บทที่ 5

สรุปผลการวิจัยและการอภิปรายผล

การวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบวิธีการประมาณขนาดประชากร เมื่อมีการเลือกตัวอย่างแบบแคบเจอร์ - รีแคบเจอร์ 3 วิธี คือ

วิธีที่ 1 วิธีการประมาณโดยอาศัยตัวประมาณ Petersen

วิธีที่ 2 วิธีการประมาณโดยอาศัย Gibbs Sampler

วิธีที่ 3 วิธีการประมาณโดยอาศัยตัวประมาณที่พัฒนาจากตัวประมาณ Petersen โดยการเปรียบเทียบนี้ จะอาศัยค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมพัทธ์สัมบูรณ์ (MAPE) หากวิธีการใดที่มีค่า MAPE ต่ำที่สุด ก็จะกล่าวว่าวิธีนั้นสามารถประมาณได้ดีกว่าวิธีอื่น ๆ โดยจะทำการเปรียบเทียบวิธีที่ 1 กับวิธีที่ 3 และเปรียบเทียบวิธีที่ 2 กับวิธีที่ 3

1. เปรียบเทียบความคลาดเคลื่อนของวิธีที่ 1 และวิธีที่ 3

กรณีที่ 1 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำถูกต้องตามทฤษฎีความน่าจะเป็น

- เมื่อวิธีที่ 3 มีจำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำถูกต้องตามทฤษฎีความน่าจะเป็นด้วยเช่นกัน จะได้ว่า เมื่อประชากรมีขนาดเล็ก วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะสามารถประมาณได้ดีกว่าวิธีที่ 1

- เมื่อวิธีที่ 3 มีค่า m_3 ผิดพลาด 10% ในทิศทางบวก จะได้ว่า เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 มากขึ้นเมื่อ n_2/n_1 มีค่าลดลง แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 1 จะประมาณได้ดีกว่า

- เมื่อวิธีที่ 3 มีค่า m_3 ผิดพลาด 10% ในทิศทางลบ จะได้ว่า วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในเกือบจะทุกค่า n_1/N และ n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_3 ผิดพลาด 20% ในทิศทางบวก จะได้ว่า วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในบางกรณี และจะประมาณได้ดีขึ้นเมื่อขนาดประชากรลดลง หรือ n_2/n_1 มีค่าลดลง

- เมื่อวิธีที่ 3 มีค่า m_3 ผิดพลาด 20% ในทิศทางลบ จะได้ว่า วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3

- เมื่อวิธีที่ 3 มีค่า m_{23} ผิดพลาด 10% ในทิศทางบวก จะได้ว่า วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในบางกรณี และจะประมาณได้ดีขึ้นเมื่อขนาดประชากรลดลง หรือ n_2/n_1 มีค่าลดลง

- เมื่อวิธีที่ 3 มีค่า m_{23} ผิดพลาด 10% ในทิศทางลบ จะได้ว่า เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะดีกว่าวิธีที่ 1 มากขึ้นเมื่อขนาดประชากรเพิ่มขึ้น หรือ n_2/n_1 มีค่าลดลง แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 มากขึ้นเมื่อขนาดประชากรเล็กลง หรือ n_2/n_1 มีค่าเพิ่มขึ้น

- เมื่อวิธีที่ 3 มีค่า m_{23} ผิดพลาด 20% ในทิศทางบวก จะได้ว่า เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะดีกว่าวิธีที่ 1 มากขึ้นเมื่อขนาดประชากรเพิ่มขึ้น หรือ n_2/n_1 มีค่าลดลง แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในเกือบทุกค่า n_1/N และ n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_{23} ผิดพลาด 20% ในทิศทางลบ จะได้ว่า เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะดีกว่าวิธีที่ 1 มากขึ้นเมื่อขนาดประชากรเพิ่มขึ้น หรือ n_2/n_1 มีค่าลดลง แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 มากขึ้นเมื่อขนาดประชากรเล็กลง หรือ n_2/n_1 มีค่าเพิ่มขึ้น

