บทที่ 6

สรุปผลการวิจัยและข้อเสนอแนะ

6.1 สรุปผลการวิจัย

ในการวิจัยครั้งนี้ได้เสนอตัวประมาณค่ารวมประชากร 3 รูปแบบ ($\hat{Y}_k$; k=1,2,3) โดยอาศัยคุณสมบัติเกี่ยวกับตัวประมาณความถดถอยและการแบ่งชั้นภูมิเมื่อเลือกตัวอย่างแล้วมาใช้เพื่อกำหนดรูปแบบและวิธีการประมาณค่ารวมประชากรที่มีค่าบางค่าสูงมาก เมื่อตัวอย่างที่เลือกมาแบบสุ่มอย่างง่ายซึ่งมิได้คัดมีค่าบางหน่วยตัวอย่างเป็นค่าสูงมากรวมอยู่ด้วย พร้อมทั้งศึกษาคุณสมบัติและเปรียบเทียบประสิทธิภาพของตัวประมาณที่เสนอแนะกับตัวประมาณค่ารวมประชากรที่เสนอโดยไมเคิลและคาดาบา ($\hat{Y}_{mkt}$; t = 1,2,3,4) และตัวประมาณ $\hat{Y}_o$ โดยศึกษาจากค่าประสิทธิภาพสัมพัทธ์เป็นหลัก ซึ่งในการวิจัยครั้งนี้กำหนดให้

ประสิทธิภาพสัมพัทธ์ของตัวประมาณ $\hat{Y}_k$ เทียบกับ $\hat{Y}_o$

$$= \frac{\text{ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของ } \hat{Y}_o}{\text{ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของ } \hat{Y}_k}$$

หรือ ประสิทธิภาพสัมพัทธ์ของตัวประมาณ $\hat{Y}_{mkt}$ เทียบกับ $\hat{Y}_o$

$$= \frac{\text{ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของ } \hat{Y}_o}{\text{ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของ } \hat{Y}_{mkt}}$$

ซึ่งผลสรุปจากการวิเคราะห์ข้อมูลในบทที่ 5 จำแนกตามรูปแบบของการอนุมานได้ดังนี้

6.1.1 เมื่อใช้การอนุมานอย่างมีเงื่อนไข

จากการทดลองหาค่าประสิทธิภาพสัมพัทธ์ของตัวประมาณที่เสนอแนะหรือตัวประมาณที่เสนอโดยไมเคิลและคาดาบา เทียบกับตัวประมาณ $\hat{Y}_o$ เมื่อใช้การอนุมานอย่างมีเงื่อนไข เพื่อใช้เป็นเกณฑ์ในการศึกษาคุณสมบัติและเปรียบเทียบประสิทธิภาพของตัวประมาณที่เสนอนี้เทียบกับตัวประมาณที่เสนอโดย ไมเคิลและคาดาบา และตัวประมาณ $\hat{Y}_o$ จำแนกผลสรุปได้ดังนี้

