บทที่ 2

สถิติที่ใช้ในงานวิจัย

ในงานวิจัยนี้ได้พิจารณาถึงตัวสถิติแคปป้า ที่ใช้แสดงความสัมพันธ์ของลักษณะของข้อมูล โดยที่ Cohen (1960) ได้มีการคิดค้นการหาค่าสัมประสิทธิ์แคปป้าขึ้นมา โดยได้สามารถแสดงอยู่ในรูป

$$\kappa = \frac{P_a - P_e}{1 - P_e} \tag{1}$$

P_a คือ ความน่าจะเป็นที่จะได้ผลที่สอดคล้องกัน

P_e คือ ความน่าจะเป็นที่จะได้ผลที่เปลี่ยนแปลงไป

$1\text{-}P_e$ คือ ความน่าจะเป็นที่จะได้ผลที่สอดคล้องกันสูงสุด

จากสมการค่าสัมประสิทธิ์แคปป้าที่ได้จะเป็นพื้นฐานในการหาค่าตัวประมาณโดยยึดหลักของสมการ (1) เป็นแนวทางต่อไป

2.1 วิธีการประมาณค่าพารามิเตอร์ที่ใช้ในงานวิจัย

การศึกษาถึงค่าพารามิเตอร์ที่ใช้ในงานวิจัยนี้ ได้แสดงถึงการหาตัวประมาณค่าสัมประสิทธิ์แคปป้า ซึ่งได้มีการพัฒนาโดยยึดหลักของสมการที่ (1) จากวิธีการหาตัวประมาณทั้ง 5 วิธี

2.1.1 หลักการของ Fleiss ในปี ค.ศ. 1971 ($\kappa1$) ได้แสดงการหาตัวประมาณสัมประสิทธิ์แคปป้าโดยอาศัยสมการ (1) โดยได้กำหนดให้

$$p_a = \frac{1}{n(n-1)}\left(\sum_{j=1}^{k} n_{ij}^{2} - n_{ij}\right)$$

$$p_e = \sum_{j=1}^{k} p_{j.}^{2}$$

จะได้ตัวประมาณค่าสัมประสิทธิ์แคปป้าอยู่ในรูป

$$K1 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k} n_{ij}^2 - Nn[1+(n-1)\sum_{j=1}^{k} p_j^2]}{Nn(n-1)(1-\sum_{j=1}^{k} p_j^2)} \qquad (2)$$

และ ยังได้แสดงถึงวิธีหาตัวประมาณความแปรปรวนของตัวประมาณที่ได้ อยู่ในรูป

$$\mathrm{var}(K1) = \frac{2}{Nn(n-1)} * \frac{\sum_j p_j^2 - (2n-3)(\sum_j p_j^2)^2 + 2(n-2)\sum_j p_j^3}{(1-\sum_j p_j^2)^2} \qquad (3)$$

โดยได้กำหนดตัวแปรต่าง ๆ ในสมการ ได้แก่

p_a เป็นตัวประมาณของความน่าจะเป็นที่จะได้ผลที่สอดคล้องกัน

p_e เป็นตัวประมาณของความน่าจะเป็นที่จะได้ผลที่เปลี่ยนแปลงไป

N หมายถึง ขนาดตัวอย่างทั้งหมดที่ใช้ในการทดสอบ

k หมายถึง จำนวนผลที่ได้ทั้งหมดที่ได้ในการทดสอบ

n_{ij} หมายถึง หน่วยตัวอย่างที่ i ที่ได้ผลการทดสอบที่ j

i = 1 , 2 , 3 , ... , N

j = 1 , 2 , 3 , ... , k

n หมายถึง จำนวนวิธีทั้งหมดที่ใช้ในการทดลอง โดยที่ $n = \sum_{j=1}^{k} n_{ij}$

p_j เป็นตัวประมาณความน่าจะเป็นที่ได้ผลการทดสอบที่ j โดยที่ $\sum_{j=1}^{k} p_j = 1$

ตัวประมาณพารามิเตอร์ต่าง ๆ ที่แสดงในข้างต้นเป็นการวิเคราะห์โดยที่ลักษณะข้อมูลใช้วิธีทั้งหมดในการทดสอบมากกว่าหรือเท่ากับ 2 วิธีขึ้นไป และจำนวนผลที่เป็นไปได้ทั้งหมดในแต่ละวิธีที่ทำการทดสอบมีผลลัพธ์ที่เท่า ๆ กันและมีจำนวนมากกว่าหรือเท่ากับ 2 ผลลัพธ์ขึ้นไป สามารถเขียนการจัดเก็บข้อมูลได้อยู่ในตารางที่ 2.1

<u>ตารางที่ 2.1</u> ตารางแสดงการจัดเก็บข้อมูลในการหาค่าสัมประสิทธิ์แคปป้าจากหลักการของของ Fleiss ในปี ค.ศ. 1971

หน่วยตัวอย่างที่	ผลที่ได้จากวิธีการทดสอบ				รวม
	ผลลัพธ์ที่1 (j=1)	ผลลัพธ์ที่ 2 (j=2)	. . . .	ผลลัพธ์ที่ k (j=k)	
1	n_{11}	n_{12}		n_{1k}	n
2	n_{21}	n_{22}		n_{2k}	n
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.
N	n_{N1}	n_{N2}		n_{Nk}	n
รวม	$n_{.1}$	$n_{.2}$	. . . .	$n_{.k}$	nN
p_j	$n_{\bullet 1}/nN$	$n_{\bullet 2}/nN$	. . . .	$n_{\bullet k}/nN$	1

สำหรับในขอบเขตของงานวิจัยนี้พิจารณาเฉพาะในกรณีที่ข้อมูลอยู่ในตาราง 2x2 ดังนั้น สามารถประยุกต์วิธีข้างต้นมาอยู่ในรูปการจัดเก็บข้อมูลในตาราง 2x2 จะได้

$$K1 = \frac{4(n_{00}+n_{11})+2(n_{01}+n_{10})-2n\sum_{i=0}^{1}p_i^2}{2n(1-\sum_{i=0}^{1}p_i^2)} \tag{4}$$

และ

$$\text{var}(K1) = \frac{\sum_{i=0}^{1}p_i^2}{n(1-\sum_{i=0}^{1}p_i^2)} \tag{5}$$

โดยกำหนดตัวแปรต่าง ๆ จากข้างต้นจะกลายเป็น

N = n หมายถึง ขนาดตัวอย่างทั้งหมดที่ใช้ในการทดสอบ

k = 2 คือ จำนวนผลที่ได้ทั้งหมดที่ได้ในการทดสอบจะมี 2 ลักษณะ ได้แก่ สิ่งที่สนใจ กับ สิ่งที่ไม่สนใจ

n = 2 คือ จำนวนวิธีที่ใช้ในการทดลองจะมีทั้งหมด 2 วิธี

$p_j = p_i$ เป็นตัวประมาณความน่าจะเป็นที่ได้ผลการทดสอบที่ i โดยที่ $\sum_{i=0}^{1} p_i = 1$

i หมายถึง ผลลัพธ์ที่ได้จากการทดสอบ โดยที่ i = 0 , 1

p_0 หมายถึง ความน่าจะเป็นที่หน่วยตัวอย่างจะทดสอบแล้วผลออกมาเป็นสิ่งที่ไม่สนใจโดยที่ p_0 มีค่าเท่ากับ $\frac{(2n_{00} + n_{01} + n_{10})}{2n}$

p_1 หมายถึง ความน่าจะเป็นที่หน่วยตัวอย่างจะทดสอบแล้วผลออกมาเป็นสิ่งที่สนใจโดยที่ p_1 มีค่าเท่ากับ $\frac{(2n_{11} + n_{01} + n_{10})}{2n}$

2.1.2 หลักการของ Fleiss ในปี ค.ศ. 1981 (K2) ได้แสดงการหาตัวประมาณสัมประสิทธิ์แคปป้าสำหรับตาราง 2x2 โดยอาศัยสมการ (1) โดยได้กำหนดให้

$$p_{ij} = \frac{n_{ij}}{n} \quad ; \quad i , j = 0 , 1$$

$$p_a = p_{00} + p_{11}$$

$$p_e = p_0 \cdot p_{\cdot 0} + p_1 \cdot p_{\cdot 1}$$

จะได้ตัวประมาณค่าสัมประสิทธิ์แคปป้าอยู่ในรูป

$$K2 = \frac{2(n_{00}n_{11} - n_{01}n_{10})}{n_0 . n_{\cdot 1} + n_1 . n_{\cdot 0}} \tag{6}$$

และ ยังได้แสดงถึงตัวประมาณความแปรปรวนของตัวประมาณที่ได้ อยู่ในรูป

$$\mathrm{var}(K2) = \frac{1}{n(1-p_e)^2} * \left(p_e + p_e^2 - \sum_{i=0}^{1} \left(p_{i\cdot} p_{\cdot i} \left(p_{i\cdot} + p_{\cdot i} \right) \right) \right) \tag{7}$$

กำหนดให้ p_{ij} เป็นตัวประมาณของความน่าจะเป็นที่หน่วยตัวอย่างจะให้ผลลัพธ์วิธีแรกเป็น i และวิธีที่สองเป็น j โดย i และ j มีค่าเท่ากับ 0 หรือ 1 ตามลำดับ (0 หมายถึง เหตุการณ์ที่เกิดเป็นเหตุการณ์ที่ไม่สนใจ ; 1 หมายถึง เหตุการณ์ที่เกิดเป็นเหตุการณ์ที่สนใจ)

2.1.3 การประมาณโดยใช้หลักการภาวะน่าจะเป็นสูงสุด (K3) ได้อาศัยการหาค่าพารามิเตอร์โดยใช้หลักการของภาวะน่าจะเป็นสูงสุด โดยในทางทฤษฎีได้กำหนดให้ข้อมูลมีลักษณะดัง ตารางที่ 2.2

ตารางที่ 2.2 ตารางแสดงลักษณะข้อมูลในทางทฤษฎีของตาราง 2x2 ที่มีความน่าจะเป็นในแต่ละวิธีการทดลองเท่ากัน

		วิธีการทดลอง A		
		0	1	
วิธีการทดลอง B	0	$E(p_i'^2)$	$E(p_i p_i')$	P'
	1	$E(p_i p_i')$	$E(p_i^2)$	P
		P'	P	1.0

P' = 1 - P และ $p_i' = 1 - p_i$

โดยที่ p_i เป็นตัวประมาณความน่าจะเป็นที่ได้ผลการทดลองที่สนใจ

p_i' เป็นตัวประมาณความน่าจะเป็นที่ได้ผลการทดลองที่ไม่สนใจ

i หมายถึงกลุ่มตัวอย่างที่นำมาทดลอง โดย i มีค่าตั้งแต่ 1 ไปจนถึงอนันต์

Bloch และ Kraemer (1989) ได้กำหนดให้ สมการที่ (1) มีค่าเท่ากับ

$$\kappa = \frac{P_a - P_e}{1 - P_e} = \frac{\delta_P^2}{PP'}$$

โดยที่ δ_P^2 = Var (p_i)

P = E (p_i)

ดังนั้นสามารถเขียนสมการที่ (1) ให้อยู่ในรูป

$$\kappa = \frac{E(p_i^2) - P^2}{PP'}$$

จากการหาค่าสัมประสิทธิ์แคปป้าข้างต้นทำให้สามารถหาค่าคาดหวังต่าง ๆ จากตารางที่ 2.2 ได้ (แสดงให้ดูในตารางที่ 2.3)

<u>ตารางที่ 2.3</u> ตารางแสดงค่าคาดหวังของตาราง 2x2 ที่มีความน่าจะเป็นในแต่ละวิธีการทดลองเท่ากัน

		วิธีการทดลอง A		
		0	1	
วิธีการทดลอง B	0	$P'^2 + KPP'$	$PP'(1-K)$	P'
	1	$PP'(1-K)$	$P^2 + KPP'$	P
		P'	P	1.0

นำค่าต่าง ๆ ที่ได้จากตารางที่ 2.3 มาหาตัวประมาณภาวะน่าจะเป็นสูงสุด โดยที่

$$\ln L(P, K | n_{00}, n_{01}, n_{10}, n_{11}) = n_{00}\ln(P'^2 + KPP') + (n_{01}+n_{10})\ln[PP'(1-K)] + n_{11}\ln(P^2 + KPP')$$

ดังนั้น

$$\frac{\partial(\ln L)}{\partial(K)} = \frac{n_{11}PP'}{P^2+KPP'} - \frac{(n_{01}+n_{10})PP'}{PP'(1-K)} + \frac{n_{00}PP'}{P'^2+KPP'} \tag{8}$$

และ

$$\frac{\partial(\ln L)}{\partial(P)} = \frac{n_{11}[2P(1-K)+K]}{P^2+KPP'} + \frac{(n_{01}+n_{10})[(1-2P)(1-K)]}{PP'(1-K)} - \frac{n_{00}[2P'(1-K)+K]}{P'^2+KPP'} \tag{9}$$

กำหนดให้สมการที่ (8) และ (9) เท่ากับ 0 จะได้

$$p = \frac{2n_{11}+n_{01}+n_{10}}{2n}$$

และจะได้ตัวประมาณภาวะน่าจะเป็นสูงสุด

$$K3 = \frac{4(n_{11}n_{00}-n_{01}n_{10})-(n_{10}-n_{01})^2}{(2n_{11}+n_{01}+n_{10})(2n_{00}+n_{01}+n_{10})} \tag{10}$$

ซึ่งสามารถเขียนค่า K3 ที่ได้ให้อยู่ในรูปแบบ $\frac{p_a - p_c}{p_{max} - p_c}$ ที่ซึ่ง

$$p_a = (n_{11} + n_{00}) / n$$

$$p_e^2 = p^2 + p'^2$$

$$p_{max} = 1$$

โดยมีค่าความแปรปรวน

$$\mathrm{var}(K3) = \left\{(1-K3)[(1-K3)(1-2(K3)) + \frac{(K3)(2-(K3))}{2pp'}]\right\} / n \qquad (11)$$

2.1.4 การประมาณโดยใช้ตัวประมาณแจ๊คไนฟ (K4) เป็นการหาตัวประมาณโดยอาศัยตัวประมาณภาวะน่าจะเป็นสูงสุดมาเป็นหลักในการคิด ซึ่งการนำตัวประมาณจากวิธีภาวะน่าจะเป็นสูงสุดมาปรับค่าโดยใช้วิธีแจ๊คไนฟ จะได้

$$K4 = \sum_{i=1}^{4} \frac{n_i}{n} J_i \qquad (12)$$

โดยที่

$$J_1 = \left\{(K3)\left[1-\frac{n}{n-1}\left(\frac{p'}{p}\right)\right]+\frac{p'}{p}\right\} \Big/ \left[1-\frac{1}{n-1}\left(\frac{p'}{p}\right)\right]$$

$$J_2 = J_3 = \left[(K3)\left(1-\frac{1}{4pp'}\right)-\left(\frac{n-1}{n-2}\right)\left(1-\frac{1}{4npp'}\right)\right] \Big/ \left(\frac{n}{2n-1}-\frac{1}{4npp'}\right)$$

$$J_4 = \left\{(K3)\left[1-\frac{n}{n-1}\left(\frac{p}{p'}\right)\right]+\frac{p}{p'}\right\} \Big/ \left[1-\frac{1}{n-1}\left(\frac{p}{p'}\right)\right]$$

และมีตัวประมาณความแปรปรวนของตัวประมาณ

$$\mathrm{var}\ (K4) = \frac{\sum_{i=1}^{4} n_i (J_i - (k4))^2}{n(n-1)} \qquad (13)$$

โดยในที่นี้ได้กำหนดให้ $n_1 = n_{11}$ $n_2 = n_{10}$

$n_3 = n_{01}$ $n_4 = n_{00}$

2.1.5 การประมาณโดยการถ่วงน้ำหนักของแคปป้า (K5) ในการหาตัวประมาณโดยวิธีนี้ใช้หลักการหาโดยใช้หลักของภาวะน่าจะเป็นสูงสุดเช่นเดียวกัน แต่วิธีการหาตัวประมาณวิธีนี้จะมีการถ่วงน้ำหนักโดยมีการประมาณค่าการถ่วงน้ำหนัก

$$r = \frac{n_{11} - n_{10}}{n_{11} - n_{01} - n_{10} + n_{00}} \quad \text{และ} \quad r' = 1 - r$$

