

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

จากการเปรียบเทียบ วิธีการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของ ข้อมูลแบบจับคู่ ทั้ง 4 วิธี ซึ่งได้แก่

1. วิธีการประมาณของ Wald (WA)
2. วิธีการประมาณของ Newcombe (NH)
3. วิธีการประมาณของ May และ Johnson (MJ)
4. วิธีการประมาณของ Zhou และ Qin (ZQ)

ผู้วิจัยได้ทำการศึกษาเปรียบเทียบในแต่ละสถานการณ์ทดลอง ประกอบด้วย

- 1) กำหนดระดับความเชื่อมั่น 3 ระดับคือ 90%, 95% และ 99%
- 2) กำหนดให้ขนาดตัวอย่างที่ใช้ $n = n_1 = n_2$ มีค่าตั้งแต่ 10,20,30,40,50,60,70,80
- 3) กำหนดค่าความน่าจะเป็นร่วมที่ทำให้ได้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างประชากร กลุ่มที่ 1 และ 2 มีค่าอยู่ใน ระดับต่ำ $0 \leq r < 0.4$, ระดับกลาง $0.4 \leq r \leq 0.6$, ระดับสูง $0.6 < r \leq 1$

4) ในแต่ละระดับความเชื่อมั่น และแต่ละระดับขนาดตัวอย่างของสองประชากร จะ กำหนดค่าผลต่างของสัดส่วนประชากร ($p_1 - p_0$) ในแต่ละระดับของค่าสัมประสิทธิ์สหสัมพันธ์ กล่าวคือ

ค่าสัมประสิทธิ์สหสัมพันธ์มีระดับต่ำ $0 \leq r < 0.4$ จะกำหนด $p_1 - p_0$ มีค่าตั้งแต่ 0-0.8

ค่าสัมประสิทธิ์สหสัมพันธ์มีระดับกลาง $0.4 \leq r \leq 0.6$ จะกำหนด $p_1 - p_0$ มีค่าตั้งแต่ 0-0.3

ค่าสัมประสิทธิ์สหสัมพันธ์มีระดับสูง $0.6 < r \leq 1$ จะกำหนด $p_1 - p_0$ มีค่าตั้งแต่ 0-0.2

โดยที่ $p_1 - p_0$ มีค่าเพิ่มทีละ 0.1 และกำหนดให้ $p_1 \geq p_0$

ในการเปรียบเทียบวิธีการประมาณผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่ในแต่ละสถานการณ์ทดลอง จะทำการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลอง และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ได้ถ้าวิธีการใด ให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองต่ำสุด และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นสั้นที่สุด จะถือว่าวิธีการประมาณ

ดังกล่าวเป็นวิธีการประมาณที่เหมาะสมที่สุดสำหรับวิธีการประมาณผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่ในแต่ละสถานการณ์นั้นๆ

5.1 สรุปผลการวิจัย

ในการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นของวิธีการประมาณทั้ง 4 วิธี ที่ระดับความเชื่อมั่น 3 ระดับ คือ 90%, 95% และ 99% ระดับค่าสัมประสิทธิ์สหสัมพันธ์ 3 ระดับ โดยพิจารณาตามขนาดของตัวอย่างและผลต่างระหว่างค่าสัดส่วนของสองประชากร จากวิธีการประมาณที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองมีค่าไม่ต่ำกว่าค่าระดับความเชื่อมั่นที่กำหนดและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุดจะถือว่าวิธีการนั้นเหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ จากผลการวิจัยในตารางที่ 4.1.1 – 4.1.48, 4.2.1 – 4.2.48 ได้ผลสรุปดังนี้

ที่ระดับสัมประสิทธิ์ระดับ ต่ำ $0 \leq r < 0.4$

กรณีขนาดตัวอย่างทั้งสองประชากรน้อยกว่า 30 วิธี NH ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุดในทุกสถานการณ์ ยกเว้นระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0 และน้อยกว่า 0.1 วิธี MJ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุด

กรณีที่ขนาดตัวอย่างตั้งแต่ 30 และน้อยกว่า 70 วิธี ZQ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุดในทุกสถานการณ์ ยกเว้นระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0 และน้อยกว่า 0.3 วิธี MJ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุด

กรณีที่ขนาดตัวอย่างมากกว่า 70 เมื่อระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0 และน้อยกว่า 0.3 วิธี WA, MJ, ZQ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำใกล้เคียงกันในทุกสถานการณ์ ยกเว้นระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0.3 และน้อยกว่า 1 วิธี WA, NH, ZQ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำใกล้เคียงกันในทุกสถานการณ์

ที่ระดับสัมประสิทธิ์ระดับ กลาง $0.4 \leq r \leq 0.6$

กรณีขนาดตัวอย่างทั้งสองประชากรน้อยกว่า 30 วิธี NH ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุดในทุกสถานการณ์ ยกเว้นระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0 และน้อยกว่า 0.1 วิธี MJ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุด

กรณีที่ขนาดตัวอย่างตั้งแต่ 30 และน้อยกว่า 70 วิธี MJ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุดในทุกสถานการณ์ ยกเว้นระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0.3 และน้อยกว่า 0.4 วิธี ZQ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุด

กรณีที่ขนาดตัวอย่างมากกว่า 70 วิธี WA, MJ, ZQ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำใกล้เคียงกันในทุกสถานการณ์ ยกเว้นระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0.3 และน้อยกว่า 0.4 วิธี WA, NH, ZQ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำใกล้เคียงกันในทุกสถานการณ์

ที่ระดับสัมประสิทธิ์ระดับ สูง $0.6 < r < 1$

กรณีขนาดตัวอย่างทั้งสองประชากรน้อยกว่า 30 วิธี NH ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุดในทุกสถานการณ์ ยกเว้นระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าตั้งแต่ 0 และน้อยกว่า 0.1 วิธี MJ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุด

กรณีที่ขนาดตัวอย่างตั้งแต่ 30 และน้อยกว่า 70 วิธี MJ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุดในทุกสถานการณ์

กรณีที่ขนาดตัวอย่างมากกว่า 70 วิธี WA, MJ, ZQ ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำใกล้เคียงกันในทุกสถานการณ์

5.2 ปัจจัยที่มีอิทธิพลต่อความยาวเฉลี่ยของช่วงความเชื่อมั่นสำหรับค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่

จากผลสรุปที่ได้จากหัวข้อ 5.1 ทำให้ทราบปัจจัยต่างๆที่มีผลต่อค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นสำหรับผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่ ทั้ง 4 วิธี คือ

1. ค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ถ้าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด มีค่าเพิ่มขึ้น จะมีผลให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นเพิ่มขึ้นเช่นกัน เพราะค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นถูกถ่วงน้ำหนักด้วยค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

2. ขนาดตัวอย่างของสองประชากร ถ้าขนาดตัวอย่างของทั้งสองประชากรมีค่าเพิ่มขึ้น จะมีผลให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นลดลง เพราะถ้าขนาดตัวอย่างเพิ่มขึ้นจะทำให้ความแปรปรวนลดลง

3. ค่าสัมประสิทธิ์สหสัมพันธ์ ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างสองประชากรมีค่าเพิ่มขึ้นในทางเดียวกัน จะมีผลให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นลดลง เพราะถ้าค่า

สัมประสิทธิ์สหสัมพันธ์ระหว่างสองประชากรมีค่าเพิ่มขึ้นในทางเดียวกัน จะทำให้ความแปรปรวนลดลง

4. ระดับผลต่างระหว่างค่าสัดส่วนของประชากร ถ้าระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าแปรผกผันกับค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง กล่าวคือ ถ้าระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่าน้อยจะมีผลให้ค่าอำนาจการทดสอบมีค่าน้อย หรือค่าความผิดพลาดประเภทที่ 1 มีค่าน้อย ดังนั้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองจะมากขึ้น แต่ถ้าระดับผลต่างระหว่างค่าสัดส่วนของประชากรมีค่ามากจะมีผลให้ค่าอำนาจการทดสอบมีค่ามาก หรือค่าความผิดพลาดประเภทที่ 1 มีค่ามาก ดังนั้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองจะน้อยลง

5.3 ข้อเสนอแนะ

ผลการวิจัยครั้งนี้มีข้อเสนอแนะ 2 ประการคือ

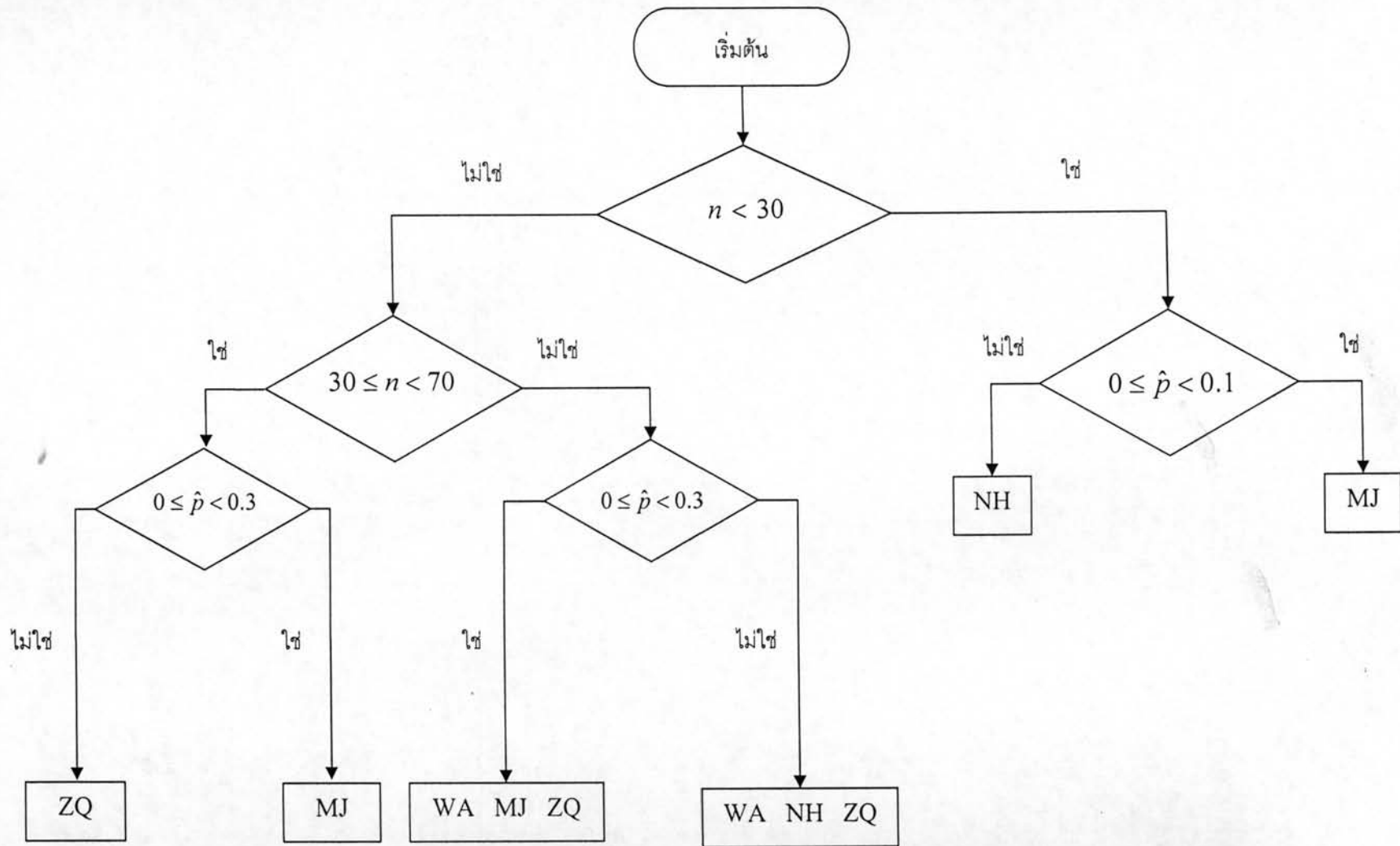
5.3.1 การนำไปใช้ประโยชน์

เพื่อเป็นแนวทางในการเลือกใช้วิธีการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่ เนื่องจากในการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่ เราจะทำการเก็บรวบรวมข้อมูลตัวอย่างขนาด $n = n_1 = n_2$ จากประชากรที่ 1 และ 2 ตามลำดับ ทำให้เราสามารถหาค่าสัดส่วนตัวอย่างของสองประชากรได้ (\hat{p}_0, \hat{p}_1) ได้เราจึงใช้ค่าผลต่างระหว่างค่าสัดส่วนตัวอย่างของสองประชากร ($\hat{p}_1 - \hat{p}_0$) แทนค่าผลต่างระหว่างค่าสัดส่วนของสองประชากร และค่าสัมประสิทธิ์สหสัมพันธ์ระดับต่างๆ (r) โดยที่

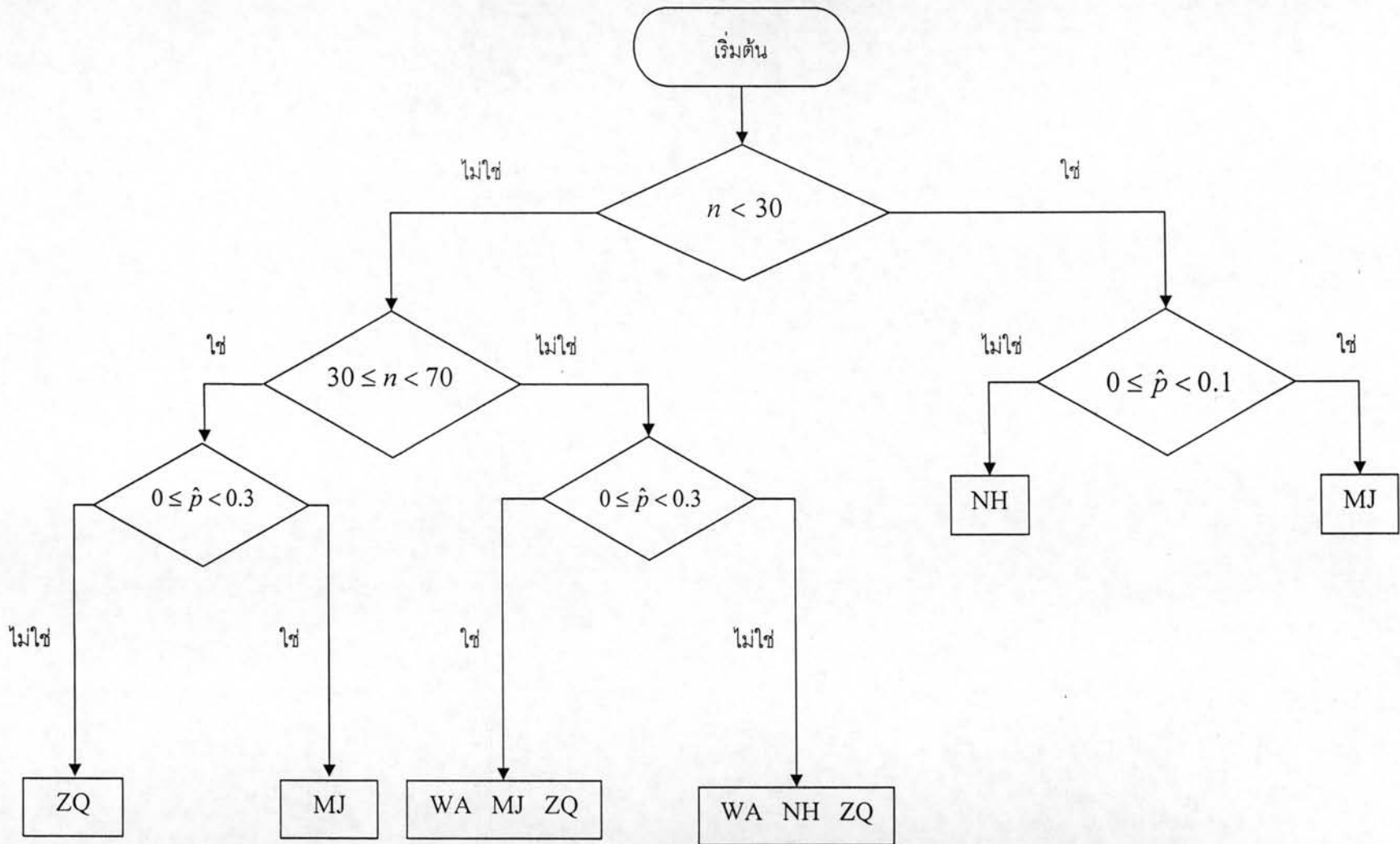
$$\hat{p}_1 = \frac{Y_1}{n}, \hat{p}_0 = \frac{Y_0}{n}, \hat{p}_{11} = \frac{Y_{11}}{n}, \hat{p}_1 \geq \hat{p}_0$$

$$r = \frac{\hat{p}_{11} - \hat{p}_0 \hat{p}_1}{\sqrt{\hat{p}_0(1 - \hat{p}_0)\hat{p}_1(1 - \hat{p}_1)}} \text{ และ } c_0 \text{ คือระดับความเชื่อมั่นที่กำหนด(0.90, 0.95 และ 0.99)}$$

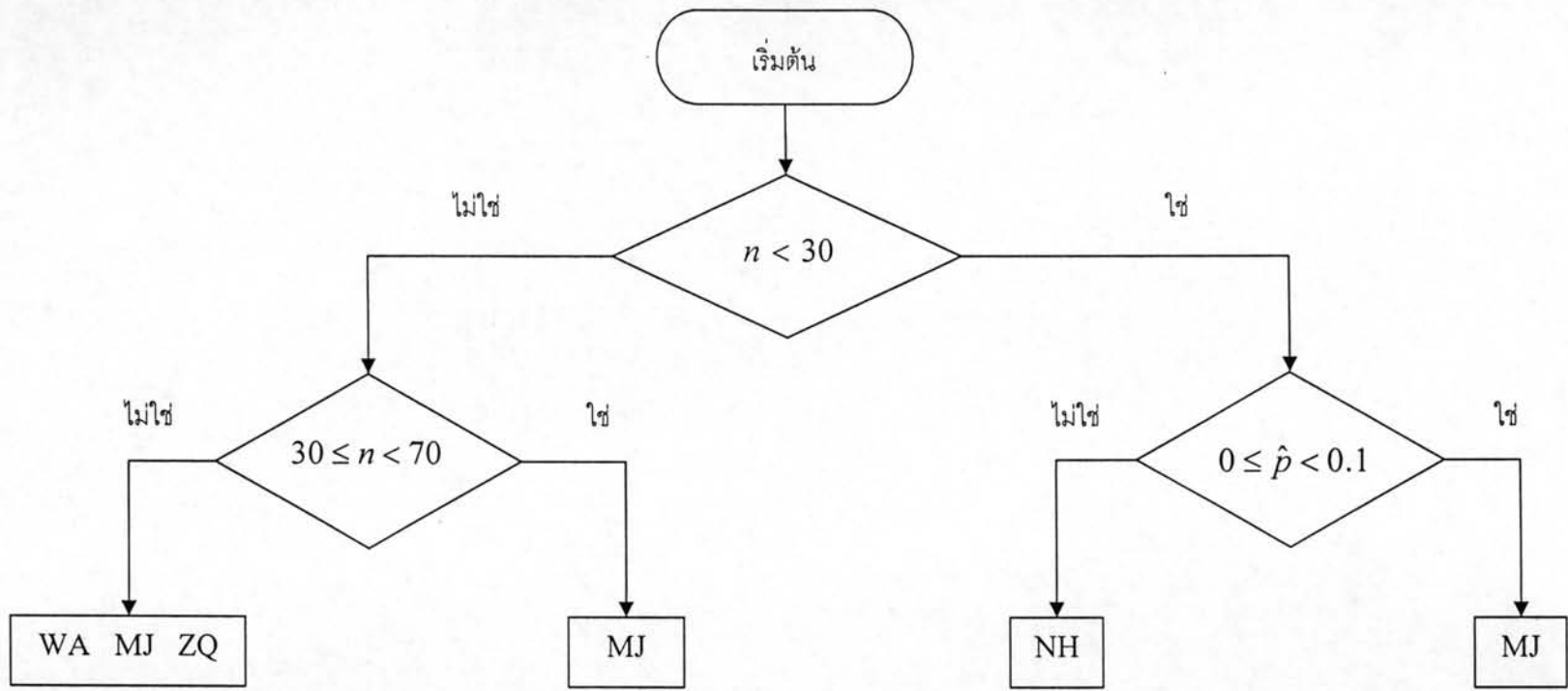
จากค่าประมาณ $\hat{p}_0, \hat{p}_1, \hat{p}_{11}, r$ นำมาใช้เป็นเกณฑ์ในการเลือกว่าวิธีประมาณช่วงความเชื่อมั่นตามแผนภูมิ รูปที่ 5.1 – 5.3



รูปที่ 5.1 แผนภูมิผลสรุปการเลือกวิธีการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่
เมื่อสัมประสิทธิ์สหสัมพันธ์มีค่าอยู่ในระดับ ต่ำ $0 \leq r < 0.4$



รูปที่ 5.2 แผนภูมิผลสรุปการเลือกวิธีการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่
เมื่อสัมประสิทธิ์สหสัมพันธ์มีค่าอยู่ในระดับ กลาง $0.4 \leq r \leq 0.6$



รูปที่ 5.3 แผนภูมิผลสรุปการเลือกวิธีการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่
เมื่อสัมประสิทธิ์สหสัมพันธ์มีค่าอยู่ในระดับ สูง $0.6 < r \leq 1$

5.3.2 การศึกษาวิจัย

1) นอกเหนือจากวิธีการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่ ทั้ง 4 วิธีดังกล่าว พบว่ายังมีวิธีการประมาณอื่นๆ อีกที่น่าสนใจ เช่น วิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องแบบต่างๆ เช่น วิธีการปรับแก้ความต่อเนื่องของ Yates หรือ วิธีการสุ่มตัวอย่างซ้ำ(Bootstrap)

2) ในการวิจัยครั้งนี้ทำการศึกษาเรื่อง วิธีการประมาณค่าแบบช่วงของผลต่างค่าสัดส่วนแบร์นูลลีของข้อมูลแบบจับคู่ ยังมีอีกหลายหัวข้อที่น่าสนใจทำการศึกษา เช่น วิธีการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของประชากรที่ไม่เป็นอิสระกันและมีข้อมูลสูญหาย (INTERVAL ESTIMATION FOR THE DIFFERENCE BETWEEN TWO DEPENDENT PROPORTIONS WITH MISSING DATA) โดยทำการเปรียบเทียบวิธีการประมาณได้แก่ The Traditional Method, The Bootstrap Method และ The Bootstrap Estimate of the standard Error Method เป็นต้น

3) เกณฑ์การเปรียบเทียบระดับความเชื่อมั่น ผู้วิจัยอาจใช้เกณฑ์การเปรียบเทียบที่เข้มมากขึ้นโดยการทดสอบสมมติฐานโดยใช้ตัวสถิติ Z ดังนี้

$$H_0 : c = c_0$$

$$H_1 : c \neq c_0$$

การใช้เกณฑ์ข้างต้นอาจพบว่ามีหลายกรณีที่วิธีการประมาณจะควบคุมไม่ได้