ก. กรณีที่ทราบจำนวนค่าลังเกตที่เป็นค่าสูงมากในประชากร

จากการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่ผู้วิจัยเสนอแนะ ($\hat{Y}_1$) ตัวประมาณที่เสนอโดยไมเคิล และคาดาบา ($\hat{Y}_{mk4}$) และตัวประมาณ $\hat{Y}_o$ พบว่า เมื่อศึกษาจากสูตรความเอนเอียงอย่างมีเงื่อนไขเมื่อกำหนดค่า n_1 แล้ว $\hat{Y}_1$ และ $\hat{Y}_{mk4}$ จะเป็นตัวประมาณที่ไม่เอนเอียงของค่ารวมประชากร แต่ $\hat{Y}_o$ เป็นตัวประมาณที่มีความเอนเอียงของ ค่ารวมประชากร และจากการวิเคราะห์ผลโดยการจำลองข้อมูลขึ้นในเครื่องคอมพิวเตอร์ IBM 370/3031 เมื่อให้ตัวแปร Y มีการแจกแจงแบบล็อกนอร์มอลหรือการแจกแจงแบบแกมม่าจะได้ว่า $\hat{Y}_1$ จะมีประสิทธิภาพลัมพัทธ์ลดลงเมื่อร้อยละของจำนวนค่าลังเกตที่เป็นค่าสูงมากที่พบในประชากรเพิ่มมากขึ้นแต่ประสิทธิภาพลัมพัทธ์ของ $\hat{Y}_1$ เทียบกับ $\hat{Y}_o$ ยังคงสูงกว่าประสิทธิภาพลัมพัทธ์ของ $\hat{Y}_{mk4}$ เทียบกับ $\hat{Y}_o$ ในทุกกรณี และถ้าตัวแปร Y ที่สนใจจะประมาณค่ารวมประชากรมีความสัมพันธ์เชิงเส้นต่อกันกับตัวแปร X หรือในตัวอย่างมีร้อยละของจำนวนหน่วยตัวอย่างที่เป็นค่าสูงมาก เพิ่มมากขึ้นแล้ว $\hat{Y}_1$ จะมีประสิทธิภาพลัมพัทธ์ เมื่อเทียบกับ $\hat{Y}_o$ สูงกว่า $\hat{Y}_{mk4}$ อย่างเห็นได้ชัด โดย $\hat{Y}_1$ จะมีค่าประสิทธิภาพลัมพัทธ์ประมาณ 2 เท่าของ $\hat{Y}_{mk4}$ ถ้าค่าสัมประสิทธิ์สหสัมพันธ์เชิงเส้นระหว่างตัวแปร Y และตัวแปร X มีค่าเท่ากับ 0.7 หรือ -0.7

ข. กรณีที่ไม่ทราบจำนวนค่าลังเกตที่เป็นค่าสูงมากในประชากร

จากการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่เสนอแนะ ($\hat{Y}_2$, $\hat{Y}_3$) ตัวประมาณที่ไมเคิลและคาดาบาเสนอ ($\hat{Y}_{mk1}$, $\hat{Y}_{mk2}$, $\hat{Y}_{mk3}$) และตัวประมาณ $\hat{Y}_o$ พบว่าเมื่อศึกษาจากสูตรความเอนเอียงอย่างมีเงื่อนไข ตัวประมาณดังกล่าวจะเป็นตัวประมาณที่มีความเอนเอียงสำหรับค่ารวมประชากร และจากการวิเคราะห์ผลโดยการจำลองข้อมูล เมื่อสมมติให้ตัวแปร Y มีการแจกแจงแบบล็อกนอร์มอล หรือการแจกแจงแบบแกมม่า พบว่าในกรณีที่ร้อยละของจำนวนค่าลังเกตที่เป็นค่าสูงมากที่พบในประชากรมีค่าเท่ากับ 1.8% ณ ระดับที่ร้อยละของจำนวนหน่วยตัวอย่างที่มีค่าสูงมากที่พบในตัวอย่างและค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร Y และตัวแปร X มีค่าน้อยจะได้ $\hat{Y}_{mk2}$ เป็นตัวประมาณที่มีประสิทธิภาพลัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงกว่าตัวประมาณอื่น ๆ ที่เปรียบเทียบด้วย (ในที่นี้ตัวประมาณอื่น ๆ คือ $\hat{Y}_2$, $\hat{Y}_3$, $\hat{Y}_{mk1}$ และ $\hat{Y}_{mk3}$) แต่ทั้งนี้ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร Y และตัวแปร X มีค่ามากกว่าหรือเท่ากับ 0.7 (หรือ -0.7) แล้ว $\hat{Y}_3$ จะเป็นตัวประมาณที่มีประสิทธิภาพลัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด ส่วน ณ ระดับที่ร้อยละของจำนวนหน่วยตัวอย่างที่มีค่าสูงมากที่พบในตัวอย่างมีค่ามาก