กรณีที่ 2 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 10% ในทิศทางบวก

- เมื่อวิธีที่ 3 มีค่า m_2 ผิดพลาด 10% ในทิศทางบวกแต่ค่าอื่น ๆ เป็นไปตามทฤษฎี วิธีที่ 3 จะดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+10%, +10%) วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+10%, +20%) วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในเกือบทุกขนาดประชากร, ค่าn_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+10%, -10%) เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในบางกรณี โดยจะดีขึ้นเมื่อขนาดประชากรเพิ่มขึ้น และเมื่อค่า n_2/n_1 เพิ่มขึ้น แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะดีกว่าวิธีที่ 1 เกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+10%, -20%) เมื่อประชากรมีขนาดเล็ก วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในบางกรณี และวิธีที่ 3 จะประมาณได้ดีขึ้นเมื่อขนาดประชากรเพิ่มขึ้น หรือค่า n_1/N และ n_2/n_1 เพิ่มขึ้น แต่เมื่อประชากรมีขนาดปานกลางวิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

กรณีที่ 3 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 10% ในทิศทางลบ

- เมื่อวิธีที่ 3 มีค่า m_2 ผิดพลาด 10% ในทิศทางลบ แต่ค่าอื่น ๆ เป็นไปตามทฤษฎี วิธีที่ 3 จะดีกว่าวิธีที่ 1 ในบางกรณี โดยจะประมาณได้ดีขึ้นเมื่อขนาดประชากรเพิ่มขึ้น หรือเมื่อ n_2/n_1 มีค่าเพิ่มขึ้น

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-10%, -10%)วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-10%, -20%)วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-10%, +10%)วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1 แต่วิธีที่ 3 จะประมาณได้ดีขึ้นเมื่อขนาดประชากรเพิ่มขึ้น ,ค่า n_1/N และค่า n_2/n_1 เพิ่มขึ้น

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-10%, +20%) เมื่อประชากรมีขนาดเล็ก วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1 แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะดีกว่าวิธีที่ 1 ในบางกรณี และจะประมาณได้ดีขึ้น เมื่อขนาดประชากรเพิ่ม หรือค่า n_1/N และ n_2/n_1 มีค่าเพิ่มขึ้น

กรณีที่ 4 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 20% ในทิศทางบวก

- เมื่อวิธีที่ 3 มีค่า m_2 ผิดพลาด 20% ในทิศทางบวกแต่ค่าอื่น ๆ เป็นไปตามทฤษฎี วิธีที่ 3 จะดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+20%, +20%) วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+20%, +10%) วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+20%, -20%) วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่าn_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (+20%, -10%) วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

กรณีที่ 5 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 20% ในทิศทางลบ

- เมื่อวิธีที่ 3 มีค่า m_2 ผิดพลาด 20% ในทิศทางลบ แต่ค่าอื่น ๆ เป็นไปตามทฤษฎี วิธีที่ 3 จะดีกว่าวิธีที่ 1 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-20%, -20%) วิธีที่ 1 จะประมาณได้ดีกว่าวิธีที่ 3 ในเกือบทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-20%, -10%) เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในบางกรณี และจะประมาณได้ดีขึ้นเมื่อขนาดประชากรเพิ่มขึ้น แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะประมาณได้ดีกว่าในเกือบทุกขนาดประชากร,ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-20%, +20%) เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในบางกรณี และจะประมาณได้ดีขึ้นเมื่อขนาดประชากรเพิ่มขึ้น แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะประมาณได้ดีกว่าในเกือบทุกขนาดประชากร,ค่า n_1/N และค่า n_2/n_1

- เมื่อวิธีที่ 3 มีค่า m_2 และ m_3 ผิดพลาด (-20%, +10%) เมื่อประชากรมีขนาดเล็ก วิธีที่ 3 จะประมาณได้ดีกว่าวิธีที่ 1 ในบางกรณี และจะประมาณได้ดีขึ้นเมื่อขนาดประชากรเพิ่มขึ้น แต่เมื่อประชากรมีขนาดปานกลาง วิธีที่ 3 จะประมาณได้ดีกว่าในเกือบทุกขนาดประชากร,ค่า n_1/N และค่า n_2/n_1

2. เปรียบเทียบความคลาดเคลื่อนของวิธีที่ 2 และวิธีที่ 3

กรณีที่ 1 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำถูกต้องตามทฤษฎีความน่าจะเป็น จะได้ว่าวิธีที่ 3 จะประมาณค่าได้ดีกว่าวิธีที่ 2 ในทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

กรณีที่ 2 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 10% ในทิศทางบวก จะได้ว่าวิธีที่ 3 จะประมาณค่าได้ดีกว่าวิธีที่ 2 ในทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

กรณีที่ 3 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 10% ในทิศทางลบ จะได้ว่าวิธีที่ 3 จะประมาณค่าได้ดีกว่าวิธีที่ 2 ในทุกขนาดประชากร, ค่า n_1/N และค่าn_2/n_1

กรณีที่ 4 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 20% ในทิศทางบวก จะได้ว่าวิธีที่ 3 จะประมาณค่าได้ดีกว่าวิธีที่ 2 ในทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

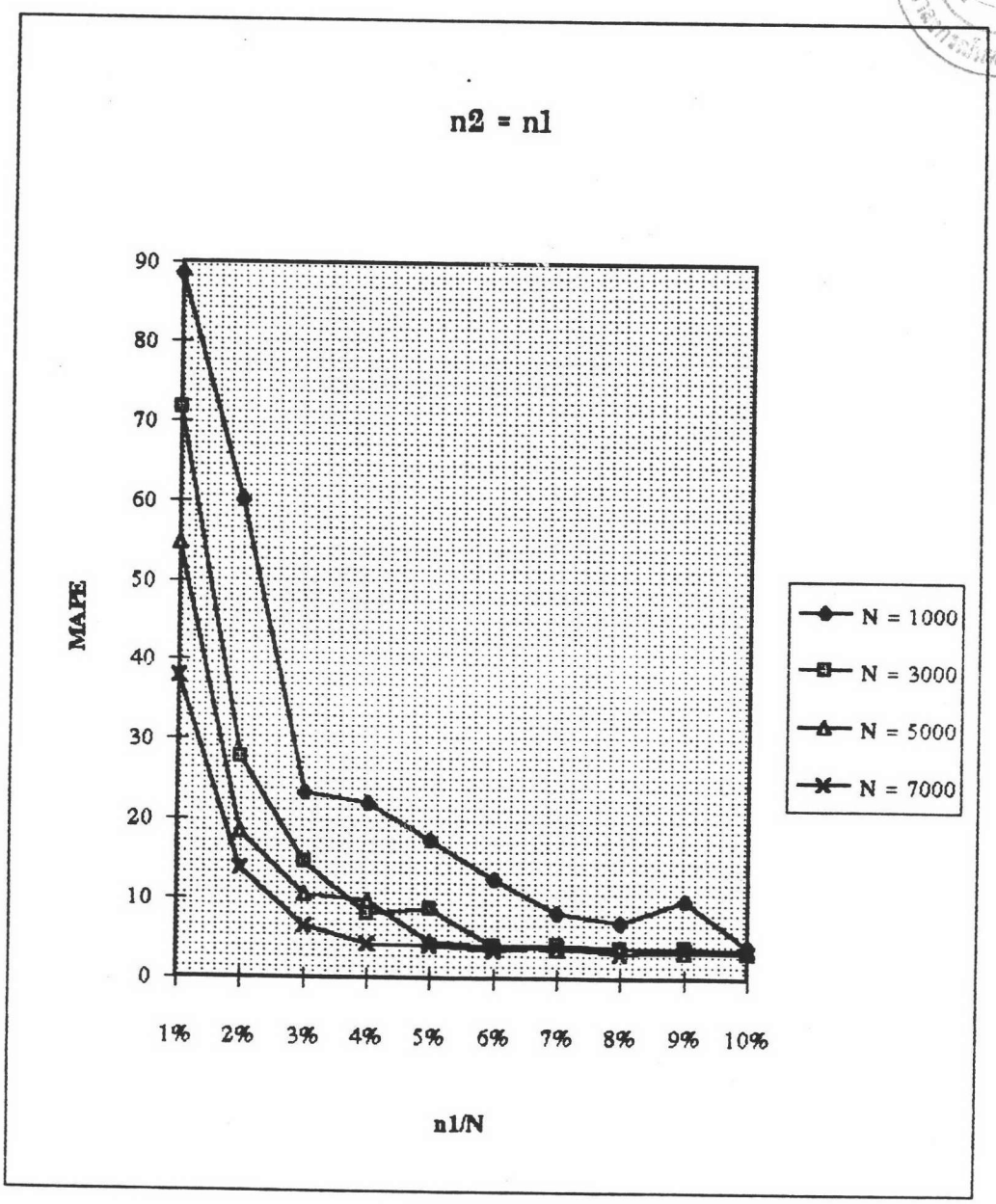
กรณีที่ 5 : จำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในการเลือกครั้งที่ 2 (m_2) ผิดพลาด 20% ในทิศทางลบ จะได้ว่าวิธีที่ 3 จะประมาณค่าได้ดีกว่าวิธีที่ 2 ในทุกขนาดประชากร, ค่า n_1/N และค่า n_2/n_1

การอภิปรายผล

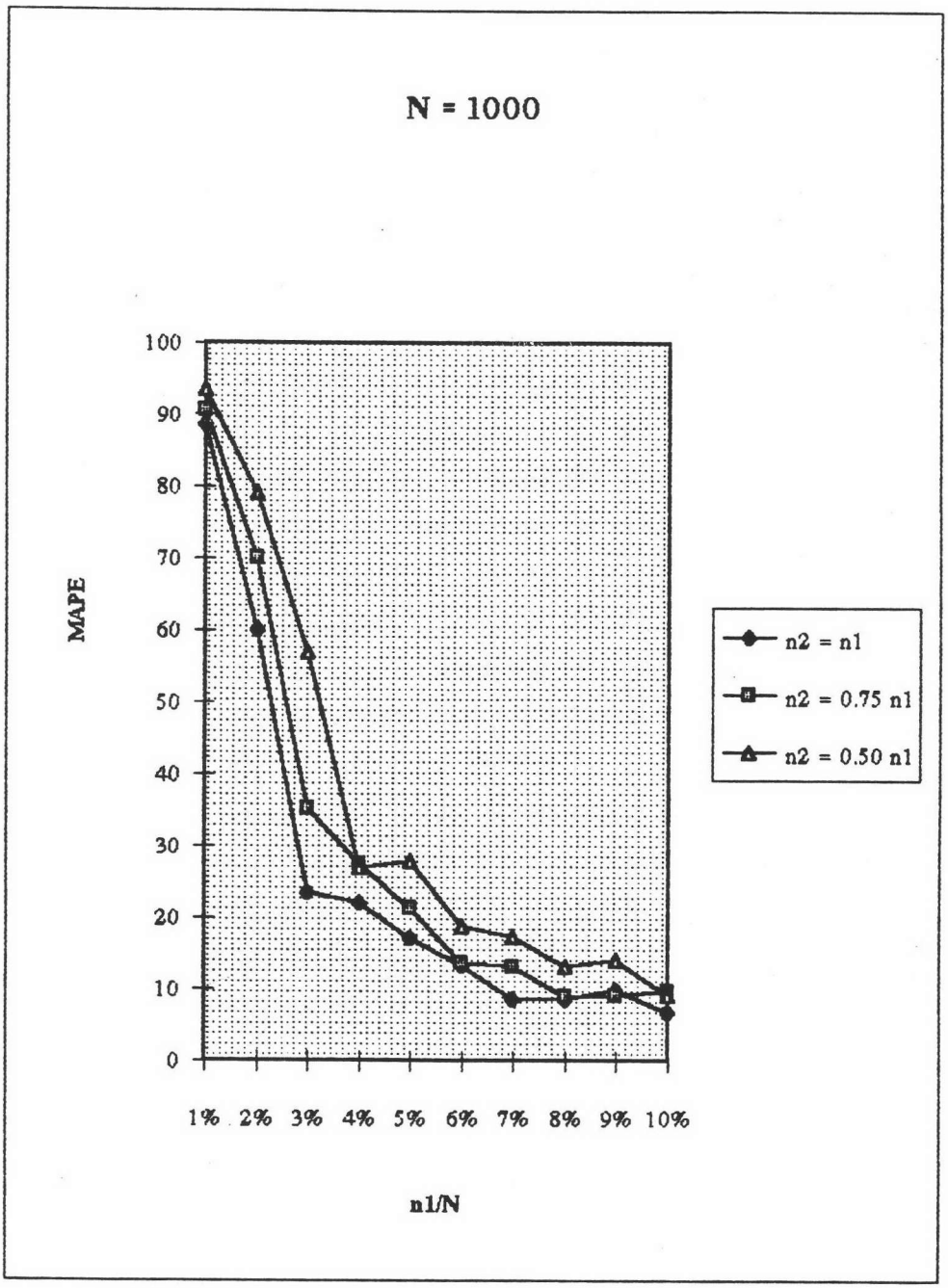
จากผลการเปรียบเทียบข้างต้น จะเห็นได้ว่า วิธีที่ 3 สามารถประมาณขนาดประชากรได้ดีกว่าวิธีที่ 1 ในเกือบทุกกรณี ยกเว้นในกรณีที่ประชากรมีขนาดเล็กและมีจำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำเป็นไปตามทฤษฎีความน่าจะเป็น และกรณีที่ m_2 มีความผิดพลาดร้อยละ 10 ในทิศทางลบ ทั้งนี้เพราะในกรณีนี้ จำนวนซ้ำมีค่ามากพอสมควร ส่วนการที่วิธีที่ 2 ประมาณได้ไม่ดีนั้น อาจเกิดจากค่าพารามิเตอร์ที่จำลองให้กับการแจกแจงเบตา (คือค่า $(r, 1-\prod_{i=1}^{3}(1-p_i))$ ไม่เหมาะสม อันเนื่องมาจากการกำหนดค่า (a,b) := (1,1) ซึ่งเปรียบเสมือนกรณีที่ไม่มีข้อมูลเบื้องต้น และเนื่องจากวิธีการประมาณทั้งสามวิธีนี้ จะประมาณได้ดีหรือไม่ขึ้นอยู่กับจำนวนหน่วยตัวอย่างที่ถูกเลือกซ้ำในแต่ละการเลือกตัวอย่าง ดังนั้นหากการเลือกตัวอย่างไม่พบหน่วยตัวอย่างซ้ำเลยตลอดการเลือกตัวอย่าง ก็ไม่สามารถนำวิธีทั้งสามนี้ไปประมาณขนาดประชากรได้

จากผลสรุปการวิจัยข้างต้นสามารถแสดงด้วยกราฟ เพื่อดูแนวโน้มของค่า MAPE ได้ดังนี้

รูปที่ 1 : แสดงแนวโน้มของค่า MAPE เมื่อพิจารณาตามขนาดประชากร

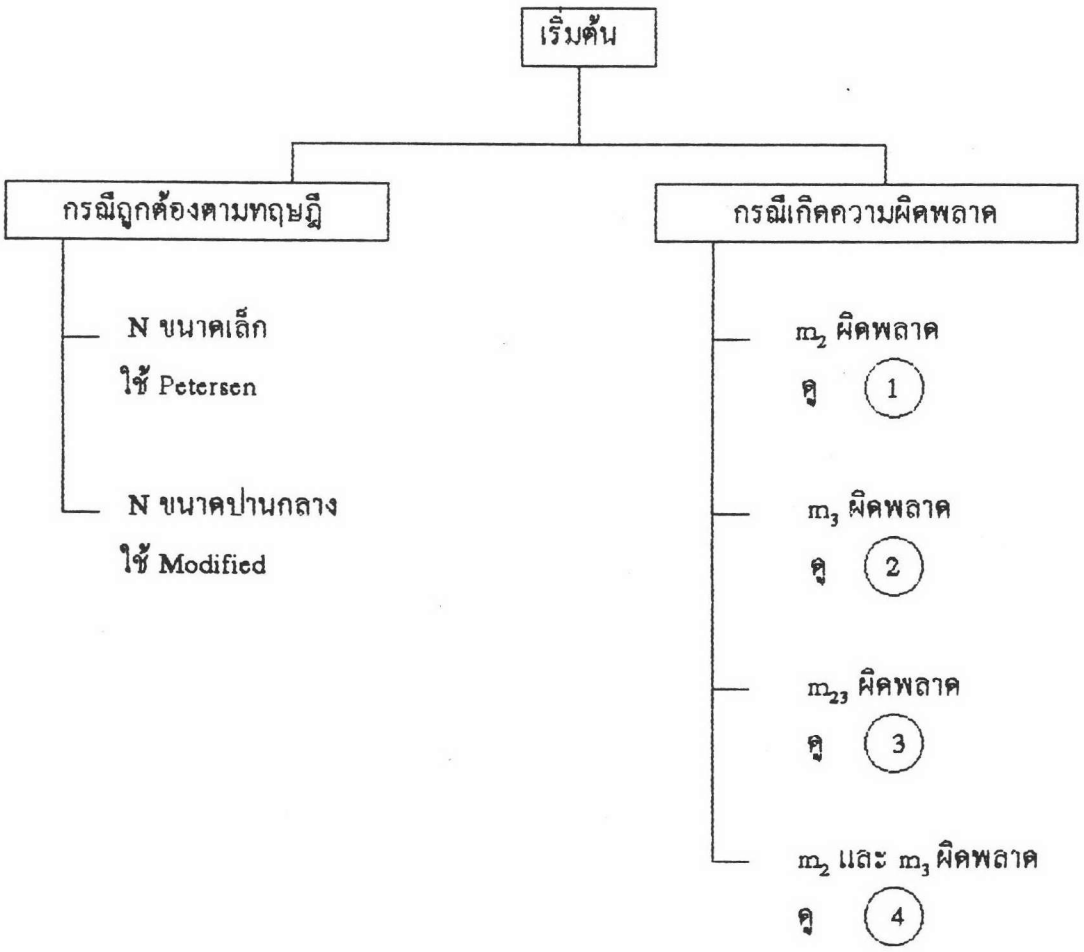


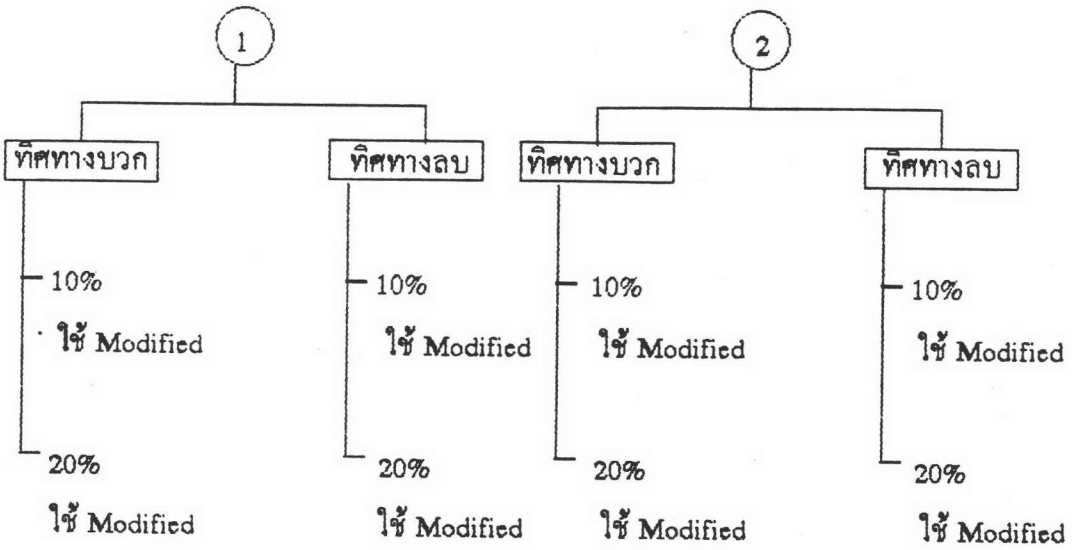
รูปที่ 2 : แสดงแนวโน้มของค่า MAPE เมื่อพิจารณาค่าสัดส่วนระหว่างขนาดตัวอย่างในการเลือกครั้งที่ 2 และครั้งที่ 1

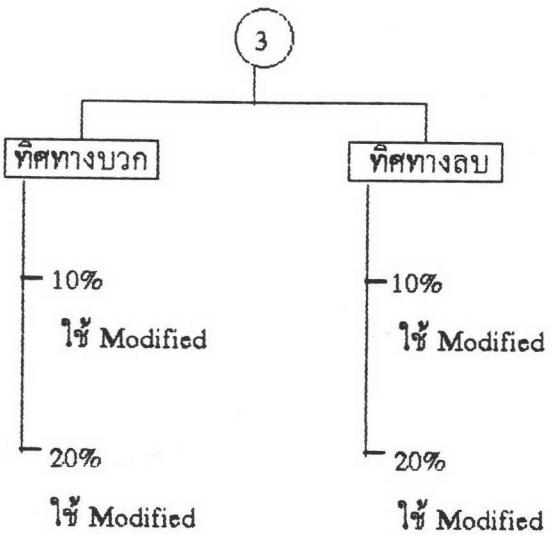


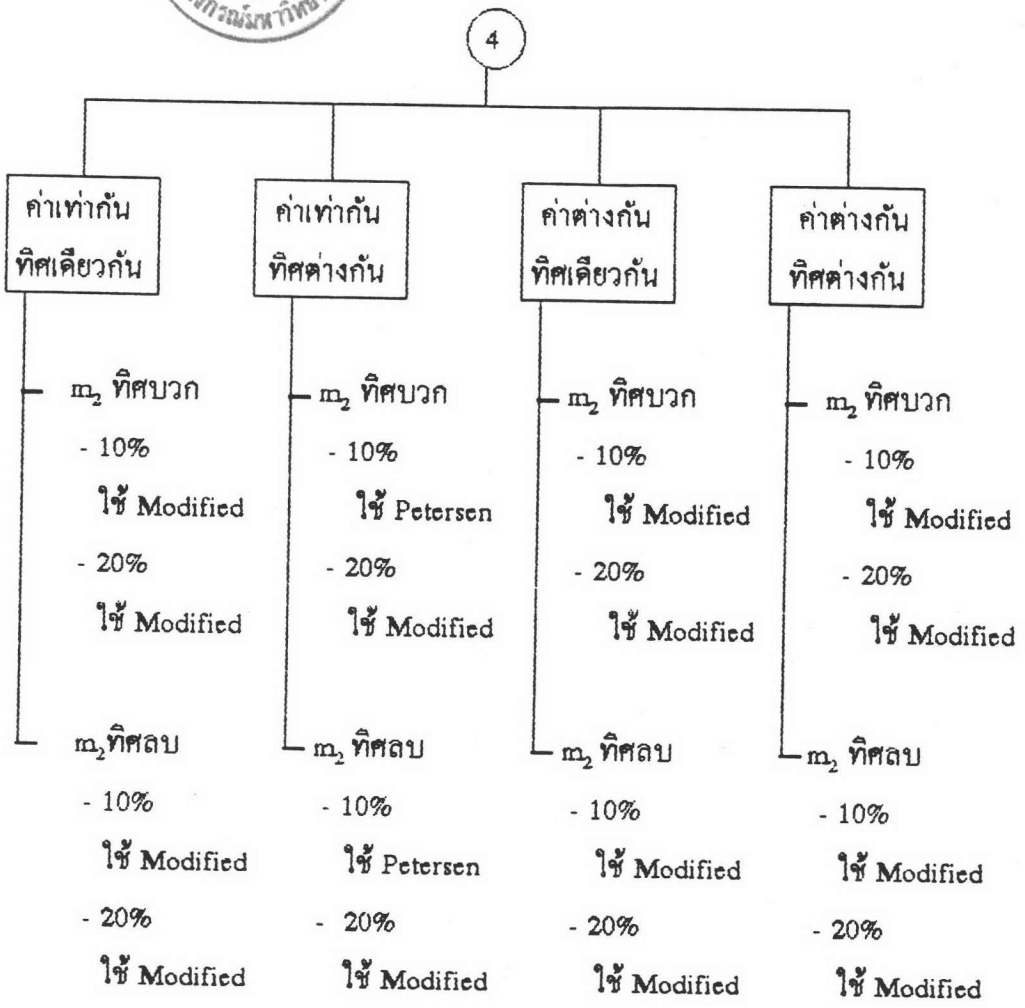
จากรูปที่ 1 เมื่อประชากรมีขนาดเพิ่มขึ้นค่า MAPE จะลดลง หมายความว่าวิธีการประมาณเมื่อมีการเลือกตัวอย่างแบบแคบเจอร์-รีแคบเจอร์ จะประมาณได้ดีถ้าประชากรมีขนาดใหญ่ และจากรูปที่ 2 เมื่อค่าสัดส่วนระหว่างขนาดตัวอย่างในการเลือกครั้งที่ 2 และครั้งที่ 1 มีค่าเพิ่มขึ้น หรือเมื่อขนาดตัวอย่างในการเลือกครั้งที่ 2 มีค่าใกล้เคียงขนาดตัวอย่างในการเลือกครั้งที่ 1 จะได้ว่าค่า MAPE จะลดลง ทั้งนี้การลดลงของค่า MAPE จะไม่แตกต่างกันมากนัก ดังนั้นในการเลือกตัวอย่างจึงไม่จำเป็นที่จะต้องเลือกขนาดตัวอย่างในครั้งที่ 2 ให้ใกล้เคียงครั้งที่ 1 โดยอาจจะเลือกให้มีขนาดเพียงครึ่งเดียวของการเลือกตัวอย่างในครั้งที่ 1 ก็ได้ นอกจากนี้จากรูปทั้งสอง แสดงให้เห็นว่า เมื่อสัดส่วนของขนาดตัวอย่างในการเลือกครั้งที่ 1 ต่อขนาดประชากรมีค่าเพิ่มขึ้น ค่า MAPE จะลดลง โดยจะลดลงเร็วมากในค่าสัดส่วนน้อย ๆ จนถึงระดับ 5% จึงลดลงน้อย ดังนั้น ก็ไม่มีความจำเป็นที่จะต้องทำการเลือกตัวอย่างให้มีขนาดใหญ่มากนัก โดยเลือกมาประมาณครึ่งหนึ่งของขนาดประชากรก็ให้ค่าประมาณที่ถูกต้องพอสมควรแล้ว

เพื่อให้ง่ายแก่การเลือกใช้วิธีการประมาณแบบต่าง ๆ ขอสรุปเป็นแผนภาพดังนี้




1
ทิศทางบวก
ทิศทางลบ
10%
ใช้ Modified
20%
ใช้ Modified
10%
ใช้ Modified
20%
ใช้ Modified
2
ทิศทางบวก
ทิศทางลบ
10%
ใช้ Modified
20%
ใช้ Modified
10%
ใช้ Modified
20%
ใช้ Modified


3
ทิศทางบวก
ทิศทางลบ
10%
ใช้ Modified
20%
ใช้ Modified
10%
ใช้ Modified
20%
ใช้ Modified



ข้อเสนอแนะ

ในการวิจัยครั้งนี้ ขอบเขตที่กำหนดยังไม่ครอบคลุมในกรณีของวิธีที่ 3 ทั้งนี้เพราะมีการกำหนดค่า (a,b) = (1,1) เพียงคู่ลำดับเดียว ทำให้การสรุปสำหรับวิธีนี้ยังไม่ดีเท่าที่ควร นอกจากนี้การกำหนดสัดส่วนให้แก่ขนาดตัวอย่างที่ทำการเลือกในครั้งที่ 1 ต่อขนาดประชากรไม่กระจายมากนัก ทั้งนี้เพราะผู้วิจัยคำนึงถึงการนำไปประยุกต์ใช้จริง ซึ่งย่อมต้องการที่จะกำหนดขนาดตัวอย่างไม่มากนัก เพื่อลดค่าใช้จ่ายในการเลือกตัวอย่างและเก็บข้อมูล สิ่งที่น่าสนใจทำการวิจัยเพิ่มเติม คือ กรณีที่ขนาดประชากรไม่คงที่ อันเนื่องมาจากการเกิด-การตาย, การอพยพเข้า-ออก

เป็นต้น,กรณีที่ขนาดตัวอย่างในการเลือกตัวอย่างครั้งที่ 2 มีค่ามากกว่าขนาดตัวอย่างในการเลือกตัวอย่างครั้งที่ 1 เป็นสัดส่วนต่าง ๆ และกรณีที่ขนาดประชากรและความน่าจะเป็นมีการแจกแจงที่แตกต่างออกไปจากการวิจัยนี้