jl}$ เป็นสมาชิกของเซ็ทตัวสถิติที่เพียงพอแล้ว MLE ของ $m_{ij.}, m_{i.l}, m_{.jl}$ คือ $x_{ij.}, x_{i.l}, x_{.jl}$ ตามลำดับ

2) MLE ของ $\{m_{ijl}\}$ จะมีเพียงชุดเดียวเท่านั้น ที่สอดคล้องกับเงื่อนไขของตัวแบบ และเงื่อนไขในข้อ 1)

วิธีคำนวณหา MLE ของ m_{ijl} นั้นมี 2 วิธี คือ

1) การประมาณโดยตรง (Direct Estimates) เป็นวิธีที่ใช้ได้สำหรับตัวแบบบางตัวแบบเท่านั้น โดยที่ m_{ijl} จะเป็นฟังก์ชันของตัวสถิติที่เพียงพอ

2) การประมาณโดยกระทำซ้ำของค่าคาดหวัง (Iterative Computation of Expect Values) จะใช้ในกรณีที่หาโดยวิธี การประมาณโดยตรง (Direct Estimate)ไม่ได้ ซึ่งต้องใช้คอมพิวเตอร์เข้ามาช่วย

สำหรับการวิจัยครั้งนี้ตัวแบบทั้ง 7 ที่กล่าวไว้ข้างต้นนั้นสามารถหา MLE ของ m_{ijl} ได้โดยวิธี Direct Estimate ซึ่งขั้นตอนการหานั้นสามารถเข้าใจได้ง่ายตามตัวอย่างที่จะกล่าวต่อไปนี้

ตัวอย่าง สำหรับตัวแบบ [13][23]

$$\ln p_{ijl} = U+U_1+U_2+U_3+U_{13}+U_{23}$$

$$m_{ijl} = \exp\,(U'+U_1+U_2+U_3+U_{13}+U_{23})$$

ตัวสถิติที่เพียงพอ คือ $N_0, x_{i..}, x_{.j.}, x_{..l}, x_{i.l}, x_{.jl}$

ตัวสถิติที่เพียงพอต่ำสุด (Minimal Sufficient Statistics) คือ $x_{i.l}, x_{.jl}$

Configulation ที่เพียงพอต่ำสุด คือ C_{13}, C_{23}

ผลของ Birch ได้ว่า $m_{i.l} = x_{i.l} = \exp(U'+U_1+U_3+U_{13}) \cdot \exp(U_2+U_{23})$

$$m_{.jl} = x_{.jl} = \exp(U'+U_2+U_3+U_{23}) \cdot \exp(U_1+U_{13})$$

$$m_{..l} = x_{..l} = \exp(U'+U_3) \cdot \exp(U_1+U_2+U_{13}+U_{23})$$

ฉนั้น $$\hat{m}_{ijl} = \frac{x_{i.l}\,x_{.jl}}{x_{..l}}$$

สำหรับค่าประมาณค่าคาดหวังของความถี่ ($\hat{m}_{ijl}$) ของสมมติฐานว่างทั้ง 7 ข้อนั้นได้แสดงไว้ในตารางที่ 2.3

2.3.3 วิธีคำนวณหาองศาแห่งความเป็นอิสระ (Degree of Freedom ; a^*)

สำหรับตัวแบบไม่อิ่มตัวนั้นสามารถคำนวณหาองศาแห่งความเป็นอิสระได้ 2 วิธีดังนี้

2.3.3.1 นับจำนวนพารามิเตอร์อิสระของเซ็ทพารามิเตอร์ที่เท่ากับศูนย์ หรือพารามิเตอร์อิสระที่ปรากฏอยู่ในสมมติฐานว่าง เช่น $H_o : U_{12}=U_{123}=0$

มีพารามิเตอร์อิสระ = $(I-1)(J-1)+(I-1)(J-1)(L-1)$

ฉนั้น องศาแห่งความเป็นอิสระ $a^* = L(I-1)(J-1)$

2.3.3.2 นับจำนวนพารามิเตอร์ที่ต้องประมาณ หรือพารามิเตอร์อิสระที่ปรากฏอยู่ในตัวแบบ (T_p) แล้วลบออกจากจำนวนเซลทั้งหมด (T_e) เช่น

$$\ln p_{ijl} = U+U_1+U_2+U_3$$

มีพารามิเตอร์อิสระที่ต้องประมาณ = $1+(I-1)+(J-1)+(L-1) = T_p$

ฉนั้น องศาแห่งความเป็นอิสระ $a^* = T_e - T_p$

$$= IJL-\{1+(I-1)+(J-1)+(L-1)\}$$

$$= IJL-(I+J+L)+2$$

สำหรับจำนวนพารามิเตอร์อิสระของตารางขนาด 3 มิติ นั้นได้แสดงไว้ใน ตารางที่ 2.3

ตารางที่ 2.2 แสดงองศาแห่งความเป็นอิสระของแต่ละ U-term สำหรับตารางขนาด 3 มิติ

เทอม U	จำนวนพารามิเตอร์อิสระ
U_{123}	(I-1)(J-1)(L-1)
U_{12}	(I-1)(J-1)
U_{13}	(I-1)(L-1)
U_{23}	(J-1)(L-1)
U_{1}	(I-1)
U_{2}	(J-1)
U_{3}	(L-1)
รวม	IJL

ตารางที่ 2.3 แสดงตัวแบบลอกการิทึมเชิงเส้นตรงที่สนใจพร้อมทั้ง ค่าคาดหวังของความถี่ และองศาแห่งความเป็นอิสระ

ตัวแบบ	ค่าคาดหวังของความถี่ ($\hat{m}_{ijl}$)	องศาแห่งความเป็นอิสระ
1.[23]	$x_{.jl}/I$	JL(I-1)
2.[13]	$x_{i.l}/J$	IL(J-1)
3.[12]	$x_{ij.}/L$	IJ(L-1)
4.[13][23]	$x_{i.l}x_{.jl}/x_{..l}$	L(I-1)(J-1)
5.[12][23]	$x_{ij.}x_{.jl}/x_{.j.}$	J(I-1)(L-1)
6.[12][13]	$x_{ij.}x_{i.l}/x_{i..}$	I(J-1)(L-1)
7.[1][2][3]	$x_{i..}x_{.j.}x_{..l}/N_0^2$	IJL-I-J-L+2

2.3.4 ตัวสถิติไลลิฮูด G^2

ในการทดสอบสมมติฐานของตัวแบบลอกการิทึมเชิงเส้นตรงในครั้งนี้ ใช้ตัวสถิติไลลิฮูด G^2 ซึ่งมีพื้นฐานมาจากการทดสอบอัตราส่วนภาวะน่าจะเป็น (Likelihood Ratio Test) โดยมีขั้นตอนการคำนวณหาตัวสถิติดังกล่าว ตามนิยามและทฤษฎี ดังต่อไปนี้

<u>นิยาม</u> การทดสอบด้วยอัตราส่วนภาวะน่าจะเป็นของสมมติฐาน $H_0 : \Omega_o \in \Omega$ กับ $H_a : \Omega_a \in \Omega$ เมื่อ $\Omega_o \cup \Omega_a = \Omega$ ใช้ตัวสถิติ

$$\lambda = \frac{L(\hat{\Omega}_o)}{L(\hat{\Omega})}$$

เมื่อ $L(\hat{\Omega}_o)$ และ $L(\hat{\Omega})$ เป็นฟังก์ชั่นภาวะน่าจะเป็นเมื่อแทนค่าพารามิเตอร์ด้วย

MLE ภายใต้ข้อกำหนดว่า $\Theta \in \Omega_o$ และ $\Theta \in \Omega$ ตามลำดับ การทดสอบนี้จะปฏิเสธสมมติฐานว่าง เมื่อ $\lambda \leq K$ เมื่อ K เป็นค่าที่ทำให้ $P(\lambda \leq K / \Theta \in \Omega_o) = \alpha$ ที่กำหนดไว้เป็นระดับนัยสำคัญ

<u>ทฤษฎี</u> ภายใต้เงื่อนไขปกติ (Regularity Conditions) ของฟังก์ชั่นความหนาแน่น $f(x; \Theta)$ ถ้าขนาดตัวอย่าง n ใหญ่ ตัวแปรสุ่ม $-2\ln \lambda$ จะมีการแจกแจงโดยประมาณแบบ χ^2 โดยมีองศาแห่งความเป็นอิสระเท่ากับความแตกต่างระหว่างจำนวนพารามิเตอร์ใน Ω และ Ω_o ที่ไม่ได้กำหนดค่า

สำหรับการวิจัยนี้ข้อมูลที่สนใจมีการแจกแจงแบบพหุนาม ซึ่งมี pdf ดังนี้

$$f(\{x_{ijl}\} : \{p_{ijl}\}) = \frac{N_o!}{\prod_{ijl} x_{ijl}!} \prod_{ijl} (p_{ijl})^{x_{ijl}}$$

เมื่อ $\Omega_o = \{p_{ijl} ; p_{ijl} = \hat{m}_{ijl}/N_o\}$ โดย $\hat{m}_{ijl}$ เป็น MLE ของค่าคาดหวังความถี่

$\Omega = \{p_{ijl} ; p_{ijl} = x_{ijl}/N_o\}$

$$L(\Omega_o) = \frac{N_o!}{\prod_{ijl} x_{ijl}!} \prod_{ijl} (m_{ijl}/N_o)^{x_{ijl}}$$

$$L(\Omega) = \frac{N_o!}{\prod_{ijl} x_{ijl}!} \prod_{ijl} (x_{ijl}/N_o)^{x_{ijl}}$$

$$-2\ln\lambda = -2 \Sigma\Sigma\Sigma\, x_{ijl}(\ln m_{ijl} - \ln x_{ijl})$$

$$G^2 = 2 \Sigma\Sigma\Sigma\, x_{ijl}(\ln x_{ijl} - \ln m_{ijl})$$

จากทฤษฎีข้างต้น G^2 จะมีการแจกแจงเข้าสู่การแจกแจงแบบไคสแควร์ โดยมีองศาแห่งความเป็นอิสระที่ a^*

2.3.5 การทดสอบความเป็นอิสระโดยใช้ตัวแบบลอกการิทึมเชิงเส้นตรง

ในการทดสอบว่าข้อมูลในตารางการณ์จรมีลักษณะเป็นไปตามข้อสมมติฐานว่างหรือไม่นั้นสามารถทดสอบได้ตามขั้นตอนต่อไปนี้

2.3.5.1 คำนวณค่าประมาณค่าคาดหวังของความถี่ (m_{ijl}) ภายใต้ตัวแบบที่ศึกษา

2.3.5.2 คำนวณค่าสถิติไลลิฮูด G^2 โดยที่

$$G^2 = 2 \sum\sum\sum x_{ijl}(\ln x_{ijl} - \ln m_{ijl})$$

ซึ่งตัวสถิติไลลิฮูด G^2 นั้นมีการแจกแจงเข้าสู่การแจกแจงแบบไคสแควร์ (Asymptotically Distributed as χ^2) ซึ่งมีองศาแห่งความเป็นอิสระที่ a^*

2.3.5.3 นำค่าสถิติไลลิฮูด G^2 ที่คำนวณได้จากข้อ 2.3.5.2 ซึ่งมีองศาแห่งความเป็นอิสระที่ a^* มาเปรียบเทียบกับค่า χ^2 จากตารางการแจกแจงแบบไคสแควร์ที่องศาแห่งความเป็นอิสระเดียวกัน

ผลการทดสอบจะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่าสถิติไลลิฮูด G^2 จากการคำนวณมีค่ามากกว่าค่า χ^2 จากตารางการแจกแจงแบบไคสแควร์ที่องศาแห่งความเป็นอิสระที่ a^* ณ. ระดับนัยสำคัญที่กำหนดไว้ นั่นคือตัวแบบดังกล่าวไม่เหมาะสมกับข้อมูลนั้น

ผลการทดสอบจะยอมรับสมมติฐานว่าง H_0 เมื่อค่าสถิติไลลิฮูด G^2 จากการคำนวณมีค่าน้อยกว่าค่า χ^2 จากตารางการแจกแจงแบบไคสแควร์ที่องศาแห่งความเป็นอิสระที่ a^* ณ.ระดับนัยสำคัญที่กำหนดไว้ นั่นคือตัวแบบดังกล่าวเหมาะสมกับข้อมูลนั้น

2.4 การแจกแจงแบบทวินาม (Binomial Distribution)

ลักษณะของการทดลองแบบทวินาม คือ

2.4.1 การทดลองซ้ำๆ กัน N ครั้ง

2.4.2 ในการทดลองแต่ละครั้ง ผลที่เกิดขึ้นมี 2 อย่างเท่านั้น คือความสำเร็จ (Success) และความล้มเหลว (Failure)

2.4.3 ความน่าจะเป็นของความสำเร็จที่เกิดจากการทดลองแต่ละครั้งที่มีค่าคงที่เท่ากับ P และความน่าจะเป็นของความล้มเหลวเท่ากับ 1-P

2.4.4 การทดลองแต่ละครั้งเป็นอิสระจากกัน

นิยาม

ตัวแปรสุ่ม x จะมีการแจกแจงแบบทวินาม ก็ต่อเมื่อการแจกแจงความน่าจะเป็นของ x คือ

$$f(x;N,P) = \binom{N}{x} P^{x}(1-P)^{N-x} \quad \text{เมื่อ } x=0,1,2,\dots,N$$

x คือจำนวนครั้งของความสำเร็จที่เกิดจากการทดลอง

N คือจำนวนครั้งทั้งหมดของการทดลอง

P คือความน่าจะเป็นของความสำเร็จที่เกิดจากการทดลองแต่ละครั้ง

การแจกแจงแบบทวินามนี้ จะมีค่า N และ P เป็นค่าพารามิเตอร์ ฉนั้นการแจกแจงนี้จะมีลักษณะอย่างไรจึงขึ้นอยู่กับค่าทั้งสอง

สำหรับการวิจัยครั้งนี้ค่า x_{1jl} มีการแจกแจงแบบทวินาม พารามิเตอร์ คือ $N = x_{.jl}$ และ P_{1jl} ซึ่งในการทดลองแต่ละครั้งผลที่เกิดขึ้นมี 2 อย่าง คือ อาการเป็นที่น่าพอใจ (Success) และ อาการไม่เป็นที่น่าพอใจ (Failure) ตัวอย่างดังที่แสดงไว้ในตารางที่ 2.1

2.5 ตัวอย่างการใช้ตัวแบบลอกการิทึมเชิงเส้นตรง

ข้อมูลตัวอย่างที่นำมาแสดงนี้ได้มาจากการทดลองของ แฮสติ้ง และคณะ (Hasting and other: 1978 :1041-1045) ซึ่งทำการทดลองเกี่ยวกับ ยาลดกรดในการป้องกันภาวะเลือดออกในกระเพาะอาหารและลำไส้แบบเฉียบพลัน (Antacid Titration in the Prevention of acute Gastrointertinal Bleeding) โดยมีวิธีการดังต่อไปนี้

2.5.1 แผนการสุ่มตัวอย่าง

การวิจัยดังกล่าวพิจารณา จำนวนของปัจจัยการเสี่ยง (No.of Risk Factor) ที่จะเป็นแผลในกระเพาะอาหาร และลำไส้ เป็นปัจจัยที่ไม่ใช่ทรีทเมนต์ (ตัวแปรอิสระ B) เพื่อใช้ในการแบ่งประชากรออกเป็น 3 ชั้นภูมิ โดยสุ่มตัวอย่างมารวมทั้งสิ้น 100 คน สามารถจำแนกตามชั้นภูมิต่างๆ ได้ดังนี้

ชั้นภูมิที่ 1 จำนวนของปัจจัยการเสี่ยง คือ 0-1 จำนวน 38 คน
ชั้นภูมิที่ 2 จำนวนของปัจจัยการเสี่ยง คือ 2 จำนวน 23 คน
ชั้นภูมิที่ 3 จำนวนของปัจจัยการเสี่ยง คือ 3-6 จำนวน 39 คน

ในแต่ละชั้นภูมิแบ่งออกเป็น 2 กลุ่มโดยสุ่ม ซึ่งกลุ่มหนึ่งจะมีจำนวน 51 คนที่ได้รับยาลดกรด (Antacid) ส่วนอีกกลุ่มหนึ่งมีจำนวน 49 คน เป็นกลุ่มควบคุม (Control)

2.5.2 การเก็บรวบรวมข้อมูล

พิจารณาจากการตอบสนองของผู้ป่วยที่ได้รับทรีทเมนต์ โดยแบ่งการตอบสนองออกเป็น 2 อย่าง คือ อาการเป็นที่น่าพอใจ (Success) และอาการไม่เป็นที่น่าพอใจ (Failure) จากการทดลองดังกล่าวได้ค่าสังเกต ดังตารางที่ 2.4

ตารางที่ 2.4 แสดงตัวอย่างข้อมูลของการทดลองเกี่ยวกับยาลดกรด (Antacid) ในการป้องกันภาวะเลือดออกในกระเพาะอาหารและลำไส้แบบเฉียบพลัน

ทรีทเมนต์ / risk / การตอบสนอง	Antacid			Control			Total
	0-1	2	3-6	0-1	2	3-6	
Success	24	8	17	13	12	12	86
Failure	0	0	2	1	3	8	14
Total	24	8	19	14	15	20	100

ส่วนการวิเคราะห์ข้อมูล และการสรุปผล จะกล่าวต่อไปในบทที่ 3