$\hat{Y}_3$ ก็จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงกว่าตัวอื่น ๆ อย่างเห็นได้ชัด และประสิทธิภาพสัมพัทธ์ของ $\hat{Y}_3$ เทียบกับ $\hat{Y}_o$ จะเทียบเท่าประสิทธิภาพสัมพัทธ์ของ $\hat{Y}_1$ เทียบกับ $\hat{Y}_o$ ถ้าจำนวนของหน่วยตัวอย่างที่มีค่าสูงมากที่พบในตัวอย่างมีค่าเข้าใกล้จำนวนค่าสังเกตที่มีค่าสูงมากทั้งหมดที่พบในประชากรมากยิ่งขึ้น ส่วนในกรณีที่ร้อยละของจำนวนค่าสังเกตที่เป็นค่าสูงมากที่พบในประชากรมีจำนวนมากยิ่งขึ้น ถึงแม้ว่าตัวประมาณ $\hat{Y}_3$ $\hat{Y}_{mk1}$ และ $\hat{Y}_{mk3}$ จะมีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ ลดลงก็ตาม ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณดังกล่าวจะให้ผลในทำนองเดียวกันยกเว้น ณ ระดับที่ร้อยละของจำนวนหน่วยตัวอย่างที่มีค่าสูงมากที่พบในตัวอย่าง มีค่าน้อย ๆ แล้ว $\hat{Y}_2$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบ $\hat{Y}_o$ สูงสุด กล่าวโดยสรุปแล้ว ถึงแม้ว่าจำนวนค่าสังเกตที่เป็นค่าสูงมากในประชากร จะมีจำนวนมากขึ้นแล้วทำให้ประสิทธิภาพสัมพัทธ์ของตัวประมาณดังกล่าวเมื่อเทียบกับ $\hat{Y}_o$ ลดลงก็ตาม แต่ถ้าตัวแปร Y และตัวแปร X มีความสัมพันธ์เชิงเส้นต่อกันเพิ่มมากขึ้นหรือมีจำนวนหน่วยตัวอย่างที่มีค่าสูงมากเพิ่มขึ้นแล้ว $\hat{Y}_3$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด

6.1.2 <u>เมื่อใช้การอนุมานอย่างไม่มีเงื่อนไข</u>

จากการทดลองหาค่าประสิทธิภาพสัมพัทธ์ของตัวประมาณที่ผู้วิจัยเสนอแนะหรือตัวประมาณที่เสนอโดยไมเคิล และคาดาบา เทียบกับตัวประมาณ $\hat{Y}_o$ ในการอนุมานอย่างไม่มีเงื่อนไข เนื่องจากหลักการของการอนุมานแบบไม่มีเงื่อนไขได้คิดเฉลี่ยจำนวนหน่วยตัวอย่างที่เป็นค่าสูงมากที่จะพบในตัวอย่างด้วยค่าโมเมนต์ที่ k ของจำนวนหน่วยตัวอย่างที่เป็นค่าสูงมาก ดังนั้น การเปลี่ยนแปลงค่าร้อยละหรือจำนวนหน่วยตัวอย่างที่มีค่าสูงมากที่พบในตัวอย่างจึงไม่มีผลกระทบต่อการเปลี่ยนแปลงประสิทธิภาพของตัวประมาณ ผลสรุปจำแนกได้ดังนี้

ก. กรณีที่ทราบจำนวนค่าสังเกตที่เป็นค่าสูงมากในประชากร

เมื่อศึกษาจากสูตรความเอนเอียงอย่างไม่มีเงื่อนไข พบว่า $\hat{Y}_1$ และ $\hat{Y}_{mk4}$ เป็นตัวประมาณที่ไม่เอนเอียงของค่ารวมประชากร แต่ $\hat{Y}_o$ เป็นตัวประมาณที่มีความเอนเอียงของค่ารวมประชากร จากการวิเคราะห์ผลโดยการจำลองข้อมูล เมื่อสมมติให้ตัวแปร Y มีการแจกแจงแบบล็อกนอร์มอลหรือการแจกแจงแบบแกมม่า ก็จะยังคงพบว่า $\hat{Y}_1$ จะมีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ ลดลง เมื่อในประชากรมีจำนวนค่าสังเกตที่เป็นค่าสูงมากเพิ่มขึ้นสำหรับในการแจกแจงแบบแกมม่าก็ตาม $\hat{Y}_1$ ก็ยังเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงกว่า $\hat{Y}_{mk4}$ และถ้าตัวแปร Y และตัวแปร X มีความสัมพันธ์เชิงเส้นกันต่อกัน