Bloch and Kraemer (1989) ได้แสดงวิธีหาตัวประมาณที่ถ่วงน้ำหนัก

$$K5 = \frac{(n_{11}n_{00} - n_{10}n_{01})}{r(n_{11} + n_{10})(n_{10} + n_{00}) + r'(n_{01} + n_{00})(n_{11} + n_{01})} \qquad (14)$$

และได้แสดงตัวประมาณความแปรปรวนของตัวประมาณที่ได้ อยู่ในรูปความแปรปรวนของ k(r) จะมีค่า

$$\text{var}(K5) = \frac{1}{n}\left(\frac{pp'qq'}{(pq'r + p'qr')}\right)v \qquad (15)$$

โดยที่ $p = (n_{11} + n_{10}) / n$; $q = (n_{00} + n_{01}) / n$

$v = 1 + 4u_x u_y \rho - (1 + 3u_x^2 + 3u_y^2)\rho^2 + 2u_x u_y \rho^3$

$$\rho = \frac{(pq'r + p'qr')}{(pp'qq')^{1/2}} K5$$

$u_x = (0.5 - p)/(pp')^{1/2}$

$u_y = (0.5 - q)/(qq')^{1/2}$

โดยกำหนดให้ P คือ ความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจในอิทธิพล A

Q คือ ความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจในอิทธิพล B

$P' = 1-P$, $Q' = 1-Q$

สำหรับในงานวิจัยนี้เป็นกรณีที่ P และ Q มีค่าเท่ากัน ดังนั้นจึงทำให้ตัวประมาณความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจ A และ ตัวประมาณความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจ B มีค่าเท่ากัน แทนตัวประมาณความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจทั้งสองในสมการที่ (15) จะได้ว่า

$$\text{var}(K5) = \text{var}(K3) \qquad (16)$$

ซึ่งเฉพาะในกรณีที่มีความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจในแต่ละวิธีการทดลองเท่ากัน

2.2 วิธีการหาช่วงความเชื่อมั่นที่ใช้ในการวิจัย

การหาช่วงความเชื่อมั่นของตัวประมาณ เพื่อที่จะมาพิจารณาว่าค่า K_0 ที่กำหนดให้มีค่าอยู่นอกช่วงความเชื่อมั่นของตัวประมาณนั้นกี่เปอร์เซ็นต์ โดยที่ถ้าช่วงความเชื่อมั่นของตัวประมาณใดมีค่าเข้าใกล้ 0.05 (ระดับนัยสำคัญที่กำหนด) มากที่สุด ก็จะแสดงว่าตัวประมาณตัวนั้นมีช่วงความเชื่อมั่นที่เหมาะสมที่สุด โดย Fleiss (1971) และ Hale and Fleiss (1993) ได้กล่าวถึง การหาช่วงความเชื่อมั่น จากการประมาณการแจกแจงของตัวประมาณสัมประสิทธิ์แคปป้าที่ได้ว่า มีการแจกแจงแบบปกติและมีค่าเฉลี่ยเท่ากับ K_0 และหลังจากการนำตัวประมาณที่ได้ลบกับค่าคงที่ที่กำหนดขึ้น หารด้วยค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Error, SE) ของตัวประมาณก็จะได้ การแจกแจงแบบปกติมีค่าเฉลี่ยเป็น 0 และ ความแปรปรวนเท่ากับ 1 ซึ่งก็จะสามารถเปิดตารางการแจกแจงปกติ หรือสามารถเขียนอยู่ในรูป

$$Z_{\frac{\alpha}{2}} = \frac{K - K_0}{\sqrt{\mathrm{var}(K)}} \tag{17}$$

ดังนั้นจะได้ช่วงความเชื่อมั่นของของตัวประมาณค่าสัมประสิทธิ์แคปป้าอยู่ระหว่าง ($K - Z_{\alpha/2}SE(K)$, $K + Z_{\alpha/2}SE(K)$)

2.3 สถิติที่ใช้ทดสอบสมมติฐานในงานวิจัย

ในงานวิจัยนี้มีการทดสอบสมมติฐานอยู่ 2 ลักษณะ ได้แก่ การทดสอบหาความคลาดเคลื่อนประเภทที่ 1 และ การทดสอบหาความคลาดเคลื่อนประเภทที่ 2 หรือ การหาอำนาจในการทดสอบ สำหรับเกณฑ์การพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 จะใช้หลักเกณฑ์ของ Bradley (ศุชสาคร พงษ์ประดิษฐ , 2531: 20-21) ได้กำหนดให้ การทดสอบจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ถ้าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่ภายในขอบเขต [0.025 , 0.075] เมื่อทดสอบที่ระดับนัยสำคัญที่กำหนดไว้ในขอบเขตของงานวิจัย (α=0.05) และถ้าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่นอกขอบเขตที่ระบุ จะถือว่าตัวสถิติทดสอบนั้นไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ซึ่งมีอยู่สองกรณีได้แก่ กรณีที่ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่ามากกว่าขอบเขตบนของเกณฑ์ที่พิจารณา จะถือว่าการทดสอบนั้นมีความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มากกว่าระดับนัยสำคัญที่กำหนด และอีกกรณีคือ กรณีที่ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่าน้อยกว่าขอบเขตบนของเกณฑ์ที่พิจารณา จะถือว่าการทดสอบนั้นมีความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าระดับนัยสำคัญที่กำหนด

สำหรับเกณฑ์ที่ใช้ในการพิจารณาความแตกต่างของอำนาจที่ใช้ในการทดสอบ (ศุขสาคร พงษ์ประดิษฐ, 2531: 21) กำหนดให้ กรณีที่อำนาจการทดสอบที่ได้มีความแตกต่างน้อยกว่าหรือเท่ากับ 0.036 จะถือว่าอำนาจการทดสอบของทั้ง 2 วิธี ไม่แตกต่างกันอย่างมีนัยสำคัญ และในกรณีที่อำนาจการทดสอบที่ได้มีความแตกต่างมากกว่า 0.036 จะถือว่าอำนาจการทดสอบของทั้ง 2 วิธี แตกต่างกันอย่างมีนัยสำคัญ ณ ระดับนัยสำคัญที่ 0.05

โดยในงานวิจัยแบ่งวิธีการที่ใช้ในการทดสอบความเป็นเอกพันธ์ของค่าสัมประสิทธิ์แคปป้าในประชากร 2 ชุดที่มีลักษณะของการเก็บข้อมูลที่เหมือนกันคือ มีวิธีทดลองสองวิธีเช่นกัน และมีผลลัพธ์ที่ได้สองลักษณะ สำหรับสมมุติฐานหลักในการทดสอบความเป็นเอกพันธ์ของค่าสัมประสิทธิ์แคปป้าในประชากร 2 ชุดคือ Ho : $\kappa_1 = \kappa_2$ โดยจะพิจารณาวิธีทดสอบทั้งสิ้น 3 วิธี ได้แก่

2.3.1 วิธีทดสอบภาวะสารูปสนิทดีโดยใช้ค่าสัมประสิทธิ์แคปป้า (Donner and Eliasziw ,1989) เป็นวิธีทดสอบที่ใช้พื้นฐานของ วิธีไคสแควร์เป็นหลัก โดยใช้ค่าสัมประสิทธิ์แคปป้าเป็นเกณฑ์ และมีสมมุติฐานหลักคือ Ho : $\kappa_1 = \kappa_2 = \ldots = \kappa_k$ โดยที่ k คือ จำนวนกลุ่มข้อมูลที่นำมาทดสอบ และกำหนดให้ตัวประมาณความน่าจะเป็นต่าง ๆ อยู่ในลักษณะ

$p_{1h}(\kappa_h) = pr(X_{i1h} = 1, X_{i2h} = 1) = p_h^2 + p_h(1- p_h)\,\kappa_h$

$p_{2h}(\kappa_h) = pr(X_{i1h} = 1, X_{i2h} = 0 \text{ or } X_{i1h} = 0, X_{i2h} = 1) = 2\,p_h(1- p_h)(1-\kappa_h)$

$p_{3h}(\kappa_h) = pr(X_{i1h} = 0, X_{i2h} = 0) = (1- p_h)^2 + p_h(1- p_h)\,\kappa_h$

ตัวประมาณที่ได้มีลักษณะเช่นเดียวกันกับตัวประมาณที่ได้จากตารางที่ 2.3 แต่ต่างกันที่จะมีชุดของตัวอย่างเพิ่มขึ้นมา โดยได้กำหนดให้

$p_{1h}(\kappa_h)$ คือ ตัวประมาณความน่าจะเป็นที่จะได้ผลที่สนใจทั้งสองการทดลอง

$p_{2h}(\kappa_h)$ คือ ตัวประมาณความน่าจะเป็นที่จะได้ผลที่สนใจหนึ่งการทดลอง และผลที่ไม่สนใจในอีกหนึ่งการทดลอง

$p_{3h}(\kappa_h)$ คือ ตัวประมาณความน่าจะเป็นที่จะได้ผลที่ไม่สนใจทั้งสองการทดลอง

h หมายถึง ลำดับที่ของกลุ่มข้อมูลที่นำมาทดสอบ โดยที่ h มีค่าตั้งแต่ 1 ไปจนถึง k

k หมายถึง จำนวนกลุ่มทั้งหมดที่นำมาทดสอบ

สำหรับการกำหนดให้ตัวประมาณสัมประสิทธิ์แคปป้าโดยใช้ตัวประมาณภาวะน่าจะเป็นสูงสุดเป็นเกณฑ์ในการตัวประมาณของแต่ละกลุ่มที่นำมาทดสอบ นอกจากนี้ยังได้กำหนดให้ตัวประมาณของสัมประสิทธิ์แคปป้าเฉลี่ย (κ_{Avg})อยู่ในรูป

$$\kappa_{Avg} = \frac{\left(\sum_{h=1}^{k} N_h p_h(1-p_h)(K3_h)\right)}{\left(\sum_{h=1}^{k} N_h p_h(1-p_h)\right)} = 1 - \frac{n_{2\bullet}}{\left(2\sum_{h=1}^{k} N_h p_h(1-p_h)\right)} \tag{18}$$

โดยกำหนดให้

$K3_h$ หมายถึง ตัวประมาณของค่าสัมประสิทธิ์แคปป้าในกลุ่มที่ h

$$n_{2\bullet} = \sum_{h=1}^{k}(n_{01h}+n_{10h})$$

ซึ่งการนำตัวประมาณค่าโดยใช้หลักการภาวะน่าจะเป็นสูงสุด (K3) มาเป็นเกณฑ์ในการคิดเนื่องจากคาดว่าตัวประมาณค่าโดยใช้หลักการภาวะน่าจะเป็นสูงสุดจะเป็นตัวประมาณที่มีค่าความคลาดเคลื่อนกำลังสองต่ำสุด สำหรับกลุ่มของข้อมูลที่นำมาทดสอบสมมุติฐานสามารถเขียนได้ดังตารางที่ 2.4

ตารางที่ 2.4 ตารางแสดงลักษณะของกลุ่มข้อมูลที่ใช้ทดสอบสมมุติฐานที่กำหนด

ระดับชั้นที่	ผลของการทดสอบ	ความถี่ของข้อมูล				รวม
		กลุ่มที่ 1	กลุ่มที่ 2	....	กลุ่มที่ g	
1	(1,1)	n11	n21	....	n1g	n1.
2	(1,0)หรือ(0,1)	n21	n22	....	n2g	n2.
3	(0,0)	n31	n23	....	n3g	n3.
	รวม	N_1	N_2	....	N_g	N

และตัวสถิติที่ใช้

$$\chi_G^2 = \sum_{h=1}^{k}\sum_{l=1}^{3}\frac{\left[n_{lh}-N_h p_{lh}(K_{Avg})\right]^2}{N_h p_{lh}(K_{Avg})} \quad (19)$$

ภายใต้สมมุติฐานหลักที่กำหนด โดย χ_G^2 ประมาณด้วยการแจกแจงแบบไคสแควร์ ที่มีองศาของความอิสระ (Degrees of Freedom) เท่ากับ k-1 ซึ่งถ้าค่าที่คำนวณได้จากสมการที่ (17) มีค่าน้อยกว่าหรือเท่ากับค่าไคสแควร์ที่เปิดจากตารางก็จะยอมรับสมมุติฐานหลัก (Accept Ho) ในทางกลับกันถ้ามีค่าที่คำนวณได้มากกว่าก็จะปฏิเสธสมมุติฐานหลัก

สำหรับในขอบเขตของงานวิจัยนี้จะพิจารณาการทดสอบกรณีที่เปรียบเทียบ 2 กลุ่มข้อมูล ดังนั้น k = 2 และตัวสถิติที่ใช้ทดสอบ

$$\chi_G^2 = \sum_{h=1}^{2}\sum_{l=1}^{3}\frac{\left[n_{lh}-N_h p_{lh}(K_{Avg})\right]^2}{N_h p_{lh}(K_{Avg})} \quad (20)$$

ภายใต้สมมุติฐานหลัก Ho : $K_1 = K_2$ ด้วยการแจกแจงแบบไคสแควร์ มีองศาของความอิสระเท่ากับ 1

2.3.2 **วิธีที่ขึ้นอยู่กับความแปรปรวนของขนาดตัวอย่างที่ใหญ่**

หลักการสำหรับวิธีนี้คือ ต้องหาตัวประมาณความแปรปรวนของค่าสัมประสิทธิ์แคปป้าที่ได้ในแต่ละกลุ่มตัวอย่าง โดยได้กำหนดค่าความแปรปรวนเมื่อขนาดตัวอย่างมีค่ามาก

$$\mathrm{var}(K_h) = \{(1-K3_h)[(1-K3_h)(1-2(K3_h)) + \frac{(K3_h)(2-(K3_h))}{2p_h p_h'}]\} / n_h$$

สำหรับตัวประมาณความแปรปรวนที่ได้จะมีลักษณะเช่นเดียวกับตัวประมาณความแปรปรวนของ K3 ในสมการที่ (11) เนื่องจากใช้ตัวประมาณสัมประสิทธิ์แคปป้า K3 เป็นตัวประมาณในแต่ละกลุ่มข้อมูล และเพื่อที่จะให้เป็นตัวประมาณที่ใช้ในกรณีเดียวกันกับวิธีทดสอบวิธีแรก หลังจากได้ตัวประมาณความแปรปรวนในแต่ละกลุ่มก็นำตัวประมาณความแปรปรวนที่ได้ไปถ่วงน้ำหนักโดยให้

$w_h = 1/\mathrm{var}(K_h)$ โดยใช้ตัวสถิติ

$$\chi_v^2 = \sum_{h=1}^{k} w_h (K_h - K_{Avg})^2 \qquad (21)$$

ภายใต้สมมุติฐานหลักที่กำหนด โดย χ_v^2 ประมาณด้วยการแจกแจงแบบไคสแควร์ ที่มีองศาของความอิสระ (Degrees of Freedom) เท่ากับ k-1 ซึ่งถ้าค่าที่คำนวณได้จากสมการที่ (17) มีค่าน้อยกว่าหรือเท่ากับค่าไคสแควร์ที่เปิดจากตารางก็จะยอมรับสมมุติฐานหลัก ในทางกลับกันถ้ามีค่าที่คำนวณได้มากกว่าก็จะปฏิเสธสมมุติฐานหลัก เช่นเดียวกับวิธีแรก และตัวประมาณสัมประสิทธิ์แคปป้าเฉลี่ย

$$K_{Avg} = \frac{\sum_{h=1}^{k} w_h K3_h}{\sum_{h=1}^{k} K3_h} \qquad (22)$$

สำหรับในขอบเขตของงานวิจัยนี้จะพิจารณาการทดสอบกรณีที่เปรียบเทียบ 2 กลุ่มข้อมูล ดังนั้น k = 2 และสามารถประยุกต์ตัวสถิติที่ใช้ทดสอบให้อยู่ในรูป

$$Z_V = \frac{(K3_1 - K3_2)}{\sqrt{\mathrm{var}(K3_1) + \mathrm{var}(K3_2)}} \qquad (23)$$

ภายใต้สมมุติฐานที่กำหนด โดย Z_V เป็นการแจกแจงแบบปกติมาตรฐาน

2.3.3 วิธีทดสอบภาวะสารูปสนิทดีโดยใช้ความน่าจะเป็น

ในที่นี้จะเป็นการทดสอบลักษณะความเหมือนกันของข้อมูล โดยใช้ไคสแควร์ ซึ่งเป็นวิธีปกติที่ใช้เพื่อที่จะทดสอบว่า ข้อมูล k กลุ่มมีลักษณะการกระจายแตกต่างกันหรือไม่ โดยใช้ตัวสถิติ

$$\chi_H^2 = \sum_{h=1}^{k}\sum_{i=0}^{1}\sum_{j=0}^{1}\frac{\left[n_{ijh} - E(n_{ijh})\right]^2}{E(n_{ijh})} \qquad (24)$$

โดยที่

i = 0 , 1

j = 0 , 1

h = 1 , 2 , ... , k

k หมายถึง จำนวนกลุ่มทั้งหมดที่นำมาทดสอบ

$$E(n_{ijh}) = \frac{\sum_{h=1}^{k} n_{ijh}}{h}$$

ทดสอบภายใต้สมมุติฐานที่กำหนด โดย χ_H^2 ประมาณด้วยการแจกแจงแบบไคสแควร์ ที่มีองศาความเป็นอิสระเท่ากับ ((2x2)-1)(k-1) หรือเท่ากับ 3(k-1) ซึ่งถ้าค่าที่คำนวณได้จากสมการที่ (24) มีค่าน้อยกว่าหรือเท่ากับค่าไคสแควร์ที่เปิดจากตารางก็จะยอมรับสมมุติฐานหลัก ในทางกลับกันถ้ามีค่าที่คำนวณได้มากกว่าก็จะปฏิเสธสมมุติฐานหลัก

สำหรับในขอบเขตของงานวิจัยนี้จะพิจารณาการทดสอบกรณีที่เปรียบเทียบ 2 กลุ่มข้อมูล ดังนั้น k = 2 และตัวสถิติที่ใช้ทดสอบ

$$\chi_H^2 = \sum_{h=1}^{2}\sum_{i=0}^{1}\sum_{j=0}^{1}\frac{\left[n_{ijh} - E(n_{ijh})\right]^2}{E(n_{ijh})} \qquad (25)$$

ภายใต้สมมุติฐานหลัก Ho : $\kappa_1 = \kappa_2$ ด้วยการแจกแจงแบบไคสแควร์ มีองศาของความอิสระเท่ากับ 3

2.4 ข้อจำกัดที่ใช้ในงานวิจัย

จากการนำเสนอตัวประมาณค่าต่าง ๆ และวิธีที่ใช้ในการทดสอบจะสังเกตเห็นว่า มีข้อจำกัดในการหาค่าต่าง ๆ เนื่องจากในแต่ละสมการส่วนที่ได้มีโอกาสที่จะเท่ากับ 0 ได้ดังนั้นในหัวข้อนี้จึงพิจารณาถึงข้อจำกัดต่าง ๆ โดยแบ่งเป็น

2.4.1 *ข้อจำกัดของพารามิเตอร์ที่ใช้ในการวิจัย*

จากตัวประมาณค่าต่างๆ ที่ได้เสนอในข้างต้นจะพบว่าตัวประมาณต่าง ๆ ไม่สามารถที่จะคำนวณค่าได้ในบางกรณี โดยจะแสดงให้เห็นว่ามีข้อจำกัดใดบ้างที่ทำให้ตัวประมาณ และตัวประมาณของความแปรปรวนไม่สามารถหาค่าได้ ภายใต้ขอบเขตงานวิจัย

2.4.1.1 ตัวประมาณสัมประสิทธิ์แคปป้าโดยหลักการของ Fleiss (1971)

พิจารณาตัวประมาณสัมประสิทธิ์แคปป้า และตัวประมาณความแปรปรวน สามารถหาค่าได้เมื่อ p_0 และ p_1 ต้องไม่เท่ากับ 0 หรือ $\frac{2n_{00}+n_{01}+n_{10}}{2n}$ และ $\frac{2n_{11}+n_{01}+n_{10}}{2n}$ ต้องไม่เท่ากับ 0 ดังนั้นจะได้ว่า n_{00} และ n_{11} ต้องไม่เท่ากับ ขนาดตัวอย่างทั้งหมด (n)

2.4.1.2 ตัวประมาณสัมประสิทธิ์แคปป้าโดยใช้หลักการของ Fleiss (1981)

พิจารณาตัวประมาณสัมประสิทธิ์แคปป้า และตัวประมาณความแปรปรวน สามารถหาค่าได้เมื่อ $(n_{00}+n_{01})(n_{01}+n_{11})+(n_{10}+n_{11})(n_{00}+n_{01})$ ต้องไม่เท่ากับ 0 หรือกล่าวได้ว่า $(n_{00}+n_{01})(n_{01}+n_{11})$ และ $(n_{10}+n_{11})(n_{00}+n_{01})$ ต้องไม่เท่ากับ 0 เนื่องจากทั้งสองพจน์มีค่ามากกว่าเท่ากับ 0

2.4.1.3 ตัวประมาณสัมประสิทธิ์แคปป้าโดยใช้หลักการภาวะน่าจะเป็นสูงสุด

พิจารณาตัวประมาณสัมประสิทธิ์แคปป้า และตัวประมาณความแปรปรวน สามารถหาค่าได้เมื่อ p(1-p) ต้องไม่เท่ากับ 0 หรือ $\frac{2n_{00}+n_{01}+n_{10}}{2n}$ และ $\frac{2n_{11}+n_{01}+n_{10}}{2n}$ ต้องไม่เท่ากับ 0 ดังนั้นจะได้ว่า n_{00} และ n_{11} ต้องไม่เท่ากับ ขนาดตัวอย่างทั้งหมด (n) เช่นเดียวกับตัวประมาณสัมประสิทธิ์แคปป้าตัวแรก

2.4.1.4 ตัวประมาณสัมประสิทธิ์แคปป้า โดยใช้ตัวประมาณแจ๊คไนฟ

พิจารณาค่า J_i และตัวประมาณความแปรปรวน สามารถหาค่าได้เมื่อสามารถหาตัวประมาณสัมประสิทธิ์แคปป้าโดยใช้ตัวประมาณภาวะน่าจะเป็นสูงสุดได้ ดังนั้น n_{00} และ n_{11} ต้องไม่เท่ากับ ขนาดตัวอย่างทั้งหมด (n) นอกจากนี้ยังต้องพิจารณา J_i แต่ละตัวด้วย

พิจารณาค่า J_1 สามารถหาค่าได้เมื่อ

$\frac{1}{n-1}\left(\frac{p'}{p}\right)$ ต้องไม่เท่ากับ 1

$\frac{1-p}{p}$ ต้องไม่เท่ากับ $n-1$

$1-p$ ต้องไม่เท่ากับ $np-p$

np ต้องไม่เท่ากับ 1

ดังนั้น p ต้องไม่เท่ากับ $\frac{1}{n}$

พิจารณาค่า J_2 และ J_3 จะสามารถหาค่าได้เมื่อ

$\frac{1}{4pp'}$ ต้องไม่เท่ากับ $\frac{n}{2n-1}$

$4p(1-p)$ ต้องไม่เท่ากับ $\frac{2n-1}{n}$ โดยที่ $p(1-p) \neq 0$ (ข้อจำกัดของ K3)

$p-p^2$ ต้องไม่เท่ากับ $\frac{2n-1}{4n}$

$p^2-p+\frac{2n-1}{4n}$ ต้องไม่เท่ากับ 0

ดังนั้น p ต้องไม่เท่ากับ $\frac{1 \pm \sqrt{\frac{2n+1}{4n}}}{2}$

พิจารณาค่า J_4 จะสามารถหาค่าได้เมื่อ

$\frac{1}{n-1}\left(\frac{p}{p'}\right)$ ต้องไม่เท่ากับ 1

p ต้องไม่เท่ากับ $(n-1)(1-p)$

$p+pn-p$ ต้องไม่เท่ากับ $n-1$

ดังนั้น p ต้องไม่เท่ากับ $\frac{n-1}{n}$

2.4.1.5 ตัวประมาณสัมประสิทธิ์แคปป้า โดยการให้น้ำหนัก

พิจารณาตัวประมาณสัมประสิทธิ์แคปป้า สามารถหาค่าได้เมื่อน้ำหนักที่ให้ (r) มีค่าอยู่ระหว่าง 0 ถึง 1 และเมื่อพิจารณาตัวประมาณความแปรปรวน สามารถหาค่าได้เมื่อสามารถหาตัวประมาณความแปรปรวนของสัมประสิทธิ์แคปป้าโดยใช้ตัวประมาณภาวะน่าจะเป็นสูงสุดได้ ดังนั้น n_{00} และ n_{11} ต้องไม่เท่ากับ ขนาดตัวอย่างทั้งหมด (n)

2.4.2 *ข้อจำกัดของวิธีที่ใช้ในการทดสอบความเป็นเอกพันธ์ของค่าสัมประสิทธิ์แคปป้า*

ข้อจำกัดของวิธีที่ใช้ในการทดสอบความเป็นเอกพันธ์ของค่าสัมประสิทธิ์แคปป้า ก็จะแบ่งออกเป็น 3 กรณีตามวิธีที่ใช้ในการทดสอบความเป็นเอกพันธ์ของค่าสัมประสิทธิ์แคปป้า ได้แก่

2.4.2.1 วิธีทดสอบภาวะสารูปสนิทดี โดยใช้ค่าสัมประสิทธิ์แคปป้า

สามารถใช้วิธีทดสอบนี้ได้เมื่อ หาค่าหาตัวประมาณสัมประสิทธิ์แคปป้าโดยใช้ตัวประมาณภาวะน่าจะเป็นสูงสุดในแต่ละกลุ่มตัวอย่างได้ ดังนั้น n_{00h} และ n_{11h} ต้องไม่เท่ากับ ขนาดตัวอย่างทั้งหมด (n_h) และ

$p_{2h}(K3)$ ต้องไม่เท่ากับ 0

$2p_h(1-p_h)(1-K_{Avg})$ ต้องไม่เท่ากับ 0

ดังนั้น K_{Avg} ต้องไม่เท่ากับ 1 และ $p_h(1-p_h)$ ต้องไม่เท่ากับ 0 (ข้อจำกัดของ K3)

2.4.2.2 วิธีทดสอบที่ขึ้นอยู่กับความแปรปรวนของขนาดตัวอย่างที่ใหญ่

ใช้วิธีทดสอบนี้ได้เมื่อ สามารถหาค่าตัวประมาณสัมประสิทธิ์แคปป้าโดยใช้ตัวประมาณภาวะน่าจะเป็นสูงสุดในแต่ละกลุ่มตัวอย่างได้ ดังนั้น n_{00h} และ n_{11h} ต้องไม่เท่ากับ ขนาดตัวอย่างทั้งหมด (n_h) และ ตัวประมาณสัมประสิทธิ์แคปป้าในแต่ละกลุ่ม ($K3_h$) ต้องไม่เท่ากับ 1 เนื่องจากจะไม่สามารถหาตัวประมาณความแปรปรวนได้

2.4.2.3 วิธีทดสอบภาวะสารูปสนิทดี โดยใช้ความน่าจะเป็น

สามารถใช้วิธีทดสอบนี้ได้เมื่อ

$E(n_{ijh})$ ต้องไม่เท่ากับ 0

ดังนั้น $\sum_{h=1}^{k} n_{ijk}$ ต้องไม่เท่ากับ 0