เพิ่มขึ้นแล้ว $\hat{Y}_1$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงกว่า $\hat{Y}_{mk4}$ อย่างเห็นได้ชัด

ข. กรณีที่ไม่ทราบจำนวนค่าสังเกตที่เป็นค่าสูงมากในประชากร

เมื่อศึกษาจากสูตรความเอนเอียงอย่างไม่มีเงื่อนไข พบว่า $\hat{Y}_2$ $\hat{Y}_3$ $\hat{Y}_{mk1}$ $\hat{Y}_{mk2}$ $\hat{Y}_{mk3}$ และ $\hat{Y}_o$ เป็นตัวประมาณที่มีความเอนเอียงของค่ารวมประชากร และจากการวิเคราะห์ผลโดยการจำลองข้อมูล เมื่อสมมติให้ตัวแปร Y มีการแจกแจงแบบล็อก-นอร์มอลนั้นพบว่า ไม่ว่าจะมีจำนวนร้อยละของจำนวนค่าสังเกตที่เป็นค่าสูงมากที่พบในประชากรเป็นเท่าใดก็ตาม เมื่อใช้ขนาดตัวอย่างเท่ากับ 50 และความสัมพันธ์เชิงเส้นระหว่างตัวแปร Y และตัวแปร X มีค่าเท่ากับ 0.1 0.3 หรือ 0.5 (หรือเท่ากับ -0.1 -0.3 หรือ - 0.5) จะได้ $\hat{Y}_{mk3}$ เป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด แต่ถ้าความสัมพันธ์เชิงเส้นระหว่างตัวแปร Y และตัวแปร X มีค่าเท่ากับ 0.7 หรือ -0.7 แล้ว $\hat{Y}_2$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด และเมื่อใช้ขนาดตัวอย่างเท่ากับ 100 หรือ 200 $\hat{Y}_{mk3}$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุดถ้าค่าสัมประสิทธิ์หลัมพันธ์มีค่าเท่ากับ 0.1 หรือ 0.3 ส่วน $\hat{Y}_2$จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด ถ้าค่าสัมประสิทธิ์หลัมพันธ์มีค่าเท่ากับ 0.5 หรือ 0.7 สำหรับกรณีที่ตัวแปร Y มีการแจกแจงแบบแกมม่า พบว่า ไม่ว่าในประชากรจะมีจำนวนค่าสังเกตที่เป็นค่าสูงมากเป็นจำนวนเท่าใดก็ตาม เมื่อใช้ขนาดตัวอย่างเท่ากับ 50 จะได้ $\hat{Y}_{mk3}$ เป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$สูงสุดถ้าความสัมพันธ์เชิงเส้นระหว่างตัวแปร Y และตัวแปร X มีค่าเท่ากับ 0.1 หรือ 0.3 แต่ถ้าความสัมพันธ์เชิงเส้นระหว่างตัวแปร Y และตัวแปร X มีค่ามากขึ้นเป็น 0.5 หรือ 0.7 แล้ว $\hat{Y}_2$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด ส่วนกรณีเมื่อใช้ขนาดตัวอย่างเพิ่มขึ้นเป็น 100 หรือ 200 แล้ว $\hat{Y}_{mk3}$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุดเฉพาะเมื่อค่าสัมประสิทธิ์หลัมพันธ์เท่ากับ 0.1 แต่เมื่อค่าสัมประสิทธิ์หลัมพันธ์ระหว่างตัวแปร Y และตัวแปร X เพิ่มขึ้นเป็น 0.3 0.5 หรือ 0.7 แล้วตัวประมาณ $\hat{Y}_2$ จะมีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด กล่าวโดยสรุปแล้ว ถึงแม้ว่าตัวประมาณ $\hat{Y}_2$ $\hat{Y}_3$ และ $\hat{Y}_{mki}$; i = 1,2,3 จะมีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ ลดลงถ้า ในประชากรมีร้อยละหรือจำนวนของค่าสังเกตที่เป็นค่าสูงมากเพิ่มมากขึ้น และเมื่อความสัมพันธ์เชิงเส้นระหว่างตัวแปร Y และตัวแปร X มีค่าน้อย จะได้ $\hat{Y}_{mk2}$ เป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุด ส่วน $\hat{Y}_2$ จะเป็นตัวประมาณที่มี

ประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุดเมื่อตัวแปร Y มีความสัมพันธ์เชิงเส้นต่อกันกับตัวแปร X เพิ่มมากขึ้น

เนื่องจากในการอนุมานอย่างไม่มีเงื่อนไข มีการคำนวณหาค่าความน่าจะเป็นสำหรับแต่ละค่าที่เป็นไปได้ของจำนวนหน่วยตัวอย่างที่มีค่าสูงมาก และการแจกแจงของจำนวนหน่วยตัวอย่างที่มีค่าสูงมากเป็นแบบ positive Hypergeometric distribution ซึ่งจะมีปัญหาในการคำนวณหาค่าความน่าจะเป็นในกรณีที่ประชากรมีขนาดใหญ่ เพื่อเป็นการแก้ปัญหาดังกล่าว วิทยานิพนธ์นี้จึงได้เสนอการประมาณค่าความน่าจะเป็นจากฟังก์ชั่นความหนาแน่นของ positive Hypergeometric distribution $p(n_1 \mid n,N,N_1)$ ด้วยฟังก์ชั่นความหนาแน่น $p(n_1)$ ซึ่งผลจากการทดสอบความแตกต่างของค่าความน่าจะเป็นที่คำนวณจากทั้งสองฟังก์ชั่นความหนาแน่นในบทที่ 5 จะเห็นได้ว่า ความน่าจะเป็นที่คำนวณจากทั้งสองฟังก์ชั่นความหนาแน่นให้ค่าไม่แตกต่างกัน ณ ระดับนัยสำคัญเท่ากับ 0.05 ทั้งนี้การนำฟังก์ชั่นความหนาแน่น $p(n_1)$ มาใช้ประมาณค่าความน่าจะเป็นจากฟังก์ชั่นความหนาแน่น $P(n_1 \mid n,N,N_1)$ จะให้ผลดีถ้า $\frac{n}{N}$ มีค่าน้อยกว่าหรือเท่ากับ 0.1

6.2 ข้อเสนอแนะ

1. จากการสรุปผลการวิเคราะห์เพื่อศึกษาคุณสมบัติเปรียบเทียบ ประสิทธิภาพของตัวประมาณที่ผู้วิจัยเสนอแนะกับตัวประมาณที่เสนอโดย ไมเคิล และคาดาบา และตัวประมาณ $\hat{Y}_o$ พบว่าไม่ว่ารูปแบบของการอนุมานจะเป็นแบบใด ประสิทธิภาพสัมพัทธ์ของตัวประมาณที่ผู้วิจัยเสนอเทียบกับ $\hat{Y}_o$ มีความสัมพันธ์ในทางเดียวกันกับค่าสัมประสิทธิ์สหสัมพันธ์ กล่าวคือ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์เพิ่มขึ้นค่าประสิทธิภาพสัมพัทธ์ของตัวประมาณที่เสนอเมื่อเทียบกับตัวประมาณ $\hat{Y}_o$ จะมีค่าเพิ่มขึ้นด้วย และถ้าค่าสัมประสิทธิ์สหสัมพันธ์มีค่าตั้งแต่ 0.7 ขึ้นไป ตัวประมาณที่เสนอแนะจะมีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงกว่าตัวประมาณที่เสนอโดย ไมเคิล และคาดาบา และมีค่ามากกว่า 1 กล่าวคือ $\hat{Y}_1$ จะมีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงกว่า $\hat{Y}_{mk4}$ เสมอ ส่วน $\hat{Y}_2$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุดในกรณีการอนุมานอย่างไม่มีเงื่อนไข และ $\hat{Y}_3$ จะเป็นตัวประมาณที่มีประสิทธิภาพสัมพัทธ์เมื่อเทียบกับ $\hat{Y}_o$ สูงสุดในกรณีการอนุมานอย่างมีเงื่อนไข เพราะฉะนั้นในการเลือกตัวอย่างแบบสุ่มอย่างง่ายและมีตัวอย่างบางหน่วยมีค่าสูงมากที่เป็นค่าที่มีอยู่จริงในประชากร ถ้าตัวแปรที่เราสนใจศึกษามีความสัมพันธ์เชิงเส้นค่อนข้างสูงต่อกันกับตัวแปรอื่น ในกรณีนี้ควรจะใช้ $\hat{Y}_1$ เป็นตัวประมาณค่า

รวมประชากร เมื่อทราบจำนวนค่าสังเกตที่เป็นค่าสูงมากในประชากร (ไม่ว่าจะใช้การอนุมานรูปแบบใดก็ตาม) สำหรับกรณีที่ไม่ทราบจำนวนค่าสังเกตที่เป็นค่าสูงมากในประชากร เมื่อใช้การอนุมานอย่างไม่มีเงื่อนไขควรจะใช้ $\hat{Y}_2$ และควรจะใช้ $\hat{Y}_3$ ในการอนุมานอย่างมีเงื่อนไข

2. ในกรณีที่ผู้วิจัยทราบจำนวนหน่วยตัวอย่างที่มีค่าสูงมากที่พบในตัวอย่าง แต่ไม่แน่ใจว่าจำนวนที่พบนั้นมีค่ามากหรือไม่เมื่อเทียบกับขนาดตัวอย่างที่ใช้และจำนวนค่าสังเกตที่เป็นค่าสูงมากที่พบในประชากร และพบว่ามีตัวแปรอื่น ที่มีความสัมพันธ์เชิงเส้นต่อกันกับตัวแปรที่เราสนใจศึกษาแล้วในกรณีที่ทราบ จำนวนค่าสังเกตที่เป็นค่าสูงมากในประชากร ควรจะใช้ตัวประมาณ $\hat{Y}_1$ และใช้ $\hat{Y}_3$ ในกรณีที่ไม่ทราบจำนวนค่าสังเกตที่เป็นค่าสูงมากในประชากร

3. เนื่องจากฟังก์ชั่นความหนาแน่น $p(n_1)$ ที่ใช้ประมาณค่าความน่าจะเป็นแบบ positive Hypergeometric distribution เมื่อประชากรมีขนาดใหญ่ จะมีปัญหาในการคำนวณหาค่าความน่าจะเป็นถ้าใช้ตัวอย่างขนาดใหญ่ ซึ่งในการวิจัยครั้งนี้ไม่ได้เสนอวิธีการประมาณค่าความน่าจะเป็นของฟังก์ชั่นความหนาแน่น $p(n_1)$ กรณีเมื่อใช้ขนาดตัวอย่างใหญ่ เพื่อเป็นการแก้ปัญหาดังกล่าว อาจจะทำการศึกษาวิธีการประมาณค่าความน่าจะเป็นของฟังก์ชั่นความหนาแน่น $p(n_1)$ ในการทำวิจัยครั้งต่อไป

4. ในการวิจัยครั้งนี้เป็นการศึกษาเฉพาะในกรณีที่มีบางหน่วยตัวอย่างที่มีค่าสูงมากและเป็นค่าที่มีอยู่จริงในประชากร ดังนั้นในการทำวิจัยครั้งต่อไปอาจจะศึกษาในกรณีที่มีบางหน่วยตัวอย่างที่มีค่าต่ำมากและเป็นค่าที่มีอยู่จริงในประชากร หรือศึกษาในกรณีที่มีบางหน่วยตัวอย่างที่มีค่าสูงมากหรือต่ำมากรวมอยู่ด้วย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย