รายงานวิจัย

เรื่อง

การกำหนดน้ำหนักใหม่ของตัวประมาณค่าที่ใช้อัตราส่วนรวมในการชักตัวอย่างแบบชั้นภูมิ

ผู้วิจัย

รศ.ดร.จิราวัลย์ จิตรถเวช

รศ.ดร.วิชิต หล่อจีระชุณห์กุล

เสนอต่อ

คณะกรรมการดำเนินงาน

กองทุนส่งเสริมและพัฒนางานวิจัย

มีนาคม 2555

กิตติกรรมประกาศ

งานวิจัยเรื่อง การกำหนดน้ำหนักใหม่ของตัวประมาณค่าที่ใช้อัตราส่วนรวมในการชักตัวอย่างแบบชั้นภูมิ ได้รับเงินทุนสนับสนุนการวิจัยจากคณะกรรมการดำเนินงานกองทุนส่งเสริมและพัฒนางานวิจัย สถาบันบัณฑิตพัฒนบริหารศาสตร์ ผู้วิจัยขอขอบคุณมา ณ โอกาสนี้ด้วย

รศ.ดร.จิราวัลย์ จิตรถเวช
รศ.ดร.วิชิต หล่อจีระชุณห์กุล

สารบัญ

สารบัญตาราง

ตารางที่ **หน้า**

บทที่ 1
บทนำ

1.1 ความสำคัญและความเป็นมา

การชักตัวอย่างแบบชั้นภูมิมีการนำไปใช้หลากหลาย ทั้งในทางวิทยาศาสตร์และทางสังคมศาสตร์ เป้าหมายในการชักตัวอย่างโดยวิธีนี้ก็เพื่อออกรายงานหรือข้อสรุปเกี่ยวกับประชากรย่อยหรือชั้นภูมิ และท้ายสุดก็เป็นการอนุมานไปยังประชากร โดยการประมาณค่าเฉลี่ยของประชากร ค่ายอดรวมของประชากร หรือสัดส่วนของประชากรของตัวแปรที่สนใจ การชักตัวอย่างแบบชั้นภูมิสามารถใช้วิธีการสุ่มที่แตกต่างกันในแต่ละชั้นภูมิได้ แต่ในที่นี้จะมีขอบเขตเฉพาะการชักตัวอย่างจากแต่ละชั้นภูมิเป็นการสุ่มอย่างง่ายทุกชั้นภูมิ ดังนั้น การประมาณค่าเฉลี่ยในแต่ละชั้นภูมิอาศัยค่าเฉลี่ยของตัวอย่าง ซึ่งเป็นตัวประมาณค่าที่ไม่เอนเอียงของชั้นภูมิ แล้วนำค่าเฉลี่ยของทุกชั้นภูมิมารวมกัน โดยใช้ค่าน้ำหนักของสัดส่วนของจำนวนหน่วยตัวอย่างของประชากรในแต่ละชั้นภูมิ หารด้วยจำนวนหน่วยตัวอย่างทั้งหมดในประชากรเป็นตัวถ่วงน้ำหนัก เพื่อประมาณค่าเฉลี่ยของประชากร ความถูกต้องของค่าประมาณค่าเฉลี่ยของประชากรจึงขึ้นอยู่กับความถูกต้องของค่าประมาณค่าเฉลี่ยในระดับชั้นภูมิและค่าน้ำหนักในการนำค่าเฉลี่ยของแต่ละชั้นภูมิมารวมกัน นักสถิติหลายคนได้พัฒนาวิธีการประมาณค่าเฉลี่ยของประชากร โดยการปรับค่าน้ำหนักให้เหมาะสมเพื่อให้ได้ค่าที่ใกล้เคียงกับค่าเฉลี่ยของประชากร โดยใช้น้ำหนักในการถ่วงหาค่าเฉลี่ยขึ้นอยู่กับสัดส่วนของหน่วยในชั้นภูมิ แต่ในบางกรณีจำนวนหน่วยตัวอย่างของประชากรในแต่ละชั้นภูมิและจำนวนหน่วยตัวอย่างทั้งหมดในประชากรผู้วิจัยอาจไม่ทราบค่าที่ถูกต้องหรือทราบค่าแต่เป็นข้อมูลที่ไม่ทันสมัยก็จะทำให้ตัวประมาณค่าเฉลี่ยของประชากรไม่ถูกต้อง

การสุ่มแบบชั้นภูมิเป็นวิธีการหนึ่งที่มีการแบ่งประชากรออกเป็นประชากรย่อยเรียกว่า ชั้นภูมิ ตามทฤษฎีในการแบ่งประชากรย่อย จะแบ่งโดยให้หน่วยตัวอย่างในประชากรย่อยมีลักษณะของหน่วยตัวอย่างคล้าย ๆ กัน แต่ระหว่างประชากรย่อยหรือชั้นภูมิจะให้หน่วยตัวอย่างในชั้นภูมิที่ต่างกันมีลักษณะที่แตกต่างกัน ในการแบ่งหน่วยตัวอย่างออกเป็นชั้นภูมินั้นตัวแปรที่สนใจจะเป็นตัวแปรที่ใช้ในการแบ่งชั้นภูมิดีที่สุด แต่ในความเป็นจริงอาจไม่สามารถหาข้อมูลเกี่ยวกับหน่วยตัวอย่างของตัวแปรที่สนใจได้ ผู้วิจัยส่วนใหญ่จึงมักอาศัยตัวแปรเสริมที่มีความสัมพันธ์กับตัวแปรที่สนใจในการศึกษามาช่วยในการแบ่งชั้นภูมิตามที่มีการศึกษาของครอกแครน (Cochran, 1977) เมื่อแบ่งชั้นภูมิเรียบร้อยแล้ว หลังจากนั้นจะมีการชักตัวอย่างในแต่ละชั้นภูมิด้วยวิธีสุ่มอย่างง่ายจำนวน n_h หน่วย จากประชากรย่อยที่มี N_h หน่วย การสุ่มนี้อาจจะกำหนดให้จำนวนหน่วยตัวอย่างเป็นสัดส่วนกับหน่วยตัวอย่างในประชากร นั่นคือ ถ้า

จำนวนหน่วยตัวอย่างของประชากรในชั้นภูมินั้นมีจำนวนมากก็จะทำการสุ่มจำนวนหน่วยตัวอย่างในชั้นภูมินั้นมาก นั่นคือ $n_h = \frac{N_h}{N} \cdot n$

เมื่อได้จำนวนหน่วยตัวอย่างในชั้นภูมิที่ h แล้วก็ทำการเก็บรวบรวมข้อมูล แต่ในบางกรณีไม่สามารถสุ่มตัวอย่างในลักษณะที่มีจำนวนหน่วยในตัวอย่างเป็นสัดส่วนกับจำนวนหน่วยในประชากรได้ เพราะจำนวนหน่วยตัวอย่างของประชากรในแต่ละชั้นภูมิแตกต่างกันมาก ทำให้หน่วยตัวอย่างในบางชั้นภูมินั้น มีจำนวนหน่วยตัวอย่างน้อยเกินไป ไม่เป็นตัวแทนที่ดีของหน่วยตัวอย่างในประชากรของชั้นภูมินั้น ในทางกลับกันอาจจะได้จำนวนหน่วยตัวอย่างมากเกินไปในชั้นภูมิที่มีจำนวนมาก ซึ่งไม่สามารถจะเก็บรวบรวมข้อมูลได้ อาจเนื่องจากทรัพยากรที่มีจำกัด ซึ่งในกรณีนี้อาจสุ่มตัวอย่างในแต่ละชั้นภูมิออกมาจำนวนเท่ากัน เพื่อให้แน่ใจว่าทุกชั้นภูมิจะมีหน่วยตัวอย่างที่เป็นตัวแทนที่ดีในระดับหนึ่งของหน่วยในประชากร เพื่อใช้ในการประมาณค่ากลับไปสู่ประชากร

การหาตัวประมาณค่าเฉลี่ย ยอดรวม หรือสัดส่วนของตัวแปรที่สนใจนั้น มีการประมาณโดยใช้สัดส่วนของจำนวนหน่วยตัวอย่างของประชากรชั้นภูมิ ต่อจำนวนหน่วยตัวอย่างทั้งหมดของประชากรเป็นค่าน้ำหนัก ซึ่งให้ผลดีในระดับหนึ่ง หรืออาจใช้การประมาณค่าแบบอัตราส่วนแยกหรือใช้อัตราส่วนรวม หรือการประมาณค่าแบบการถดถอย (Cochran, 1977) เพื่อเพิ่มความถูกต้องให้กับตัวประมาณค่าโดยใช้ตัวแปรเสริมช่วยในการแบ่งชั้นภูมิและการประมาณค่า

สำหรับในงานวิจัยนี้ จะทำการศึกษาหาน้ำหนักของการรวมค่าเฉลี่ยของชั้นภูมิทั้งหมดโดยไม่อาศัยจำนวนหน่วยตัวอย่างในชั้นภูมิ ทั้งนี้ เนื่องจากในความเป็นจริงจำนวนหน่วยตัวอย่างในชั้นภูมิอาจจะไม่ทราบแน่ชัดว่ามีจำนวนเท่ากับเท่าใด และความถูกต้องของการประมาณค่าเฉลี่ยของประชากรขึ้นอยู่กับความถูกต้องของจำนวนหน่วยในชั้นภูมิ การวิจัยนี้จะคำนวณหาน้ำหนักของการถ่วงค่าเฉลี่ยของแต่ละชั้นภูมิ โดยไม่ใช้จำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ แต่จะใช้ค่าของตัวสถิติในแต่ละชั้นภูมิ เพื่อประมาณค่าของน้ำหนักเพื่อใช้ในการประมาณค่าเฉลี่ยของประชากร โดยให้มีค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองระหว่างตัวประมาณค่าเฉลี่ยของประชากรกับค่าเฉลี่ยของประชากรมีค่าต่ำที่สุด เหตุผลที่ทำเช่นนี้เนื่องจากตัวสถิติของชั้นภูมิมีความเสถียรมากกว่าจำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ เมื่อมีการเปลี่ยนแปลงในจำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ นอกเสียจากว่าจำนวนหน่วยตัวอย่างที่เพิ่มขึ้นได้ทำให้มีค่านอกเกณฑ์ (outlier) เข้ามา การชักตัวอย่างแบบชั้นภูมิ โดยใช้ตัวแปรเสริมในการแบ่งชั้นภูมิและวิธีการประมาณค่าโดยใช้การประมาณค่าแบบอัตราส่วนรวมเป็นวิธีการที่นิยมใช้กันอย่างแพร่หลายทั้งทางวิทยาศาสตร์และสังคมศาสตร์

1.2 วัตถุประสงค์และขอบเขตของงานวิจัย

1.2.1 หาค่าน้ำหนักใหม่เพื่อใช้ถ่วงในการหาตัวประมาณค่าเฉลี่ย ยอดรวม และค่าสัดส่วนที่สนใจของประชากรเมื่อมีการสุ่มแบบชั้นภูมิอย่างง่าย โดยอาศัยค่าสถิติของตัวแปรช่วยไม่อิงกับจำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ

1.2.2 ศึกษาข้อจำกัดของการใช้ตัวประมาณค่าที่สร้างขึ้นในข้อ (1.2.1)

1.2.3 หาค่าความเอนเอียงและค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนตัวประมาณค่าที่นำเสนอในข้อ (1.2.1)

1.2.4 ศึกษาประสิทธิภาพของตัวประมาณค่าในข้อ (1.2.1) เปรียบเทียบกับตัวประมาณค่าที่ใช้กันแพร่หลายทั่วไปและสามารถเปรียบเทียบกันได้ โดยใช้การจำลองข้อมูลและใช้ข้อมูลจริงอย่างน้อย 2 ชุด

1.3 วิธีดำเนินการวิจัย

1.3.1 ทำการศึกษาตัวประมาณค่าแบบอัตราส่วนรวมต่าง ๆ ที่มีการนำเสนอไว้แล้ว เช่น ตัวประมาณค่าของ ซิสโซเดีย และดวายดิ (Sisodia and Dwivedi, 1981) และตัวประมาณค่าของคาดิลาร์ และซินจิ ในปี 2003 (Kadilar and Cingi, 2003)

1.3.2 พัฒนาตัวประมาณค่าเฉลี่ยของประชากร โดยให้ตัวประมาณค่าของ $\overline{x}_{st}$ ประมาณค่าเฉลี่ยประชากร $\overline{X}$ โดยมีค่าของ $MSE(\overline{x}_{st})$ มีค่าต่ำที่สุด และตัวประมาณค่าของ $\overline{y}_{st}$ ประมาณค่าเฉลี่ยประชากร $\overline{Y}$ มีค่า $MSE(\overline{y}_{st})$ มีค่าต่ำสุด

1.3.3 พัฒนาตัวประมาณค่ายอดรวมของประชากร และสัดส่วนของตัวประมาณค่าที่สนใจของประชากร

1.3.4 หาค่าความเอนเอียงของตัวประมาณค่า และค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนของตัวประมาณค่าที่นำเสนอ

1.3.5 ทำการจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่า โดยการจำลองข้อมูลจะกำหนดค่าสหสัมพันธ์ของตัวแปรช่วยx กับตัวแปรที่สนใจ y ให้มีระดับต่าง ๆ กัน เพื่อดูผลการประมาณค่าของตัวประมาณเทียบกับตัวประมาณค่าของอัตราส่วนรวมแบบคลาสสิค (classical combined ratio) กับตัวประมาณค่าของซิสโซเดีย และดวายดิ ที่นำเสนอในปี 2003 และตัวประมาณค่าของ คาดิลาร์ และซินจิที่เสนอในปี 2005

1.3.6 นำตัวประมาณค่าที่ได้ใช้กับข้อมูลจริงอย่างน้อย 2 ชุด เพื่อดูประสิทธิภาพเทียบกับตัวประมาณค่าที่ได้กล่าวไว้แล้ว

1.4 ประโยชน์ที่จะได้รับ

1.4.1 เป็นแนวคิดใหม่ในการประมาณค่าเฉลี่ยของประชากรโดยไม่ใช้จำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ ทำให้สามารถประมาณค่าได้เมื่อไม่ทราบจำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ การสอนในวิชาเทคนิคการชักตัวอย่างวิชาเทคนิคตัวแปรเดียวกับการประยุกต์และวิชาที่เกี่ยวข้องกับการชักตัวอย่างซึ่งเป็นวิชาในหลักสูตรสถิติ

1.4.2 สามารถนำไปใช้ในสาขาวิชาอื่น ๆ ได้ที่ต้องการประมาณค่าพารามิเตอร์และใช้วิธีการสุ่มแบบชั้นภูมิ

1.4.3 ทำเป็นบทความวิจัยเพื่อลงตีพิมพ์ในวารสารในระดับนานาชาติ

บทที่ 2

ทบทวนวรรณกรรม

2.1 ตัวประมาณค่าเฉลี่ย

เมื่อมีการชักตัวอย่างแบบชั้นภูมิ ตัวประมาณค่าเฉลี่ยของประชากรที่เป็นตัวประมาณค่าพื้นฐาน ได้แก่ ตัวประมาณค่าโดยใช้ค่าเฉลี่ยของตัวอย่างในแต่ละชั้นภูมินำมาถ่วงน้ำหนักด้วย $w_h = \frac{N_h}{N}$ เมื่อ N_h คือ จำนวนหน่วยตัวอย่างในประชากรที่ h และ N คือ จำนวนหน่วยตัวอย่างทั้งหมดในประชากร ตัวประมาณค่าเฉลี่ยของประชากร ได้แก่

$$\bar{y}_{st} = \sum_{h=1}^{k} w_h \bar{y}_h \tag{2.1}$$

เมื่อ $\bar{y}_h$ คือ ค่าเฉลี่ยของตัวอย่างในชั้นภูมิที่ h

w_h คือ ค่าถ่วงน้ำหนักในชั้นภูมิที่ h

k คือ จำนวนของชั้นภูมิ

ตัวประมาณค่าเฉลี่ย $\bar{y}_{st}$ เป็นตัวประมาณค่าที่ไม่เอนเอียงของค่าเฉลี่ยของประชากร $\bar{Y}$ และมีความแปรปรวนเท่ากับ

$$V(\bar{y}_{st}) = \sum_{h=1}^{k} w_h^2 \gamma_h S_{yh}^2 \tag{2.2}$$

เมื่อ $$\gamma_h = \frac{1}{n_h} - \frac{1}{N_h}$$

n_h คือ ขนาดของตัวอย่างในชั้นภูมิที่ h

และ S_{yh}^2 คือ ค่าความแปรปรวนของตัวแปร y ในชั้นภูมิที่ h

ตัวประมาณค่าเฉลี่ยของประชากรในสมการ (2.1) เป็นตัวประมาณค่าที่ไม่ได้ใช้ตัวแปรเสริมมาช่วยในการประมาณค่า แต่ในกรณีที่ตัวแปรที่สนใจมีสัมประสิทธิ์สหสัมพันธ์กับตัวแปรช่วยx สูงตามเงื่อนไข (2.3)

$$\rho_{yx} > \frac{C_x}{2C_y} \tag{2.3}$$

เมื่อ C_x และ C_y เป็นสัมประสิทธิของความแปรปรวนของตัวแปรเสริมและตัวแปรที่สนใจตามลำดับ การประมาณค่าควรจะใช้ตัวแปรเสริมเข้ามาช่วย เพื่อให้ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองมีค่าน้อยลง (Cochran, 1977) ซึ่งอาจจะใช้วิธีการของการถดถอยหรือใช้วิธีการของอัตราส่วนมาช่วยในการประมาณค่า ซึ่งก็อาจทำได้ใน 2 กรณี คือ การใช้อัตราส่วนแยกของแต่ละชั้นภูมิของ $R_h = \dfrac{\bar{Y}_h}{\bar{X}_h}$ มาช่วยในการประมาณค่าหรืออาจใช้อัตราส่วนรวมของประชากรมาช่วยในการประมาณค่าในกรณีนี้ R มีค่าเท่ากับ $\dfrac{\bar{Y}}{\bar{X}}$ ในงานวิจัยชิ้นนี้จะใช้อัตราส่วนรวมค่าเฉลี่ยของตัวแปรที่สนใจกับค่าเฉลี่ยของตัวแปรเสริมมาช่วยในการประมาณค่า ในกรณีที่ตัวแปรช่วยx มีสัมประสิทธิ์สหสัมพันธ์กับตัวแปรสมใจสูงกว่าค่าที่กำหนดตามเงื่อนไข (2.3) ตัวประมาณค่าดังกล่าวนี้เป็นตัวประมาณค่าที่มีความเอนเอียง แต่ค่าของความแปรปรวนจะต่ำกว่าตัวประมาณค่าที่ไม่ได้ใช้อัตราส่วน (Cochran, 1977)

ตัวประมาณค่าเฉลี่ยของประชากรในกรณีที่ใช้อัตราส่วนรวมมาช่วยในการประมาณค่าคือ

$$\bar{Y}_{CR} = \frac{\bar{y}_{st}}{\bar{x}_{st}}\bar{X} \tag{2.4}$$

เมื่อ $\bar{y}_{st} = \sum_{h=1}^{k} w_h \bar{y}_h$

$\bar{x}_{st} = \sum_{h=1}^{k} w_h \bar{x}_h$

และ $\bar{x}_h$ คือ ค่าเฉลี่ยของตัวแปรช่วยx ในชั้นภูมิ h

ค่าความเอนเอียงของตัวประมาณค่า $\bar{Y}_{CR}$ (Cochran, 1977; Cingi, 1950) มีค่าเท่ากับ

$$\text{Bias}(\bar{y}_{CR}) = -\frac{\text{Cov}(\bar{y}_{st}, \bar{x}_{st})}{\bar{X}} \tag{2.5}$$

และ

$$\text{MSE}(\bar{y}_{CR}) = \sum_{h=1}^{k} w_h^2 \gamma_h (S_{yh}^2 - 2RS_{yxh} + R^2 S_{xh}^2) \tag{2.6}$$

เมื่อ $R = \dfrac{\bar{Y}}{\bar{X}}$ เป็นอัตราส่วนระหว่างค่าเฉลี่ยของตัวแปรที่สนใจ y กับค่าเฉลี่ยของตัวแปรช่วยx

S_{yxh} คือ ค่าความแปรปรวนร่วมระหว่างตัวแปรที่สนใจ y กับตัวแปรช่วยx ในชั้นภูมิ h

และ S_{xh}^2 คือ ค่าความแปรปรวนของตัวแปรช่วยx ในชั้นภูมิ h

คาดิลาร์ และซินจิ ในปี 2003 (Kadilar and Cingi, 2003) ได้ทำการปรับปรุงตัวประมาณค่าที่ใช้อัตราส่วนรวมในการสุ่มอย่างง่ายของ ซิสโซเดีย และดวายดิ (Sisodia and Dwivedi, 1981) และของยูเพดยายา และสิงห์ (Upadhyaya and Singh, 1999) เพื่อใช้ในการชักตัวอย่างแบบชั้นภูมิ

ตัวประมาณค่าที่ปรับมาจากตัวประมาณค่าที่ซิสโวเดีย และดวายดินำเสนอ คือ

$$\bar{y}_{SD} = \bar{y}_{st}\frac{\bar{X}_{SD}}{\bar{x}_{SD}} \tag{2.7}$$

โดยมีค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองและความเอนเอียงของตัวประมาณค่าเท่ากับ

$$MSE(\bar{y}_{SD}) = \sum_{h=1}^{k} w_h^2\gamma_h\bar{y}_{st}(S_{yh}^2 - 2R_{SD}S_{yxh} + R_{SD}^2S_{xh}^2) \tag{2.8}$$

$$Bias(\bar{y}_{SD}) = \frac{1}{\bar{X}_{SD}}\sum_{h=1}^{k} w_h^2\gamma_h(R_{SD}S_{xh}^2 - S_{yxh}) \tag{2.9}$$

เมื่อ

$$\bar{x}_{SD} = \sum_{h=1}^{k} w_h(\bar{x}_h + C_{xh})$$

$$\bar{X}_{SD} = \sum_{h=1}^{k} w_h(\bar{X}_h + C_{xh})$$

$$R_{SD} = \frac{\bar{Y}}{\bar{X}_{SD}}$$

และ C_{xh} คือ ค่าสัมประสิทธิ์การกระจายในประชากรของตัวแปรช่วยx ในชั้นภูมิที่ h

ตัวประมาณค่าที่ปรับมาจากตัวประมาณค่าของยูเพดยายา และสิงห์ ได้แก่

$$\bar{y}_{us1} = \bar{y}_{st}\frac{\sum_{h=1}^{k} w_h\{\bar{X}_h\beta_{2h}(x) + C_{xh}\}}{\sum_{h=1}^{k} w_h\{\bar{x}_h\beta_{2h}(x) + C_{xh}\}} \tag{2.10}$$

และ

$$\bar{y}_{us2} = \bar{y}_{st}\frac{\sum_{h=1}^{k} w_h\{\bar{X}_hC_{xh} + \beta_{2h}(x)\}}{\sum_{h=1}^{k} w_h\{\bar{x}_hC_{xh} + \beta_{2h}(x)\}} \tag{2.11}$$

เมื่อ $\beta_{2h}(x)$ คือ ความโด่งของตัวแปรช่วยx ของประชากรในชั้นภูมิ h

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองและค่าความเอนเอียงของตัวประมาณค่า $\bar{y}_{us1}$ และ $\bar{y}_{us2}$ นั้นอยู่ในรูปแบบเดียวกันกับค่าเฉลี่ยความคลาดเคลื่อนและค่าความเอนเอียงของตัว

ประมาณค่า $\overline{y}_{SD}$ ใน (2.8) และ (2.9) แต่แทนค่า R_{SD} ด้วย R_{us1} และ R_{us2} แทนค่า $\overline{X}_{SD}$ ด้วย $\overline{X}_{us1}$ และ $\overline{X}_{us2}$ และแทนค่า $\overline{x}_{SD}$ ด้วย $\overline{x}_{us1}$ และ $\overline{x}_{us2}$ ตามลำดับ

ตัวประมาณค่าของ คาดิลาร์ และซินจิ (Kadilar and Cingi, 2005) เป็นตัวประมาณค่าที่ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\overline{y}_{KC}$ มีค่าต่ำสุด คือ

$$\overline{y}_{KC} = \kappa \frac{\overline{y}_{st}}{\overline{x}_{st}} \overline{X} \tag{2.12}$$

ค่าพารามิเตอร์ของ K มีค่าเท่ากับ

$$\kappa = \frac{\overline{Y}^2}{\overline{Y}^2 + \sum_{h=1}^{k} w_h^2 \gamma_h (S_{yh}^2 - 2RS_{yxh} + R^2 S_{xh}^2)} \tag{2.13}$$

ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนของตัวประมาณค่า $\overline{y}_{KC}$

$$MSE(\overline{y}_{KC}) = \kappa^2 \sum_{h=1}^{k} w_h^2 \gamma_h (S_{yh}^2 - 2RS_{yxh} + R^2 S_{xh}^2) + (\kappa - 1)^2 \overline{Y}^2 \tag{2.14}$$

ซึ่งตัวประมาณค่า $\overline{y}_{KC}$ จะมีค่า $MSE(\overline{y}_{KC})$ น้อยกว่า $MSE(\overline{y}_{CR})$ เสมอ เนื่องจากค่า κ เป็นค่าบวกที่มีค่าน้อยกว่า 1, $0 < \kappa < 1$ และค่าของ $\overline{y}_{KC}$ ก็จะมีค่าน้อยกว่า $\overline{y}_{CR}$ เสมอ ดังนั้นการหาตัวประมาณค่า $\overline{y}_{KC}$ คือ การหาค่าของ κ แล้วนำมาคูณกับตัวประมาณค่าตัวดั้งเดิมคือ $\overline{y}_{CR}$ (Kadilar and Cingi, 2005)

สิงห์และคณะ ในปี 2008 (Singh et al., 2008) ได้ใช้การแปลงค่าด้วยการยกกำลัง (power transformation) ไปปรับค่าตัวประมาณค่าเฉลี่ยที่ใช้อัตราส่วนรวมในการสุ่มแบบอย่างง่ายและการสุ่มแบบชั้นภูมิ เพื่อเพิ่มประสิทธิภาพสัมพัทธ์ของตัวประมาณค่า

$$\overline{y}_{STSK}(\alpha) = \overline{y}_{st} \left(\frac{X_{us1}}{x_{us1}} \right)^{\alpha} \tag{2.15}$$

$$\overline{y}_{STSK}(\delta) = \overline{y}_{st} \left(\frac{X_{us2}}{x_{us2}} \right)^{\delta} \tag{2.16}$$

เมื่อ

$$\alpha = \frac{\sum_{h=1}^{k} w_h^2 \gamma_h \beta_{2h}(x) S_{yxh}}{\sum_{h=1}^{k} w_h^2 \gamma_h \beta_{2h}^2(x) S_{xh}^2 \hat{R}_{us1}} \tag{2.17}$$

$$\delta = \frac{\sum_{h=1}^{k} w_h^2 \gamma_h C_{xh} S_{yxh}}{\sum_{h=1}^{k} w_h^2 \gamma_h C_{xh}^2 S_{xh}^2 \hat{R}_{us2}} \tag{2.18}$$

$$\hat{R}_{us1} = \frac{\sum_{h=1}^{k} w_h \bar{y}_h}{\sum_{h=1}^{k} w_h \left\{\bar{x}_h \beta_{2h}(x) + C_{xh}\right\}} \tag{2.19}$$

$$\hat{R}_{us2} = \frac{\sum_{h=1}^{k} w_h \bar{y}_h}{\sum_{h=1}^{k} w_h \left\{\bar{x}_h C_{xh} + \beta_{2h}(x)\right\}} \tag{2.20}$$

ตัวประมาณค่าที่กล่าวมาทั้งหมด เป็นตัวประมาณค่าที่ใช้น้ำหนักถ่วงจากการใช้ N_h/N ทั้งสิ้น

ในงานวิจัยนี้จะสร้างตัวประมาณค่าเฉลี่ยของประชากรใหม่ โดยการหาค่าของน้ำหนัก w_h ใหม่ โดยไม่อิงกับค่าของขนาดของประชากรในชั้นภูมิ แต่อิงกับค่าสถิติในชั้นภูมิ ซึ่งเมื่อมีการเปลี่ยนแปลงค่า N_h ค่าสถิติของชั้นภูมิมีความเสถียรมากกว่า โดยมีประสิทธิภาพใกล้เคียงหรือดีกว่าตัวประมาณค่าตัวเดิม เพื่อเป็นอีกทางเลือกหนึ่ง ในกรณีที่ขนาดของประชากรในชั้นภูมิ ทราบค่าโดยประมาณหรือมีการเปลี่ยนแปลงไป

2.2 วรรณกรรมเกี่ยวกับการแบ่งชั้นภูมิ

การแบ่งชั้นภูมิในการชักตัวอย่าง (sampling) มีวัตถุประสงค์หลายประการ คือ

1. ด้านการบริหารจัดการ หน่วยงานที่ทำการสำรวจอาจมีหน่วยงานในภาคสนามซึ่งกำกับดูแลการสำรวจในส่วนหนึ่ง ๆ ของประชากร จึงทำให้ต้องแบ่งประชากรออกเป็นชั้นภูมิตามหน้าที่ของหน่วยงานในภาคสนาม

2. ด้านประชากร ประชากรอาจไม่มีความเป็นเอกพันธ์ (homogeneity) ซึ่งอาจแบ่งออกได้เป็นประชากรย่อยที่มีความเป็นเอกพันธ์หลายประชากรย่อย เรียกว่า ชั้นภูมิ (stratum) เนื่องจากชั้นภูมิแต่ละชั้นมีความเป็นเอกพันธ์ ความแตกต่างระหว่างหน่วยในชั้นภูมิจึงมีน้อย ทำให้ได้ค่าประมาณที่แม่นยำของค่าเฉลี่ยชั้นภูมิ (stratum mean) จากตัวอย่างที่เล็กได้

3. ด้านการชักตัวอย่าง ปัญหาการชักตัวอย่างอาจแตกต่างกันมากในส่วนต่าง ๆ ของประชากร ในประชากรที่เกี่ยวกับคน คนที่อาศัยอยู่ตามสถาบันต่าง ๆ เช่น วัด โรงพยาบาล หอพัก เรือนจำ ฯลฯ มักจะถูกจัดให้อยู่ในชั้นภูมิที่แตกต่างไปจากคนที่อาศัยอยู่ในบ้านอยู่อาศัยธรรมดา เพราะเทคนิคการชักตัวอย่างจากชั้นภูมิบ้านอยู่อาศัยกับชั้นภูมิเรือนจำ ชั้นภูมิ

โรงพยาบาล ฯลฯ ย่อมแตกต่างกัน ในการชักตัวอย่างเกี่ยวกับธุรกิจ นิติบุคคลขนาดใหญ่มักจะถูกแยกให้อยู่คนละชั้นภูมิกับนิติบุคคลขนาดย่อม

การออกแบบการชักตัวอย่างสุ่มแบบชั้นภูมิ (stratified random sampling design) เป็นการวางแผนชักตัวอย่างในประชากร U ซึ่งแบ่งออกเป็น k ประชากรย่อยที่แยกขาดจากกัน ไม่มีการทับซ้อนระหว่างประชากรย่อย ที่เรียกว่า ชั้นภูมิ

$$U = \bigcup_{i=1}^{k} U_i \ ; \ U_i \bigcap U_j = \phi, \quad i \neq j \in \left[1, 2, \ldots k\right)$$

ในทัศนะของการอนุมาน การแบ่งชั้นภูมิมีวัตถุประสงค์หลักเพื่อทำให้ตัวประมาณค่า $\hat{\theta}$ มีความแม่นยำที่สุด หรืออีกนัยหนึ่งต้องการให้ค่า $\mathrm{MSE}(\hat{\theta})$ มีค่าน้อยที่สุด เมื่อ $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอียงเฉของ θ $\mathrm{MSE}(\hat{\theta})$ ก็คือ ความแปรปรวนของ $\hat{\theta}$

การออกแบบการชักตัวอย่างสุ่มแบบชั้นภูมิ ควรพิจารณาจำนวนชั้นภูมิที่ใช้ การจัดสรรตัวอย่างระหว่างชั้นภูมิ และการกำหนดขอบเขตของชั้นภูมิ ครอกแครน (Cochran, 1977) ได้พัฒนาตัวแบบที่แสดงถึงการลดลงโดยประมาณของความแปรปรวนของค่าเฉลี่ยตัวอย่างแบบชั้นภูมิ (stratified sample mean) เปรียบเทียบกับค่าเฉลี่ยจากการชักตัวอย่างสุ่มอย่างง่าย (simply random sampling) และได้สรุปว่า การลดลงของความแปรปรวนจะน้อยมากภายหลังที่มีจำนวนชั้นภูมิมากกว่า 6 ชั้น ในประเด็นการจัดสรรตัวอย่างระหว่างชั้นภูมินั้น เนย์แมน (Neyman, 1934) ได้เสนอวิธีการกระจายตัวอย่าง n หน่วย ระหว่างชั้นภูมิ k ชั้น เพื่อให้ความแปรปรวนของค่าเฉลี่ยตัวอย่างแบบชั้นภูมิมีค่าน้อยที่สุด โดยจัดสรรจำนวนตัวอย่าง n_h ให้แก่ชั้นภูมิ h แปรตาม $N_h S_{yh}$

$$n_h = n \frac{N_h S_{yh}}{\sum_{h=1}^{k} N_h S_{yh}} \tag{2.21}$$

ในทางปฏิบัติการจัดสรรตัวอย่างของเนย์แมนกระทำไม่ได้ เพราะตัวแปร y เป็นตัวแปรที่ยังไม่ทราบคุณลักษณะใด ๆ เลย เป็นตัวแปรที่กำลังศึกษาอยู่ แดลเนียส (Dalenius, 1950) ได้เสนอแนะให้จัดสรรตัวอย่างตามแนวคิดของเนย์แมน แต่ใช้คุณลักษณะของตัวแปรช่วยx ที่มีสหสัมพันธ์สูงกับตัวแปร y แทนและได้พิสูจน์ว่าความแปรปรวนของค่าเฉลี่ยตัวอย่างแบบชั้นภูมิ $\overline{x}_{st}$ มีค่าต่ำสุด เมื่อขอบเขตชั้นภูมิ b_h เป็นไปตาม (2.22)

$$\frac{S_{xh}^2 - (b_h - \overline{x}_h)^2}{S_{xh}^2} = \frac{S_{x(h+1)}^2 - (b_h - \overline{x}_{h+1})^2}{S_{x(h+1)}^2} \tag{2.22}$$

อย่างไรก็ตาม การกำหนดขอบเขตชั้นภูมิ b_h โดยใช้ (2.22) กระทำไม่ง่ายนัก เพราะ $\bar{x}_h$ และ S_{xh}^2 ขึ้นอยู่กับ b_h แดลเนียส และฮอดจ์ (Dalenius and Hodge, 1959) ได้พัฒนาการประมาณค่าจาก (2.22) โดยวิธีการจากความถี่สะสม (cumulative root frequency)

ให้ f(x) เป็นการแจกแจงของตัวแปรช่วย $x \in (-\infty, \infty)$

$$H = \int_{-\infty}^{\infty} \sqrt{f(x)}dx \tag{2.23}$$

$$C_K = \frac{H}{k} \tag{2.24}$$

ชั้นภูมิทั้ง k ชั้น จะมีขอบเขตดังนี้

$$(-\infty, \frac{H}{k}], (\frac{H}{k}, \frac{2H}{k}], \ldots, (\frac{k-1}{k}H, \infty)$$

จะเห็นได้ว่า

$$b_h = \int_{-\infty}^{x} \sqrt{f(x)}dc = \frac{hH}{k} \tag{2.25}$$

ครอกแครน (Cochran, 1961) ได้เสนอการประมาณเพื่อหาขอบเขตของชั้นภูมิตาม (2.25) โดยมีอัลกอริทึมดังนี้

1) นับความถี่ของตัวแปรที่อยู่ในแต่ละช่วงที่กำหนดขึ้น f_j, j = 1, 2, …, M M ควรมีค่ามากเพื่อที่จะได้ขอบเขตของชั้นภูมิที่ใกล้อุตตมะ

2) คำนวณ $\sqrt{f_j}$ และหาค่าสะสมของ $\sqrt{f_j}$

3) กำหนดจำนวนชั้นภูมิที่ต้องการ k

4) ขอบเขตของชั้นภูมิจะถูกกำหนดขึ้น โดยการรวมช่วงที่อยู่ติดกันเป็น k กลุ่ม หรือชั้นภูมิ เพื่อให้ $\Sigma\sqrt{f_j}$ ในแต่ละชั้นภูมิมีค่าใกล้เคียงกัน

เฮดลิน (Hedlin, 2000) ได้ตั้งข้อสังเกตว่า ขอบเขตของชั้นภูมิขึ้นอยู่กับการกำหนดค่า M และยังไม่มีทฤษฎีที่อาจใช้ในการกำหนดค่า M ได้

เอคแมน (Ekman, 1959) ได้เสนอการเลือกค่า b_h ที่ทำให้ความแปรปรวนของค่าเฉลี่ยตัวอย่างมีค่าต่ำที่สุด โดย

$$w_h(b_h - b_{h-1}) = C_K,\ h = 1, 2, \ldots, k \tag{2.26}$$

โดย $w_h = N_h/N$ และ C_K เป็นค่าคงที่ขึ้นอยู่กับจำนวนชั้นภูมิ k

ครอกแครน (Cochran, 1961) ได้ตั้งข้อสังเกตว่า $\sum_{h=1}^{k} w_h(b_h - b_{h-1})$ ไม่ใช่ค่าคงที่ แต่ขึ้นอยู่กับค่า k และ b_h และเฮดลิน (Hedlin, 2000) ยังชี้ให้เห็นว่า (2.26) หาคำตอบที่แท้จริงได้ยาก และแม้ว่าจะสามารถหาคำตอบโดยประมาณได้ แต่ก็มิอาจทราบได้ว่าจะมีคำตอบที่ดีกว่าคำตอบโดยประมาณที่พบหรือไม่

แดลเนียส และฮอดจ์ (Dalenius and Hodges, 1959) ได้สรุปว่า ในประชากรหลาย ๆ ประชากร ความแปรปรวนสัมพัทธ์ในแต่ละชั้นภูมิไม่ได้มีค่าแตกต่างกันมากนัก หากมีจำนวนชั้นภูมิมากและขอบเขตของชั้นภูมิมิได้กำหนดขึ้นอย่างมีเหตุผล

ลาเวลเล และฮิดิโรโกลส์ (Lavellée and Hidiroglous, 1988) ได้ตั้งข้อสังเกตสำหรับประชากรที่เบ้ว่า ค่าสัมประสิทธิ์ของการกระจายในระดับชั้นภูมิ มีแนวโน้มจะเท่ากัน ถ้าการจัดสรรตัวอย่างเป็นการจัดสรรอุตตมะ

กันนิง และฮอร์แกน (Gunning and Horgan, 2004) ได้ใช้ฐานคติการกระจายในชั้นภูมิเป็นเอกรูป แสดงให้เห็นว่า ถ้าค่าสัมประสิทธิ์ความแปรปรวนในแต่ละชั้นภูมิเท่ากัน จะทำให้ขอบเขตของชั้นภูมิเป็นไปตาม (2.26) และการแบ่งชั้นภูมิจะเป็นแบบเรขาคณิต (geometric stratification) โดย

$$b_h = b_0\gamma^h, \quad h = 0, 1, 2, \ldots, k \tag{2.27}$$

โดย b_0 และ b_h เป็นค่าต่ำสุดและค่าสูงสุดของตัวแปรช่วยx และ $\gamma = (b_k/b_0)^{1/k}$

ในการศึกษานี้ จะแบ่งชั้นภูมิตามข้อเสนอแนะของกันนิงและฮอร์แกน โดยจะปรับให้สัมประสิทธิของการกระจายของตัวแปรเสริมในระดับชั้นภูมิมีค่าใกล้เคียงกันโดยประมาณ

บทที่ 3
การพัฒนาค่าน้ำหนักของตัวประมาณค่า

การศึกษาในบทนี้เป็นการพัฒนาค่าน้ำหนักในการประมาณค่าเฉลี่ยของประชากรที่มาจากตัวสถิติของชั้นภูมิแทนจำนวนหน่วยของประชากรในชั้นภูมิ ซึ่งเป้าหมายหลักในการศึกษาเพื่อนำมาใช้ถ่วงน้ำหนักของค่าเฉลี่ยของตัวอย่างในแต่ละชั้นภูมิ โดยให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณค่ามีค่าต่ำที่สุด

3.1 การหาเกณฑ์ที่ใช้ในการประมาณค่า

การประมาณค่าเฉลี่ยของประชากรเมื่อมีการชักตัวอย่างแบบชั้นภูมิ คือ ตัวประมาณค่า $\bar{y}_{st}$ ซึ่งค่าความแปรปรวนของตัวประมาณค่า $\bar{y}_{st}$ มีค่าต่ำกว่าค่าความแปรปรวนของตัวประมาณค่า $\hat{\bar{Y}}_{CR}$ ที่ใช้วิธีการอัตราส่วนรวมและตัวแปรเสริมในการประมาณค่า ถ้าตัวแปรที่สนใจและตัวแปรเสริมมีสหสัมพันธ์ต่ำหรือมีค่าสัมประสิทธิ์สหสัมพันธ์ ρ_{yx} ไม่เป็นไปตามอสมการ

$$\rho_{yx} > \frac{C_x}{2C_y} \tag{3.1}$$

เมื่อ C_x และ C_y เป็นค่าสัมประสิทธิ์การกระจายในประชากรของตัวแปรช่วยx และตัวแปรที่สนใจ y ตามลำดับ (Cochran, 1977)

ทฤษฎีบทที่ 3.1 ในการชักตัวอย่างแบบชั้นภูมิตัวประมาณค่าพารามิเตอร์ $\bar{y}_{st}$ มีค่าความแปรปรวนมากกว่าตัวประมาณค่า $\hat{\bar{Y}}_{CR}$ เมื่อค่าของสัมประสิทธิ์สหสัมพันธ์ส่วนรวมของทุกชั้นภูมิ (combined correlation coefficient across all strata), $\rho_{c\bar{y}_{st}\bar{x}_{st}}$ มีค่าสอดคล้องกับอสมการ

$$\rho_{c\bar{y}_{st}\bar{x}_{st}} > \frac{C^2_{\bar{x}_{st}} + \left\{3-E(M_{\bar{x}_{st}})\right\}\left\{1-E(M_{\bar{x}_{st}})\right\}C^2_{\bar{y}_{st}}}{2C_{\bar{x}_{st}}C_{\bar{y}_{st}}\left\{2-E(M_{\bar{x}_{st}})\right\}} \tag{3.2}$$

เมื่อค่าสัมประสิทธิ์สหสัมพันธ์ส่วนรวม $\rho_{c\bar{y}_{st}\bar{x}_{st}}$ เท่ากับ (Singh and Vishwakarmam, 2008)

$$\rho_{c\bar{y}_{st}\bar{x}_{st}} = \frac{\sum_{h=1}^{k} w_h^2 \gamma_h \rho_h S_{yh} S_{xh}}{\sqrt{\sum_{h=1}^{k} w_h^2 \gamma_h S_{yh}^2 \sum_{h=1}^{k} w_h^2 \gamma_h S_{xh}^2}} \tag{3.3}$$

และ $$C_{\bar{y}_{st}} = \sqrt{\sum_{h=1}^{k} w_h^2 \gamma_h S_{yh}^2} \Big/ E(\bar{y}_{st}) \quad C_{\bar{x}_{st}} = \sqrt{\sum_{h=1}^{k} w_h^2 \gamma_h S_{xh}^2} \Big/ E(\bar{x}_{st})$$

และ $M_{\bar{x}_{st}} = \dfrac{\bar{x}_{st}}{\bar{X}}$ ทั้งนี้ โดยมิได้สูญเสียความเป็นทั่วไปของผลลัพธ์แต่อย่างใด ซึ่งสามารถตั้งฐานคติให้ $C_{\bar{y}_{st}}$ และ $C_{\bar{x}_{st}}$ มีค่าเป็นบวก

พิสูจน์

จาก (2.4) $$\hat{\bar{Y}}_{CR} = \frac{\bar{y}_{st}}{\bar{x}_{st}} \bar{X}$$

ดังนั้น $$\hat{\bar{Y}}_{CR} = \frac{\bar{y}_{st}}{\dfrac{\bar{x}_{st}}{\bar{X}}} = \frac{\bar{y}_{st}}{1 + \dfrac{\bar{x}_{st} - \bar{X}}{\bar{X}}}$$

จากการประมาณค่าด้วยอนุกรมเทเลอร์ (Taylor's series) องศาที่ 1 ภายใต้ฐานคติ $\dfrac{\bar{x}_{st} - \bar{X}}{\bar{X}} << 1.$ จะได้

$$\hat{\bar{Y}}_{CR} \cong \bar{y}_{st}\left(1 - \frac{\bar{x}_{st} - \bar{X}}{\bar{X}}\right) \cong \bar{y}_{st}\left(2 - \frac{\bar{x}_{st}}{\bar{X}}\right)$$

ความแปรปรวนของตัวประมาณค่า $\hat{\bar{Y}}_{CR}$ จึงเขียนได้เป็น

$$V(\hat{\bar{Y}}_{CR}) = \left\{E\left(\frac{\bar{y}_{st}}{\bar{X}}\right)\right\}^2 V(\bar{x}_{st}) + \left\{E\left(2 - \frac{\bar{x}_{st}}{\bar{X}}\right)\right\}^2 V(\bar{y}_{st})$$

$$+ 2E(\bar{y}_{st})E\left(2 - \frac{\bar{x}_{st}}{\bar{X}}\right)Cov\left(\bar{y}_{st}, 2 - \frac{\bar{x}_{st}}{\bar{X}}\right) \tag{3.4}$$

พิจารณา

$$Cov\left(\bar{y}_{st}, 2 - \frac{\bar{x}_{st}}{\bar{X}}\right) = E\left\{\bar{y}_{st}\left(2 - \frac{\bar{x}_{st}}{\bar{X}}\right)\right\} - E(\bar{y}_{st})E\left(2 - \frac{\bar{x}_{st}}{\bar{X}}\right)$$

$$= 2E(\bar{y}_{st}) - \frac{E(\bar{y}_{st}\bar{x}_{st})}{\bar{X}} - 2E(\bar{y}_{st}) + \frac{E(\bar{y}_{st})E(\bar{x}_{st})}{\bar{X}}$$

$$= -\left\{\frac{E(\bar{y}_{st}\bar{x}_{st})}{\bar{X}} - \frac{E(\bar{y}_{st})E(\bar{x}_{st})}{\bar{X}}\right\}$$

$$= -\frac{1}{\bar{X}}\left\{E(\bar{y}_{st}\bar{x}_{st}) - E(\bar{y}_{st})E(\bar{x}_{st})\right\}$$

$$= -\frac{1}{\bar{X}}\mathrm{Cov}(\bar{y}_{st}, \bar{x}_{st})$$

$$= -\frac{1}{\bar{X}}\rho_{c\bar{y}_{st}\bar{x}_{st}}\sigma_{\bar{y}_{st}}\sigma_{\bar{x}_{st}} \tag{3.5}$$

เมื่อนำผลของ $\mathrm{Cov}\left(\bar{y}_{st}, 2-\frac{\bar{x}_{st}}{\bar{X}}\right)$ ใน (3.5) ไปแทนค่า $V(\hat{\bar{Y}}_{CR})$ ใน (3.4) จะได้

$$V(\hat{\bar{Y}}_{CR}) = \left\{E\left(\frac{\bar{y}_{st}}{\bar{X}}\right)\right\}^2 V(\bar{x}_{st}) + \left\{\left(2 - E\frac{(\bar{x}_{st})}{\bar{X}}\right)\right\}^2 V(\bar{y}_{st})$$

$$-\frac{2E(\bar{y}_{st})\left\{2 - E\left(\frac{\bar{x}_{st}}{\bar{X}}\right)\right\}\rho_{c\bar{y}_{st}\bar{x}_{st}}\sigma_{\bar{y}_{st}}\sigma_{\bar{x}_{st}}}{\bar{X}}$$

$$= \frac{\{E(\bar{y}_{st})\}^2}{\bar{X}^2}V(\bar{x}_{st}) + \left\{2 - \frac{E(\bar{x}_{st})}{\bar{X}}\right\}^2 V(\bar{y}_{st})$$

$$-\frac{2E(\bar{y}_{st})}{\bar{X}}\left\{2 - \frac{E(\bar{x}_{st})}{\bar{X}}\right\}\rho_{c\bar{y}_{st}\bar{x}_{st}}\sigma_{\bar{y}_{st}}\sigma_{\bar{x}_{st}}$$

หารทั้ง 2 ข้างด้วย $V(\bar{y}_{st})$ จะได้

$$\frac{V(\hat{\bar{Y}}_{CR})}{V(\bar{y}_{st})} = \frac{\{E(\bar{y}_{st})\}^2}{\bar{X}^2}\frac{V(\bar{x}_{st})}{V(\bar{y}_{st})} + \left\{2 - E(M_{\bar{x}_{st}})\right\}^2$$

$$-2\frac{E(\bar{y}_{st})}{\bar{X}V(\bar{y}_{st})}\left\{2 - E(M_{\bar{x}_{st}})\right\}\rho_{c\bar{y}_{st}\bar{x}_{st}}\sigma_{\bar{y}_{st}}\sigma_{\bar{x}_{st}}$$

$$= \left(\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\right)^2 + \left\{2 - E(M_{\bar{x}_{st}})\right\}^2 - 2\left\{2 - E(M_{\bar{x}_{st}})\right\}\rho_{c\bar{y}_{st}\bar{x}_{st}}\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}$$

โดย $\quad C_{\bar{y}_{st}} = \frac{\sigma_{\bar{y}_{st}}}{E(\bar{y}_{st})}, C_{\bar{x}_{st}} = \frac{\sigma_{\bar{x}_{st}}}{E(\bar{x}_{st})}$

จัดพจน์ของอัตราส่วน $\frac{V(\hat{\bar{Y}}_{CR})}{V(\bar{y}_{st})}$ ใหม่จะได้

$$\frac{V(\hat{\bar{Y}}_{CR})}{V(\bar{y}_{st})} = \left(\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\right)^2 + \left\{2 - E(M_{\bar{x}_{st}})\right\}^2 - 2\left\{2 - E(M_{\bar{x}_{st}})\right\}\left(\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\cdot\rho_{c\bar{y}_{st}\bar{x}_{st}}\right)$$

$$\frac{V(\hat{\hat{Y}}_{CR})}{V(\bar{y}_{st})}-1=\left(\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\right)^2+\left\{2-E(M_{\bar{x}_{st}})\right\}^2-1-2\left\{2-E(M_{\bar{x}_{st}})\right\}\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\cdot\rho_{c\bar{y}_{st}\bar{x}_{st}}$$

$$\frac{V(\hat{\hat{Y}}_{CR})-V(\bar{y}_{st})}{V(\bar{y}_{st})}=\left(\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\right)^2+\left\{1-E(M_{\bar{x}_{st}})\right\}\left\{3-E(M_{\bar{x}_{st}})\right\}$$

$$-2\left\{2-E(M_{\bar{x}_{st}})\right\}\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\rho_{c\bar{y}_{st}\bar{x}_{st}}$$

ถ้า $\rho_{c_{\bar{y}_{st}\bar{x}_{st}}} > \dfrac{C^2_{\bar{x}_{st}}+\left\{1-E(M_{\bar{x}_{st}})\right\}\left\{3-E(M_{\bar{x}_{st}})\right\}C^2_{\bar{y}_{st}}}{2C_{\bar{x}_{st}}C_{\bar{y}_{st}}\left\{2-E(M_{\bar{x}_{st}})\right\}}$

$$> \frac{C_{\bar{y}_{st}}}{C_{\bar{x}_{st}}}\frac{\left(\frac{C_{\bar{x}_{st}}}{C_{\bar{y}_{st}}}\right)^2+\left\{1-E(M_{\bar{x}_{st}})\right\}\left\{3-E(M_{\bar{x}_{st}})\right\}}{2\left\{2-E(M_{\bar{x}_{st}})\right\}}$$

จะทำให้ $V(\bar{y}_{st}) > V(\hat{\hat{Y}}_{CR})$

ช.ต.พ.

จะเห็นได้ว่า เมื่อค่าเฉลี่ยของตัวแปรเสริม, $\bar{x}_{st}$ เข้าใกล้ค่าเฉลี่ยของประชากร, $\bar{X}$ ค่าที่อยู่ทางขวามือของอสมการ (3.2) จะเข้าใกล้ค่า $C_{\bar{x}_{st}}/(2C_{\bar{y}_{st}})$ ซึ่งอยู่ในรูปแบบคล้ายคลึงกับอสมการ (3.1) ค่าทางด้านขวามือของอสมการ (3.2) สามารถประมาณค่าได้โดยการใช้ความแปรปรวนของตัวอย่างของตัวแปรที่สนใจ s^2_{yh} แทนค่าของความแปรปรวนของประชากร S^2_{yh} และในกรณีที่ไม่ทราบค่า S^2_{xh} ก็อาจใช้ความแปรปรวนของตัวอย่าง s^2_{xh} แทนได้ ซึ่งจะได้ว่า

$$\frac{\sum_{h=1}^{k} w_h^2\gamma_h s_{xh}^2\Big/\bar{x}_{st}+(3-\bar{x}_{st}/\bar{X})(1-\bar{x}_{st}/\bar{X})\sum_{h=1}^{k} w_h^2\gamma_h s_{xh}^2\Big/\bar{y}_{st}}{2\left(\sqrt{\sum_{h=1}^{k} w_h^2\gamma_h s_{xh}^2\Big/\bar{x}_{st}}\right)\left(\sqrt{\sum_{h=1}^{k} w_h^2\gamma_h s_{xh}^2\Big/\bar{y}_{st}}\right)(2-\bar{x}_{st}/\bar{X})} \qquad (3.6)$$

และควรจะใช้ค่านี้ในกรณีที่ขนาดของตัวอย่างมีขนาดเล็ก ดังนั้น ถ้าอสมการ (3.2) ไม่เป็นจริง ตัวประมาณค่า $\bar{y}_{st}$ จะถูกใช้ในการประมาณค่า $\bar{Y}$ แทน $\hat{\hat{Y}}_{CR}$

3.2 การพัฒนาค่าน้ำหนักของตัวประมาณค่า

ค่าน้ำหนักใหม่ที่ใช้ในการคำนวณค่า $\bar{y}_{st}$ ที่นำเสนอไม่ได้อาศัยจำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ แต่จะเป็นค่าน้ำหนักที่ทำให้ค่าเฉลี่ยกำลังสองของค่าความคลาดเคลื่อนของ $\bar{y}_{st}$ มีค่าต่ำที่สุด อย่างไรก็ตาม การหาค่าต่ำที่สุดไม่สามารถคำนวณได้เพราะไม่ทราบค่าเฉลี่ยของประชากร แต่ถ้าตัวแปรช่วยx มีค่าสัมประสิทธิ์สหสัมพันธ์สูงกับตัวแปรที่สนใจ y น้ำหนักที่ทำให้ค่าเฉลี่ยกำลังสองของค่าความคลาดเคลื่อนของ $\bar{x}_{st}$ มีค่าต่ำที่สุด อาจใช้แทนน้ำหนักที่ทำให้ค่าเฉลี่ยกำลังสองของค่าความคลาดเคลื่อนของ $\bar{y}_{st}$ มีค่าต่ำที่สุดได้

การประมาณค่าดังกล่าวข้างต้นมีแนวคิดคล้ายคลึงกับการจัดสรรขนาดตัวอย่างของเนย์แมน (Neyman) โดยอาศัยค่าความแปรปรวนของตัวแปรช่วยx ที่ทราบค่าแทนค่าความแปรปรวนของตัวแปรที่สนใจ y ซึ่งไม่ทราบค่า (Dalenius, 1950) ดังนั้น การศึกษาในครั้งนี้จึงนำเสนอน้ำหนักของชั้นภูมิที่ใช้ในการประมาณค่าเฉลี่ยของประชากรจากค่าสถิติของชั้นภูมิ ซึ่งค่าน้ำหนักเหล่านี้จะมีความเสถียรกว่าการใช้จำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิเมื่อขนาดของหน่วยตัวอย่างของประชากรในชั้นภูมิมีการเปลี่ยนแปลง

น้ำหนักใหม่ที่ใช้ในการถ่วงน้ำหนักค่าเฉลี่ยของตัวอย่างในชั้นภูมิ เพื่อประมาณค่าเฉลี่ยของประชากร สามารถทำได้โดยการหาค่าต่ำสุดที่มีข้อจำกัด นั่นคือ

$$\min_{\substack{w_{xh} \\ 1\le h\le k}} MSE(\hat{\bar{X}}_{wx}) = \min_{\substack{w_{xh} \\ 1\le h\le k}} E(\sum_{h=1}^{k} w_{xh}\bar{x}_h - \bar{X})^2 \tag{3.7}$$

ภายใต้ข้อจำกัด $\sum_{h=1}^{k} w_{xh} = 1$ และ $w_{xh} \ge 0\ ;\ h = 1, 2, \ldots, k$

ค่า w_{xh}^* ซึ่งเป็นผลเฉลยของปัญหาการหาค่าต่ำสุดที่มีข้อจำกัดใน (3.7) เป็นฟังก์ชันของตัวสถิติในชั้นภูมิของตัวแปรช่วยx ในทำนองเดียวกัน น้ำหนักของตัวประมาณค่า $\hat{\bar{Y}}_{wy} = \bar{y}_{st\_wy} = \sum_{h=1}^{k} w_{yh}\bar{y}_h$ คือ น้ำหนักที่ทำให้ค่าความคลาดเคลื่อนของ $\hat{\bar{Y}}_{wy}$ มีค่าต่ำสุด นั่นคือ

$$\min_{\substack{w_{yh} \\ 1\le h\le k}} MSE(\hat{\bar{Y}}_{wy}) = \min_{\substack{w_{yh} \\ 1\le h\le k}} E(\sum_{h=1}^{k} w_{yh}\bar{y}_h - \bar{Y})^2 \tag{3.8}$$

ภายใต้ข้อจำกัด $\sum_{h=1}^{k} w_{yh} = 1$ และ $w_{yh} \ge 0\ ;\ h = 1, 2, \ldots, k$

การประมาณค่าของน้ำหนัก, w_{yh}^* ซึ่งเป็นผลเฉลยของปัญหาการหาค่าต่ำสุดที่มีข้อจำกัดใน (3.8) มิอาจทำได้นอกจากต้องทำการประมาณค่าเฉลี่ยของประชากรของตัวแปรที่สนใจ, y ก่อน

ทฤษฎีบทที่ 3.2 ในการชักตัวอย่างสุ่มแบบชั้นภูมิที่ประชากรแบ่งออกเป็น k ชั้นภูมิ ซึ่งมีจำนวนหน่วยตัวอย่างของประชากรในชั้นภูมิ h เท่ากับ N_h และในแต่ละชั้นภูมิใช้วิธีการชักตัวอย่างสุ่มแบบง่ายโดยวิธีการไม่แทนที่ ให้ n_h เป็นขนาดตัวอย่างในชั้นภูมิ h น้ำหนักของชั้นภูมิที่เป็นผลเฉลยจากการแก้ปัญหา (3.7) และ (3.8) เท่ากับ

$$w_{xh}^* = \frac{E(\bar{x}_h)}{V(\bar{x}_h)} \frac{\frac{1}{E(\bar{x}_h)}\left[1+\sum_{j=1}^{k}\frac{\{E(\bar{x}_j)\}^2}{V(\bar{x}_j)}\right]+\bar{X}\sum_{j=1}^{k}\frac{1}{V(\bar{x}_j)}-\left\{1+\frac{\bar{X}}{E(\bar{x}_h)}\right\}\sum_{j=1}^{k}\frac{E(\bar{x}_j)}{V(\bar{x}_j)}}{\sum_{j=1}^{k}\frac{1}{V(\bar{x}_j)}\left[1+\sum_{j=1}^{k}\frac{\{E(\bar{x}_j)\}^2}{V(\bar{x}_j)}\right]-\left\{\sum_{j=1}^{k}\frac{E(\bar{x}_j)}{V(\bar{x}_j)}\right\}^2} \tag{3.9}$$

$$w_{yh}^* = \frac{E(\bar{y}_h)}{V(\bar{y}_h)} \frac{\frac{1}{E(\bar{y}_h)}\left[1+\sum_{j=1}^{k}\frac{\{E(\bar{y}_j)\}^2}{V(\bar{y}_j)}\right]+\bar{Y}\sum_{j=1}^{k}\frac{1}{V(\bar{y}_j)}-\left\{1+\frac{\bar{Y}}{E(\bar{y}_h)}\right\}\sum_{j=1}^{k}\frac{E(\bar{y}_j)}{V(\bar{y}_j)}}{\sum_{j=1}^{k}\frac{1}{V(\bar{y}_j)}\left[1+\sum_{j=1}^{k}\frac{\{E(\bar{y}_j)\}^2}{V(\bar{y}_j)}\right]-\left\{\sum_{j=1}^{k}\frac{E(\bar{y}_j)}{V(\bar{y}_j)}\right\}^2} \tag{3.10}$$

ตามลำดับ

เมื่อ w_{xh} มีค่าเป็นลบ สำหรับ $h \in S_x$ และ w_{yh}^* มีค่าเป็นลบ สำหรับ $h \in S_y$ ให้คำนวณ w_{xh}^* และ w_{yh}^* ใหม่ ดังนี้

$$\underset{h \notin S_x}{w_{xh}^*} = \frac{E(\bar{x}_h)}{V(\bar{x}_h)} \frac{\frac{1}{E(\bar{x}_h)}\left[1+\sum_{\substack{j=1\\ j\notin S_x}}^{k}\frac{\{E(\bar{x}_j)\}^2}{V(\bar{x}_j)}\right]+\bar{X}\sum_{\substack{j=1\\ j\notin S_x}}^{k}\frac{1}{V(\bar{x}_j)}-\left\{1+\frac{\bar{X}}{E(\bar{x}_h)}\right\}\sum_{\substack{j=1\\ j\notin S_x}}^{k}\frac{E(\bar{x}_j)}{V(\bar{x}_j)}}{\sum_{\substack{j=1\\ j\notin S_x}}^{k}\frac{1}{V(\bar{x}_j)}\left[1+\sum_{\substack{j=1\\ j\notin S_x}}^{k}\frac{\{E(\bar{x}_j)\}^2}{V(\bar{x}_j)}\right]-\left\{\sum_{\substack{j=1\\ j\notin S_x}}^{k}\frac{E(\bar{x}_j)}{V(\bar{x}_j)}\right\}^2}$$

$$\underset{h \in S_x}{w_{xh}^*} = 0$$

$$w^*_{\substack{yh \\ h\notin S_y}} = \frac{E(\bar{y}_h)}{V(\bar{y}_h)} \frac{\frac{1}{E(\bar{y}_h)}\left[1+\sum\limits_{\substack{j=1\\ j\notin S_y}}^{k}\frac{\{E(\bar{y}_j)\}^2}{V(\bar{y}_j)}\right] + \bar{Y}\sum\limits_{\substack{j=1\\ j\notin S_y}}^{k}\frac{1}{V(\bar{y}_j)} - \left\{1+\frac{\bar{Y}}{E(\bar{y}_h)}\right\}\sum\limits_{\substack{j=1\\ j\notin S_y}}^{k}\frac{E(\bar{y}_j)}{V(\bar{y}_j)}}{\sum\limits_{\substack{j=1\\ j\notin S_y}}^{k}\frac{1}{V(\bar{y}_j)}\left[1+\sum\limits_{\substack{j=1\\ j\notin S_y}}^{k}\frac{\{E(\bar{y}_j)\}^2}{V(\bar{y}_j)}\right] - \left\{\sum\limits_{\substack{j=1\\ j\notin S_y}}^{k}\frac{E(\bar{y}_j)}{V(\bar{y}_j)}\right\}^2}$$

$$w^*_{\substack{yh \\ h\in S_y}} = 0$$

พิสูจน์ Lagrangian ในการแก้ปัญหาต่ำสุดภายใต้ข้อจำกัด เพื่อหาค่าน้ำหนักของ w_{xh} เขียนได้ดังนี้

$$L = E(\sum_{h=1}^{k} w_h\bar{x}_h - \bar{X})^2 + \lambda(\sum_{h=1}^{k} w_h - 1) - \sum_{h=1}^{k}\nu_h w_h \tag{3.11}$$

$$= \sum_{h=1}^{k} w_{xh}^2 V(\bar{x}_h) + \left\{E(\sum_{h=1}^{k} w_{xh}\bar{x}_h - \bar{X})\right\}^2 + \lambda\sum_{h=1}^{k}(x_{wh} - 1) - \sum_{h=1}^{k}\nu_h w_{xh}$$

$$\frac{\partial L}{\partial w_j} = 2w_j V(\bar{x}_j) + 2\left\{E(\sum_{h=1}^{k} w_{xh}\bar{y}_h) - \bar{X}\right\}E(\bar{x}_j) + \lambda - \nu_j = 0, \quad j = 1, 2, \ldots, k$$

$$w_j V(\bar{x}_j) + E(\bar{x}_j)\sum_{h=1}^{k} w_{xh}E(\bar{x}_h) = \bar{X}E(\bar{x}_j) - \frac{\lambda}{2} + \frac{\nu_j}{2}$$

ซึ่งสามารถเขียนในรูปของเมทริกซ์ได้ดังนี้

$$(U + \mathbf{zz}')\mathbf{w} = \bar{X}\mathbf{z} - \frac{\lambda}{2}\mathbf{1} + \frac{\boldsymbol{\nu}}{2} \tag{3.12}$$

การหาค่าต่ำสุดของ Lagrangian ภายใต้เงื่อนไขของ KKT (Karush-Kuhn-Tucker) ประกอบด้วย (3.12) และ

$$\mathbf{1}'\mathbf{w}^*_x = 1 \tag{3.13}$$

$$\nu_j w^*_{xj} = 0, j=1,2,\ldots,k \tag{3.14}$$

เมื่อ **1** เป็นเวคเตอร์ขนาด $k \times 1$ ที่มีสมาชิกทุกตัวเป็น 1

$$U = \begin{bmatrix} V(\bar{x}_1) & 0 & \cdots & 0 \\ 0 & V(\bar{x}_2) & \cdots & \vdots \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & V(\bar{x}_k) \end{bmatrix}, \ \mathbf{z} = \begin{bmatrix} E(\bar{x}_1) \\ E(\bar{x}_2) \\ \vdots \\ E(\bar{x}_k) \end{bmatrix} \ \text{และ} \ \nu = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_k \end{bmatrix}$$

การหาเมทริกซ์ผกผันของ $(\mathrm{U} + \mathbf{zz}')$ สามารถหาได้ดังนี้ (Henderson และ Searle, 1981)

$$(\mathrm{U} + \mathbf{zz}')^{-1} = \mathrm{U}^{-1} - \frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}} \tag{3.15}$$

$$\mathrm{U}^{-1} = \begin{bmatrix} \frac{1}{\mathrm{V}(\overline{\mathrm{x}}_1)} & 0 & \cdots & 0 \\ 0 & \frac{1}{\mathrm{V}(\overline{\mathrm{x}}_2)} & \cdots & \vdots \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\mathrm{V}(\overline{\mathrm{x}}_k)} \end{bmatrix}$$

$$\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1} = \begin{bmatrix} \left\{\frac{\mathrm{E}(\overline{\mathrm{x}}_1)}{\mathrm{V}(\overline{\mathrm{x}}_1)}\right\}^2 & \frac{\mathrm{E}(\overline{\mathrm{x}}_1)\mathrm{E}(\overline{\mathrm{x}}_2)}{\mathrm{V}(\overline{\mathrm{x}}_1)\mathrm{V}(\overline{\mathrm{x}}_2)} & \cdots & \frac{\mathrm{E}(\overline{\mathrm{x}}_1)\mathrm{E}(\overline{\mathrm{x}}_k)}{\mathrm{V}(\overline{\mathrm{x}}_1)\mathrm{V}(\overline{\mathrm{x}}_k)} \\ \frac{\mathrm{E}(\overline{\mathrm{x}}_2)\mathrm{E}(\overline{\mathrm{x}}_1)}{\mathrm{V}(\overline{\mathrm{x}}_2)\mathrm{V}(\overline{\mathrm{x}}_1)} & \left\{\frac{\mathrm{E}(\overline{\mathrm{x}}_2)}{\mathrm{V}(\overline{\mathrm{x}}_2)}\right\}^2 & \cdots & \frac{\mathrm{E}(\overline{\mathrm{x}}_2)\mathrm{E}(\overline{\mathrm{x}}_k)}{\mathrm{V}(\overline{\mathrm{x}}_2)\mathrm{V}(\overline{\mathrm{x}}_k)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathrm{E}(\overline{\mathrm{x}}_k)\mathrm{E}(\overline{\mathrm{x}}_1)}{\mathrm{V}(\overline{\mathrm{x}}_k)\mathrm{V}(\overline{\mathrm{x}}_1)} & \frac{\mathrm{E}(\overline{\mathrm{x}}_k)\mathrm{E}(\overline{\mathrm{x}}_2)}{\mathrm{V}(\overline{\mathrm{x}}_k)\mathrm{V}(\overline{\mathrm{x}}_2)} & \cdots & \left\{\frac{\mathrm{E}(\overline{\mathrm{x}}_k)}{\mathrm{V}(\overline{\mathrm{x}}_k)}\right\}^2 \end{bmatrix} \tag{3.16}$$

$$\mathbf{z}'\mathrm{U}^{-1}\mathbf{z} = \sum_{h=1}^{k} \left\{\frac{\mathrm{E}(\overline{\mathrm{x}}_h)}{\mathrm{V}(\overline{\mathrm{x}}_h)}\right\}^2 \tag{3.17}$$

ดังนั้น

$$(\mathrm{U} + \mathbf{zz}')^{-1} = \begin{bmatrix} \frac{1}{\mathrm{V}(\overline{\mathrm{x}}_1)} & 0 & \ldots & 0 \\ 0 & \frac{1}{\mathrm{V}(\overline{\mathrm{x}}_2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{\mathrm{V}(\overline{\mathrm{x}}_k)} \end{bmatrix}$$

$$-\frac{1}{1+\sum_{h=1}^{k}\frac{\{E(\bar{x}_h)\}^2}{V(\bar{x}_h)}}\begin{bmatrix} \left\{\frac{E(\bar{x}_1)}{V(\bar{x}_1)}\right\}^2 & \frac{E(\bar{x}_1)E(\bar{x}_2)}{V(\bar{x}_1)V(\bar{x}_2)} & \cdots & \frac{E(\bar{x}_1)E(\bar{x}_k)}{V(\bar{x}_1)V(\bar{x}_k)} \\ \frac{E(\bar{x}_2)E(\bar{x}_1)}{V(\bar{x}_2)V(\bar{x}_1)} & \left\{\frac{E(\bar{x}_2)}{V(\bar{x}_2)}\right\}^2 & \cdots & \frac{E(\bar{x}_2)E(\bar{x}_k)}{V(\bar{x}_2)V(\bar{x}_k)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{E(\bar{x}_k)E(\bar{x}_1)}{V(\bar{x}_k)V(\bar{x}_1)} & \frac{E(\bar{x}_k)E(\bar{x}_2)}{V(\bar{x}_k)V(\bar{x}_2)} & \cdots & \left\{\frac{E(\bar{x}_k)}{V(\bar{x}_k)}\right\}^2 \end{bmatrix} \tag{3.18}$$

คูณ (3.12) ด้วยเมทริกซ์ $(\mathrm{U} + \mathbf{zz}')^{-1}$ ทางด้านซ้ายของสมการทั้ง 2 ข้าง จะได้

$$\mathbf{w}_x^* = (\mathrm{U} + \mathbf{zz}')^{-1}\left(\bar{\mathrm{X}}\mathbf{z} + \frac{\boldsymbol{\nu}}{2}\right) - \frac{\lambda}{2}(\mathrm{U} + \mathbf{zz}')^{-1}\mathbf{1} \tag{3.19}$$

นำค่า $\mathbf{w}_x^*$ จาก (3.19) แทนใน (3.13) จะได้

$$\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\left(\bar{\mathrm{X}}\mathbf{z} + \frac{\boldsymbol{\nu}}{2}\right) - \frac{\lambda}{2}\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\mathbf{1} = 1$$

$$\frac{\lambda}{2} = \frac{\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\left(\bar{\mathrm{X}}\mathbf{z} + \frac{\boldsymbol{\nu}}{2}\right) - 1}{\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\mathbf{1}} \tag{3.20}$$

แทนค่า $\lambda/2$ จาก (3.20) ใน (3.19) จะได้

$$\mathbf{w}_x^* = (\mathrm{U} + \mathbf{zz}')^{-1}\left(\bar{\mathrm{X}}\mathbf{z} + \frac{\boldsymbol{\nu}}{2}\right) - \frac{\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\left(\bar{\mathrm{X}}\mathbf{z} + \frac{\boldsymbol{\nu}}{2}\right) - 1}{\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\mathbf{1}}(\mathrm{U} + \mathbf{zz}')^{-1}\mathbf{1} \tag{3.21}$$

ถ้า $\mathbf{w}_x^* > \mathbf{0}$, $\boldsymbol{\nu}$ มีค่าเท่ากับ $\mathbf{0}$ เพื่อสอดคล้องกับ Complementary slackness (3.14) ดังนั้น ถ้า $\mathbf{w}_\mathbf{x}^* > \mathbf{0}$ (3.21) จะลดรูปลงเหลือ

$$\mathbf{w}_x^* = (\mathrm{U} + \mathbf{zz}')^{-1}\bar{\mathrm{X}}\mathbf{z} - \frac{\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\bar{\mathrm{X}}\mathbf{z} - 1}{\mathbf{1}'(\mathrm{U} + \mathbf{zz}')^{-1}\mathbf{1}}(\mathrm{U} + \mathbf{zz}')^{-1}\mathbf{1} \tag{3.22}$$

แทนค่า $(\mathrm{U} + \mathbf{zz}')^{-1}$ จาก (3.15) ลงใน (3.22) จะได้

$$w_x^* = \left(\mathrm{U}^{-1} - \frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right)\bar{\mathrm{X}}\mathbf{z} - \frac{\left\{\mathbf{1}'\left(\mathrm{U}^{-1} - \frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right)\bar{\mathrm{X}}\mathbf{z} - 1\right\}}{\mathbf{1}'\left(\mathrm{U}^{-1} - \frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right)\mathbf{1}}\left(\mathrm{U}^{-1} - \frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right)\mathbf{1}$$

$$= \left\{\frac{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}) - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right\}\bar{\mathrm{X}}\mathbf{z} - \frac{\left[\mathbf{1}'\left\{\frac{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}) - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right\}\bar{\mathrm{X}}\mathbf{z} - 1\right]}{\frac{\mathbf{1}'\left\{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}) - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\right\}\mathbf{1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}}$$

$$\left(\mathrm{U}^{-1} - \frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right)\mathbf{1}$$

$$= \frac{\bar{\mathrm{X}}\mathrm{U}^{-1}\mathbf{z}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}} - \frac{\left[\mathbf{1}'\left\{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}) - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\right\}\bar{\mathrm{X}}\mathbf{z} - (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\right]}{\mathbf{1}'\left\{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}) - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\right\}\mathbf{1}}\left(\mathrm{U}^{-1}\frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right)\mathbf{1}$$

$$= \frac{\bar{\mathrm{X}}\mathrm{U}^{-1}\mathbf{z}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}} - \frac{\bar{\mathrm{X}}\mathbf{1}'\mathrm{U}^{-1}\mathbf{z} - (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})}{\mathbf{1}'\left\{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}) - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\right\}\mathbf{1}}\left(\mathrm{U}^{-1}\frac{\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}}{1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}}\right)\mathbf{1}$$

พิจารณาตัวเศษ

$$= \bar{\mathrm{X}}\mathrm{U}^{-1}\mathbf{z}\mathbf{1}'\left\{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z}) - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\right\}\mathbf{1} - \left\{\bar{\mathrm{X}}(\mathbf{1}'\mathrm{U}^{-1}\mathbf{z}) - (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\right\}$$
$$\left\{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\mathbf{1} - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\mathbf{1}\right\}$$

$$= (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\left\{\bar{\mathrm{X}}\mathrm{U}^{-1}\mathbf{z}\mathbf{1}'\mathrm{U}^{-1}\mathbf{1} + \mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\mathbf{1} - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\mathbf{1}\right\}$$
$$- \bar{\mathrm{X}}\mathrm{U}^{-1}\mathbf{z}\mathbf{1}'\mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\mathbf{1} - \bar{\mathrm{X}}\mathbf{1}'\mathrm{U}^{-1}\mathbf{z}\left\{\mathrm{U}^{-1}(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\mathbf{1} - \mathrm{U}^{-1}\mathbf{zz}'\mathrm{U}^{-1}\mathbf{1}\right\}$$

$$= (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\left\{\bar{\mathrm{X}}\mathbf{1}'\mathrm{U}^{-1}\mathbf{1}\mathrm{U}^{-1}\mathbf{z} + (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\mathrm{U}^{-1}\mathbf{1} - \mathbf{z}'\mathrm{U}^{-1}\mathbf{1}\mathrm{U}^{-1}\mathbf{z}\right\}$$
$$-(1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\bar{\mathrm{X}}\mathbf{1}'\mathrm{U}^{-1}\mathbf{z}\mathrm{U}^{-1}\mathbf{1}$$

$$= (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\left\{\bar{\mathrm{X}}\mathbf{1}'\mathrm{U}^{-1}\mathbf{1}\mathrm{U}^{-1}\mathbf{z} + (1 + \mathbf{z}'\mathrm{U}^{-1}\mathbf{z})\mathrm{U}^{-1}\mathbf{1} - \mathbf{z}'\mathrm{U}^{-1}\mathbf{1}\mathrm{U}^{-1}\mathbf{z} - \bar{\mathrm{X}}\mathbf{1}'\mathrm{U}^{-1}\mathbf{z}\mathrm{U}^{-1}\mathbf{1}\right\}$$

ดังนั้นค่าของน้ำหนักมีค่าเท่ากับ

$$\mathbf{w}_x^* = \frac{\bar{X}(\mathbf{1}'U^{-1}\mathbf{1})U^{-1}\mathbf{z} - \bar{X}(\mathbf{1}'U^{-1}\mathbf{z})U^{-1}\mathbf{1} + (1+\mathbf{z}'U^{-1}\mathbf{z})U^{-1}\mathbf{1} - (\mathbf{z}'U^{-1}\mathbf{1})U^{-1}\mathbf{z}}{\mathbf{1}'\left\{U^{-1}(1+\mathbf{z}'U^{-1}\mathbf{z}) - U^{-1}\mathbf{z}\mathbf{z}'U^{-1}\right\}\mathbf{1}}$$

$$= \frac{(1+\mathbf{z}'U^{-1}\mathbf{z})U^{-1}\mathbf{1} + \bar{X}(\mathbf{1}'U^{-1}\mathbf{1})U^{-1}\mathbf{z} - \bar{X}(\mathbf{1}'U^{-1}\mathbf{1})U^{-1}\mathbf{1} - (\mathbf{1}'U^{-1}\mathbf{z})U^{-1}\mathbf{z}}{(1+\mathbf{z}'U^{-1}\mathbf{z})\mathbf{1}'U^{-1}\mathbf{1} - \mathbf{1}'U^{-1}\mathbf{z}\mathbf{z}'U^{-1}\mathbf{1}}$$

$$= \frac{(1+\mathbf{z}'U^{-1}\mathbf{z})U^{-1}\mathbf{1} + \bar{X}(\mathbf{1}'U^{-1}\mathbf{1})U^{-1}\mathbf{z} - (\mathbf{1}'U^{-1}\mathbf{z})U^{-1}\mathbf{z} - \bar{X}(\mathbf{1}'U^{-1}\mathbf{z})U^{-1}\mathbf{1}}{(1+\mathbf{z}'U^{-1}\mathbf{z})\mathbf{1}'U^{-1}\mathbf{1} - \mathbf{1}'U^{-1}\mathbf{z}\mathbf{z}'U^{-1}\mathbf{1}}$$

$$= \frac{(1+\mathbf{z}'U^{-1}\mathbf{z})U^{-1}\mathbf{1} + \bar{X}(\mathbf{1}'U^{-1}\mathbf{1})U^{-1}\mathbf{z} - (\mathbf{1}'U^{-1}\mathbf{z})U^{-1}(\mathbf{z} + \bar{X}\mathbf{1})}{(1+\mathbf{z}'U^{-1}\mathbf{z})\mathbf{1}'U^{-1}\mathbf{1} - \mathbf{1}'U^{-1}\mathbf{z}\mathbf{z}'U^{-1}\mathbf{1}} \qquad (3.23)$$

เมื่อนำค่าของ U และ $\mathbf{z}$ กลับมาแทนใน (3.23) จะได้

$$\mathbf{w}_x^* = \frac{\left(1 + \sum_{h=1}^{k}\frac{\{E(\bar{x}_h)\}^2}{V(\bar{x}_h)}\right)\begin{bmatrix}\frac{1}{V(\bar{x}_1)}\\ \vdots \\ \frac{1}{V(\bar{x}_k)}\end{bmatrix} + \bar{X}\sum_{h=1}^{k}\frac{1}{V(\bar{x}_h)}\begin{bmatrix}\frac{E(\bar{x}_1)}{V(\bar{x}_1)}\\ \vdots \\ \frac{E(\bar{x}_k)}{V(\bar{x}_k)}\end{bmatrix} - \sum_{h=1}^{k}\frac{E(\bar{x}_h)}{V(\bar{x}_h)}\begin{bmatrix}\frac{E(\bar{x}_1)}{V(\bar{x}_1)} + \frac{\bar{X}}{V(\bar{x}_1)}\\ \vdots \\ \frac{E(\bar{x}_k)}{V(\bar{x}_k)} + \frac{\bar{X}}{V(\bar{x}_k)}\end{bmatrix}}{\sum_{h=1}^{k}\frac{1}{V(\bar{x}_h)}\left\{1 + \sum_{h=1}^{k}\frac{\{E(\bar{x}_h)\}^2}{V(\bar{x}_h)}\right\} - \left\{\sum_{h=1}^{k}\frac{E(\bar{x}_h)}{V(\bar{x}_h)}\right\}^2} \qquad (3.24)$$

ดังนั้น

$$w_{xj}^* = \frac{E(\bar{x}_j)}{V(\bar{x}_j)}\frac{\frac{1}{E(\bar{x}_j)}\left[1 + \sum_{h=1}^{k}\frac{\{E(\bar{x}_h)\}^2}{V(\bar{x}_h)}\right] + \bar{X}\sum_{h=1}^{k}\frac{1}{V(\bar{x}_h)} - \left\{1 + \frac{\bar{X}}{E(\bar{x}_j)}\right\}\sum_{h=1}^{k}\frac{E(\bar{x}_h)}{V(\bar{x}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{x}_h)}\left[1 + \sum_{h=1}^{k}\frac{\{E(\bar{x}_h)\}^2}{V(\bar{x}_h)}\right] - \left\{\sum_{h=1}^{k}\frac{E(\bar{x}_h)}{V(\bar{x}_h)}\right\}^2} \qquad (3.25)$$

พิจารณา

$$\sum_{j=1}^{k}\frac{1}{V(\bar{x}_j)}\left[1 + \sum_{j=1}^{k}\frac{\{E(\bar{x}_j)\}^2}{V(\bar{x}_j)}\right] > \sum_{j=1}^{k}\frac{1}{V(\bar{x}_j)}\sum_{j=1}^{k}\frac{\{E(\bar{x}_j)\}^2}{V(\bar{x}_j)} = \sum_{j=1}^{k}\left\{\frac{E(\bar{x}_j)}{V(\bar{x}_j)}\right\}^2 + \sum_{i=1}^{k}\sum_{\substack{j=1\\ j\neq i}}^{k}\frac{\{E(\bar{x}_i)\}^2}{V(\bar{x}_i)V(\bar{x}_j)}$$

$$> \sum_{j=1}^{k}\left\{\frac{E(\bar{x}_j)}{V(\bar{x}_j)}\right\}^2 + \sum_{i=1}^{k}\sum_{\substack{j=1\\ j\neq i}}^{k}\frac{E(\bar{x}_i)E(\bar{x}_j)}{V(\bar{x}_i)V(\bar{x}_j)} = \left\{\sum_{j=1}^{k}\frac{E(\bar{x}_j)}{V(\bar{x}_j)}\right\}^2 ,$$

เพราะว่า $\dfrac{\{E(\bar{x}_i)\}^2}{V(\bar{x}_i)V(\bar{x}_j)} + \dfrac{\{E(\bar{x}_i)\}^2}{V(\bar{x}_i)V(\bar{x}_j)} > 2\dfrac{E(\bar{x}_i)E(\bar{x}_j)}{V(\bar{x}_i)V(\bar{x}_j)}$, ถ้า $E(\bar{x}_i) \neq E(\bar{x}_j)$ สำหรับ $i \neq j$

ซึ่งเป็นกรณีปกติของการชักตัวอย่างแบบชั้นภูมิ ในกรณีที่ไม่ทราบค่าพารามิเตอร์ของตัวแปรเสริมในชั้นภูมิ การคำนวณค่าน้ำหนักใน (3.25) จะแทนค่าพารามิเตอร์ทุกตัวด้วยค่าสถิติที่มาจากตัวอย่าง ซึ่งจะได้ค่าน้ำหนักดังนี้

$$\hat{w}^*_{xj} = \frac{\bar{x}_j}{\gamma_j s^2_{xj}} \frac{\dfrac{1}{\bar{x}_j}\left(1 + \sum\limits_{h=1}^{k}\dfrac{\bar{x}_h^2}{\gamma_h s^2_{xh}}\right) + \bar{X}\sum\limits_{h=1}^{k}\dfrac{1}{\gamma_h s^2_{xh}} - \left(1 + \dfrac{\bar{X}}{\bar{x}_j}\right)\sum\limits_{h=1}^{k}\dfrac{\bar{x}_h}{\gamma_h s^2_{xh}}}{\sum\limits_{h=1}^{k}\dfrac{1}{\gamma_h s^2_{xh}}\left(1 + \sum\limits_{h=1}^{k}\dfrac{\bar{x}_h^2}{\gamma_h s^2_{xh}}\right) - \left(\sum\limits_{h=1}^{k}\dfrac{\bar{x}_h}{\gamma_h s^2_{xh}}\right)^2} \quad ; \; j = 1, 2, \ldots, k$$

(3.26)

ในทำนองเดียวกัน เมื่อ $w^*_y > 0$, w^*_{yj} สามารถเขียนได้

$$w^*_{yj} = \frac{E(\bar{y}_j)}{V(\bar{y}_j)} \frac{\dfrac{1}{E(\bar{y}_j)}\left[1 + \sum\limits_{h=1}^{k}\dfrac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right] + \bar{Y}\sum\limits_{h=1}^{k}\dfrac{1}{V(\bar{y}_h)} - \left\{1 + \dfrac{\bar{Y}}{E(\bar{y}_j)}\right\}\sum\limits_{h=1}^{k}\dfrac{E(\bar{y}_h)}{V(\bar{y}_h)}}{\sum\limits_{h=1}^{k}\dfrac{1}{V(\bar{y}_h)}\left\{1 + \sum\limits_{h=1}^{k}\dfrac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\} - \left\{\sum\limits_{h=1}^{k}\dfrac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2} \qquad (3.27)$$

แม้จะแทนค่าสถิติชั้นภูมิด้วยค่าสถิติตัวอย่าง การคำนวณค่าน้ำหนัก w^*_{yj} ใน (3.27) ยังต้องการค่าประมาณค่าเฉลี่ยของประชากร $\bar{Y}$ ซึ่งอาจคำนวณได้โดยใช้น้ำหนัก w^*_{xh} ใน (3.25) ดังนี้

$$\hat{\bar{Y}}_{wx} = \sum_{h=1}^{k}\hat{w}^*_{xh}\bar{y}_h \qquad (3.28)$$

ค่าประมาณ $\hat{w}^*_{yj}$ จึงเขียนได้เป็น

$$\hat{w}^*_{yj} = \frac{\bar{y}_j}{\gamma_j s^2_{yj}} \frac{\frac{1}{\bar{y}_j}\left(1+\sum_{h=1}^{k}\frac{\bar{y}_h^2}{\gamma_h s^2_{yh}}\right)+\bar{y}_{wx}\sum_{h=1}^{k}\frac{1}{\gamma_h s^2_{yh}}-\left(1+\frac{\bar{y}_{wx}}{\bar{y}_j}\right)\sum_{h=1}^{k}\frac{\bar{y}_h}{\gamma_h s^2_{yh}}}{\sum_{h=1}^{k}\frac{1}{\gamma_h s^2_{yh}}\left(1+\sum_{h=1}^{k}\frac{\bar{y}_h^2}{\gamma_h s^2_{yh}}\right)-\left(\sum_{h=1}^{k}\frac{\bar{y}_h}{\gamma_h s^2_{yh}}\right)^2} \tag{3.29}$$

ทฤษฎีบทที่ 3.3 ในการชักตัวอย่างสุ่มแบบชั้นภูมิที่มีประชากรแบ่งออกเป็น k ชั้นภูมิ ตัวประมาณค่า $\hat{\bar{Y}}_{st\_wy} = \sum_{h=1}^{k} w^*_{yh}\bar{y}_h$ เป็นตัวประมาณค่าที่เอนเอียง แต่คงเส้นคงวา (consistent) ของ $\bar{Y}$

พิสูจน์

$$\sum_{h=1}^{k} w^*_{yh}E(\bar{y}_h) = \frac{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\left[1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right]+\bar{Y}\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

$$-\frac{\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\left\{1+\frac{\bar{Y}}{E(\bar{y}_h)}\right\}\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

$$= \frac{\bar{Y}\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}+\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}+\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

$$-\frac{\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}+\bar{Y}\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

$$\sum_{j=1}^{k} w^{*}_{yj}E(\bar{y}_j) = \frac{\bar{Y}\left[\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{j=1}^{k}\frac{E(\bar{y}_j)}{V(\bar{y}_j)}\right\}^2\right]-\bar{Y}\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}+\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

$$E(\hat{\hat{Y}}_{st\_wy}) = \bar{Y}+\frac{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}-\bar{Y}\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

ดังนั้น $$Bias(\hat{\hat{Y}}_{st\_wy}) = \frac{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}-\bar{Y}\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

$$= \frac{\sum_{h=1}^{k}\frac{n_hE(\bar{y}_h)}{V(\bar{y}_h)}-\bar{Y}\sum_{h=1}^{k}\frac{n_h}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{n_h}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{n_h\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\}-\left\{\sum_{h=1}^{k}\frac{n_hE(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2}$$

$$< \frac{\sum_{h=1}^{k}\frac{n_hE(\bar{y}_h)}{V(y_h)}-\bar{Y}\sum_{h=1}^{k}\frac{n_h}{V(y_h)}}{\sum_{h=1}^{k}\frac{n_h}{V(y_h)}\sum_{h=1}^{k}\frac{n_h\{E(\bar{y}_h)\}^2}{V(y_h)}-\left\{\sum_{h=1}^{k}\frac{n_hE(\bar{y}_h)}{V(y_h)}\right\}^2}$$

$$\left|Bias(\hat{\hat{Y}}_{st\_wy})\right| \le \left|\frac{n_{hmin}\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(y_h)}-\bar{Y}n_{hmax}\sum_{h=1}^{k}\frac{1}{V(y_h)}}{n^2_{hmin}\sum_{h=1}^{k}\frac{1}{V(y_h)}\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(y_h)}-n^2_{hmax}\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(y_h)}\right\}^2}\right|$$

$$\le \frac{1}{n_{hmax}}\frac{\frac{n_{hmin}}{n_{hmax}}\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(y_h)}-\bar{Y}\sum_{h=1}^{k}\frac{1}{V(y_h)}}{\left(\frac{n_{hmin}}{n_{hmax}}\right)^2\sum_{h=1}^{k}\frac{1}{V(y_h)}\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(y_h)}-\left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(y_h)}\right\}^2}.$$

$$\lim_{\substack{n_{h\min}\to\infty \\ n_{h\max}\to\infty}} \left|\mathrm{Bias}(\hat{\hat{\bar{Y}}}_{st\_wy})\right| \le \lim_{\substack{n_{h\min}\to\infty \\ n_{h\max}\to\infty}} \left| \frac{1}{n_{hmax}} \frac{\frac{n_{hmin}}{n_{hmax}} \sum_{h=1}^{k} \frac{E(\bar{y}_h)}{V(y_h)} - \bar{Y} \sum_{h=1}^{k} \frac{1}{V(y_h)}}{\left(\frac{n_{hmin}}{n_{hmax}}\right)^2 \sum_{h=1}^{k} \frac{1}{V(y_h)} \sum_{h=1}^{k} \frac{\{E(\bar{y}_h)\}^2}{V(y_h)} - \left\{\sum_{h=1}^{k} \frac{E(\bar{y}_h)}{V(y_h)}\right\}^2} \right|$$

ในทำนองเดียวกัน ถ้า $E(\bar{y}_i) \neq E(\bar{y}_j)$ สำหรับ $i \neq j$ สามารถแสดงได้ว่า

$$\sum_{h=1}^{k} \frac{1}{V(y_h)} \sum_{h=1}^{k} \frac{\{E(\bar{y}_h)\}^2}{V(y_h)} > \left\{\sum_{h=1}^{k} \frac{E(\bar{y}_h)}{V(y_h)}\right\}^2$$

จึงกล่าวได้ว่า ค่าสัมบูรณ์ของค่าความเอนเอียงของตัวประมาณค่า $\hat{\hat{\bar{Y}}}_{st\_wy}$ มีค่าเข้าใกล้ 0 เมื่อ $n_{hmax} \to \infty$ และ $n_{hmin} \to \infty$ ดังนั้นตัวประมาณค่ามีคุณสมบัติคงเส้นคงวา

ช.ต.พ.

3.3 ตัวประมาณค่าที่พัฒนาขึ้นใหม่

การพัฒนาตัวประมาณค่ามี 2 แนวทาง คือ การไม่ใช้ตัวแปรเสริมในการประมาณค่า และการใช้ตัวแปรเสริมในการประมาณค่า

3.3.1 การไม่ใช้ตัวแปรเสริมในการประมาณค่า เนื่องจากความสัมพันธ์ของตัวแปรเสริมและตัวแปรที่สนใจ ไม่สอดคล้องกับอสมการ 3.2 ในทฤษฎีบทที่ 3.1 เป็นเกณฑ์ในการตัดสินใจ

1) การประมาณค่าเฉลี่ยของประชากร

ในกรณีนี้ตัวประมาณค่าที่พัฒนาขึ้นใหม่ คือ

$$\hat{\hat{\bar{Y}}}_{st\_wy} = \bar{y}_{st\_wy} = \sum_{h=1}^{k} w_{yh}\bar{y}_h \tag{3.30}$$

เป็นตัวประมาณค่าที่ใช้ค่าน้ำหนักที่คำนวณมาจากค่าสถิติของตัวแปรที่สนใจ y ดังนั้นความแปรปรวนมีค่าเท่ากับ

$$V(\bar{y}_{st\_wy}) = \sum_{h=1}^{k} w_{yh}^2 \gamma_h S_{yh}^2 \tag{3.31}$$

2) การประมาณค่ายอดรวม

ตัวประมาณค่าคือ

$$\hat{Y}_{st\_wy} = N\bar{y}_{st\_wy} \tag{3.32}$$

ความแปรปรวนมีค่าเท่ากับ

$$V(\hat{Y}_{st\_wy}) = N^2 V(\bar{y}_{st\_wy}) \tag{3.33}$$

เมื่อ N คือขนาดของประชากร

3) การประมาณค่าสัดส่วนของหน่วยที่มีคุณสมบัติที่สนใจ
ตัวประมาณค่าสัดส่วนได้แก่

$$\hat{P}_{st\_wy} = \sum_{h=1}^{k} w_{yh} p_h \tag{3.34}$$

และความแปรปรวนของตัวประมาณค่าสัดส่วนมีค่าเท่ากับ

$$V(\hat{P}_{st\_wy}) = \sum_{h=1}^{k} w_{yh}^2 \gamma_h S_h^2 \tag{3.35}$$

เมื่อ p_h คือค่าของสัดส่วนของหน่วยที่มีคุณสมบัติที่สนใจในตัวอย่างในชั้นภูมิที่ h
และ S_h^2 เป็นค่าความแปรปรวนของสัดส่วนในชั้นภูมิที่ h

3.3.2 การใช้ตัวแปรเสริมในการประมาณค่า ในกรณีที่ตัวแปรเสริมและตัวแปรที่สนใจมีความสัมพันธ์กันสูง การพิจารณาที่จะใช้ตัวประมาณค่าอัตราส่วนรวมมีเกณฑ์ดังนี้ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรเสริมและตัวแปรที่สนใจสอดคล้องตามเงื่อนไขในอสมการที่ 3.2 ในทฤษฎีบทที่ 3.1 ค่าเฉลี่ยของประชากรจะประมาณโดยใช้ตัวประมาณค่าอัตราส่วนรวมที่ใช้ตัวแปรเสริม

ในทางปฏิบัติการหาค่าน้ำหนักที่ดีที่สุดเพื่อให้ค่าความคลาดเคลื่อนของตัวประมาณค่ามีค่าต่ำสุดนั้น สามารถประมาณค่าของน้ำหนัก w_{yh}, w_{xh} ได้จากการใช้ค่าสถิติของชั้นภูมิโดยแทนค่าความคาดหมายและค่าความแปรปรวนของ $\bar{x}_h$ และของ $\bar{y}_h$ ในชั้นภูมิ การประมาณค่าน้ำหนัก, w_{xh} และ w_{yh} ไม่ได้ขึ้นอยู่กับค่าของ N_h โดยตรงเหมือนการประมาณค่าจากการชักตัวอย่างแบบชั้นภูมิที่ใช้กันอยู่โดยทั่วไปซึ่งใช้ค่าถ่วงน้ำหนักเท่ากับ $\frac{N_h}{N}$ แต่ค่าน้ำหนักที่ใช้ถ่วงในกรณีนี้ใช้ค่า N_h ทางอ้อมโดยใช้ผ่านปัจจัยปรับประชากรอันตะ γ_h ตัวประมาณค่าที่นำเสนอมีการใช้น้ำหนักที่แตกต่างกัน 3 ตัวด้วยกันคือ

1) ใช้น้ำหนักของตัวแปรช่วยx ในการประมาณค่าเฉลี่ยของตัวแปรที่สนใจและตัวแปรเสริม

$$\hat{\bar{Y}}_{Rwx} = \frac{\bar{y}_{st\_wx}}{\bar{x}_{st\_wx}} \cdot \bar{X}$$

$$= \frac{\sum_{h=1}^{k} \hat{w}^{*}_{xh}\bar{y}_h}{\sum_{h=1}^{k} \hat{w}^{*}_{xh}\bar{x}_h} \cdot \bar{X} \qquad (3.36)$$

การประมาณค่าเฉลี่ยในกรณีนี้จะใช้ตัวแปรเสริมเข้ามาช่วยในการให้ข้อสนเทศเกี่ยวกับ y และใช้วิธีการประมาณค่าที่ใช้อัตราส่วนรวมมาช่วยในการประมาณค่าเฉลี่ย $(\bar{Y})$ ของประชากร

$$\text{Bias}(\hat{\bar{Y}}_{Rwx}) = \frac{1}{\bar{X}} \sum_{h=1}^{k} \hat{w}^{*}_{xh}\gamma_h(R_{wx}S^2_{xh} - S_{yxh}) \qquad (3.37)$$

และ

$$\text{MSE}(\hat{\bar{Y}}_{Rwx}) = \sum_{h=1}^{k} \hat{w}^{*}_{xh}\gamma_h(S^2_{yh} - 2R_{wx}S_{yxh} + R^2_{wx}S^2_{xh}) \qquad (3.38)$$

เมื่อ

$$R_{wx} = \frac{\sum_{h=1}^{k} \hat{w}^{*}_{xh}\bar{y}_h}{\sum_{h=1}^{k} w^{2}_{xh}\bar{x}_h}$$

2) ใช้น้ำหนักของตัวแปรที่สนใจคือ y ในการประมาณค่าเฉลี่ยของตัวแปรที่สนใจ $\bar{y}_{st}$ และตัวแปรช่วย $\bar{x}_{st}$ ตามลำดับ

$$\hat{\bar{Y}}_{Rwy} = \frac{\bar{y}_{st\_wy}}{\bar{x}_{st\_wy}} \cdot \bar{X} \qquad (3.39)$$

ค่าความเอนเอียงและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองมีค่าเช่นเดียวกับสมการ (3.34) และ (3.35) แต่แทนค่าของน้ำหนักถ่วงเปลี่ยน $\hat{w}^{*}_{xh}$ เป็น $\hat{w}^{*}_{yh}$ และ R_{wx} เป็น R_{wy}

เมื่อ

$$R_{wy} = \frac{\sum_{h=1}^{k} \hat{w}^{*}_{yh}\bar{y}_h}{\sum_{h=1}^{k} \hat{w}^{*}_{yh}\bar{x}_h},$$

3) ใช้น้ำหนักของตัวแปรช่วยx ประมาณค่าเฉลี่ยของ $\bar{x}_{st}$ และใช้น้ำหนักของตัวแปรที่สนใจ y ประมาณค่าเฉลี่ยของ $\bar{y}_{st}$ ดังนั้น ตัวประมาณค่าเฉลี่ยของประชากรได้แก่

$$\hat{\bar{Y}}_{Rwxy} = \frac{\bar{y}_{st\_wy}}{\bar{x}_{st\_wx}} \cdot \bar{X} \tag{3.40}$$

ค่าความเอนเอียงและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\hat{\bar{Y}}_{Rwxy}$ มีค่าเท่ากับ

$$\text{Bias}(\hat{\bar{Y}}_{Rwyx}) = \frac{1}{\bar{X}}\left(R_{wy} \sum_{h=1}^{k} \hat{w}_{xh}^{*2} \gamma_h S_{xh}^2 - \sum_{h=1}^{k} \hat{w}_{xh}^{*} \hat{w}_{yh}^{*} \gamma_h S_{yxh} \right) \tag{3.41}$$

และ

$$\text{MSE}(\hat{\bar{Y}}_{Rwyx}) = \sum_{h=1}^{k} \hat{w}_{xh}^{*2} \gamma_h S_{yh}^2 - 2R_{wyx} \sum_{h=1}^{k} \hat{w}_{xh}^{*} \hat{w}_{yh}^{*} \gamma_h S_{yxh} + R_{wyx}^2 \sum_{h=1}^{k} \hat{w}_{xh}^{*2} \gamma_h S_{xh}^2 \tag{3.42}$$

เมื่อ $$R_{wyx} = \frac{\sum_{h=1}^{k} \hat{w}_{yh}^{*} \bar{y}_h}{\sum_{h=1}^{k} \hat{w}_{yh}^{*} \bar{x}_h}$$

บทที่ 4
การเปรียบเทียบตัวประมาณค่า

ในบทนี้จะทำการเปรียบเทียบตัวประมาณค่าที่พัฒนาขึ้นทั้ง 4 ตัว กับตัวประมาณค่าตัวประมาณของค่าเฉลี่ย $\hat{\bar{Y}}_{st}$ ตัวประมาณของค่าเฉลี่ย $\hat{\bar{Y}}_{CR}$ ที่ใช้ตัวแปรช่วยเข้ามาช่วยในกรณีที่ตัวแปรที่สนใจ y และตัวแปรช่วยx มีความสัมพันธ์กัน ตัวประมาณค่าเฉลี่ยของคาดีลาร์และซินจิและตัวประมาณค่าของซิสโซเดียและดวายดิ

4.1 ตัวประมาณค่าที่พัฒนา

ตัวประมาณค่าเฉลี่ยของประชากรที่ได้นำเสนอใหม่ 4 ตัว พัฒนาขึ้นจากค่าน้ำหนักในการถ่วงโดยใช้ค่าสถิติของชั้นภูมิ เพื่อให้ได้ค่าประมาณของค่าเฉลี่ยของประชากรมีค่าเฉลี่ยของค่าความคลาดเคลื่อนต่ำสุด และการนำตัวประมาณค่าที่พัฒนาขึ้นไปใช้ในกรณีที่ไม่ได้ใช้ตัวแปรเสริมเข้ามาช่วยนั้น เป็นกรณีที่เงื่อนไขไม่สอดคล้องกับอสมการ (3.2) ของทฤษฎีบทที่ 1 แต่ถ้าเงื่อนไขสอดคล้องกับอสมการ (3.2) ตัวประมาณค่าอาจใช้ตัวประมาณค่า (3.33), (3.36) และ (3.37) ตัวใดตัวหนึ่ง โดยตัวประมาณค่าใหม่ 4 ตัว สรุปได้เป็น

$$\hat{\bar{Y}}_{st\_wy} = \sum_{h=1}^{k} \hat{w}^*_{yh}\bar{y}_h \tag{4.1}$$

$$\hat{\bar{Y}}_{Rwx} = \frac{\bar{y}_{st\_wx}}{\bar{x}_{st\_wx}}\cdot\bar{X} = \frac{\sum_{h=1}^{k}\hat{w}^*_{xh}\bar{y}_h}{\sum_{h=1}^{k}\hat{w}^*_{xh}\bar{x}_h}\bar{X} = R_{wx}\bar{X} \tag{4.2}$$

$$\hat{\bar{Y}}_{Rwy} = \frac{\bar{y}_{st\_wy}}{\bar{x}_{st\_wy}}\bar{X} = \frac{\sum_{h=1}^{k}\hat{w}^*_{yh}\bar{y}_h}{\sum_{h=1}^{k}\hat{w}^*_{yh}\bar{x}_h}\bar{X} = R_{wy}\bar{X} \tag{4.3}$$

$$\hat{\bar{Y}}_{Rwyx} = \frac{\bar{y}_{st\_wy}}{\bar{x}_{st\_wx}}\bar{X} = \frac{\sum_{h=1}^{k}\hat{w}^*_{yh}\bar{y}_h}{\sum_{h=1}^{k}\hat{w}^*_{xh}\bar{x}_h}\cdot\bar{X} = R_{wyx}\bar{X} \tag{4.4}$$

ค่าที่ใช้ในการเปรียบเทียบได้แก่ ค่าความเอนเอียง (Bias) และประสิทธิภาพของตัวประมาณค่า

ตัวประมาณค่า $\hat{\bar{Y}}_{st\_wy}$ ใน (4.1) เป็นตัวประมาณค่าที่เอนเอียง แต่คงเส้นคงวาตามทฤษฎีบทที่ 3 ดังนั้น ค่าความเอนเอียงมีค่าเท่ากับ

$$\text{Bias}(\hat{\hat{Y}}_{st\_wy}) = \frac{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)} - \bar{Y}\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}}{\sum_{h=1}^{k}\frac{1}{V(\bar{y}_h)}\left\{1+\sum_{h=1}^{k}\frac{\{E(\bar{y}_h)\}^2}{V(\bar{y}_h)}\right\} - \left\{\sum_{h=1}^{k}\frac{E(\bar{y}_h)}{V(\bar{y}_h)}\right\}^2} \quad (4.5)$$

และความแปรปรวนของ $\hat{\hat{Y}}_{st\_wy}$ มีค่าเท่ากับ

$$V\left(\hat{\hat{Y}}_{st\_wy}\right) = \sum_{h=1}^{k}\hat{w}_{yh}^{*2}\gamma_h S_{yh}^2 \quad (4.6)$$

ซึ่งจะเห็นได้ว่า ในการประมาณค่าโดยใช้องศาที่ 1 ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสองมีค่าเท่ากับความแปรปรวน (4.6) สำหรับตัวประมาณค่าใน (4.2) (4.3) และ (4.4) ก็เป็นตัวประมาณค่าที่มีความเอนเอียง ดังนั้น ค่าความเอนเอียงสำหรับตัวประมาณค่าใน (4.2) ซึ่งเขียนได้ในรูปเดียวกันกับ (2.9) เท่ากับ

$$\text{Bias}\left(\hat{\hat{Y}}_{Rwx}\right) = \frac{1}{\bar{\bar{X}}}\left\{\sum_{h=1}^{k}\hat{w}_{xh}^{*2}\gamma_h\left(R_{wx}S_{xh}^2 - S_{yxh}\right)\right\} \quad (4.7)$$

ในการประมาณค่าโดยใช้องศาที่ 1 ค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง ซึ่งเขียนได้ในรูปเดียวกันกับ (2.8) มีค่าเท่ากับค่าความแปรปรวน

$$\text{MSE}\left(\hat{\hat{Y}}_{Rwx}\right) = \sum_{h=1}^{k}\hat{w}_{xh}^{*2}\gamma_h\left(S_{yh}^2 - 2R_{wx}S_{yxh} + R_{wx}^2S_{xh}^2\right) \quad (4.8)$$

ตัวประมาณค่าใน (4.3) จะมีค่าความเอนเอียงและค่าเฉลี่ยความคลาดเคลื่อนเท่ากับตัวประมาณค่าใน (4.2) แต่แทนค่าน้ำหนักที่ใช้ถ่วงด้วย w_{yh} เท่านั้น สำหรับตัวประมาณค่าใน (4.4) มีค่าความเอนเอียง (3.38) และค่าเฉลี่ยความคลาดเคลื่อน (3.39) เท่ากับ

$$\text{Bias}\left(\hat{\hat{Y}}_{Rwyx}\right) = \frac{1}{\bar{\bar{X}}}\left(R_{wy}\sum_{h=1}^{k}\hat{w}_{xh}^{*2}\gamma_h S_{xh}^2 - \sum_{h=1}^{k}\hat{w}_{xh}^{*}\hat{w}_{yh}^{*}\gamma_h S_{yxh}\right) \quad (4.9)$$

และ

$$\text{MSE}\left(\hat{\hat{Y}}_{Rwyx}\right) = \sum_{h=1}^{k}\hat{w}_{yh}^{*2}\gamma_h S_{yh}^2 - 2R_{wyx}\sum_{h=1}^{k}\hat{w}_{yh}\hat{w}_{xh}\gamma_h S_{yxh} + R_{wyx}^2\sum_{h=1}^{k}\hat{w}_{xh}^{*2}\gamma_h S_{xh}^2 \quad (4.10)$$

ตัวประมาณค่าที่นำเสนอจะมีค่า $\text{MSE}\left(\hat{\hat{Y}}_{Rwx}\right)$ น้อยกว่า $\text{MSE}\left(\hat{\hat{Y}}_{Rwy}\right)$ และ $\text{MSE}\left(\hat{\hat{Y}}_{Rwyx}\right)$ ถ้า

$$S^2_{\bar{y}_{st\_wx}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2+1-\rho^2_{cwx}\right\}<S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wy}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwy}\right)^2+1-\rho^2_{cwy}\right\}$$

และ

$$S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wy}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwy}\right)^2+1-\rho^2_{cwy}\right\}<S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwyx}\right)^2+1-\rho^2_{cwyx}\right\} \quad (4.11)$$

$C_{\bar{y}_{st\_wy}}, C_{\bar{x}_{st\_wy}}, C_{\bar{y}_{st\_wx}}$ และ $C_{\bar{x}_{st\_wx}}$ เป็นสัมประสิทธิ์ของการกระจายของ $\bar{y}_{st\_wy}$ $\bar{x}_{st\_wy}$ $\bar{y}_{st\_wx}$ และ $\bar{x}_{st\_wx}$ ค่า ρ_{cwy}, ρ_{cwx} และ ρ_{cwyx} เป็นค่าสัมประสิทธิ์สหสัมพันธ์ส่วนรวมของค่าเฉลี่ยของตัวแปรที่สนใจ y และตัวแปรช่วยx ในแต่ละชั้นภูมิที่ใช้ค่าน้ำหนัก w_{yh}, w_{xh} และน้ำหนักของทั้ง w_{yh} และ w_{xh} ตามลำดับ

ในทำนองเดียวกันสามารถพิสูจน์ได้ว่า $MSE\left(\hat{\bar{Y}}_{Rwy}\right)$ มีค่าน้อยกว่า $MSE\left(\hat{\bar{Y}}_{Rwx}\right)$ และ $MSE\left(\hat{\bar{Y}}_{Rwyx}\right)$ ถ้าเป็นไปตามเงื่อนไข

$$S^2_{\bar{y}_{st\_wx}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2+1-\rho^2_{cwx}\right\}<S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wy}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwy}\right)^2+1-\rho^2_{cwy}\right\}$$

และ

$$S^2_{\bar{y}_{st\_wx}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2+1-\rho^2_{cwx}\right\}<S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwyx}\right)^2+1-\rho^2_{cwyx}\right\}$$

(4.12)

และจะได้ว่า $MSE\left(\hat{\bar{Y}}_{Rwyx}\right)$ น้อยกว่า $MSE\left(\hat{\bar{Y}}_{Rwx}\right)$ และ $MSE\left(\hat{\bar{Y}}_{Rwy}\right)$ ถ้า

$$S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wy}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwy}\right)^2+1-\rho^2_{cwy}\right\}<S^2_{\bar{y}_{st\_wx}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2+1-\rho^2_{cwx}\right\}$$

(4.13)

และ

$$S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wy}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwy}\right)^2+1-\rho^2_{cwy}\right\}<S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwyx}\right)^2+1-\rho^2_{cwyx}\right\}$$

(4.14)

ทฤษฎีบทที่ 4.1 ตัวประมาณค่า $\hat{\bar{Y}}_{Rwx}$ มีค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง $MSE\left(\hat{\bar{Y}}_{Rwx}\right)$ น้อยกว่า $MSE\left(\hat{\bar{Y}}_{Rwy}\right)$ และ $MSE\left(\hat{\bar{Y}}_{Rwyx}\right)$ ถ้าเงื่อนไขสอดคล้องกับอสมการ (4.11)

พิสูจน์

หาก (3.35) $MSE\left(\hat{\bar{Y}}_{Rwx}\right)$ สามารถเขียนใหม่ได้เป็น

$$MSE\left(\hat{\bar{Y}}_{Rwx}\right)=S^2_{\bar{y}_{st\_wx}}+R^2_{wx}S^2_{\bar{x}_{st\_wx}}-2R_{wx}S_{\bar{y}_{st\_wx}\bar{x}_{st\_wx}}$$

$$=S^2_{\bar{y}_{st\_wx}}\left\{1+R^2_{wx}\frac{S^2_{\bar{x}_{st\_wx}}}{S^2_{\bar{y}_{st\_wx}}}-2R_{wx}\rho_{cwx}\frac{S_{\bar{y}_{st\_wx}}S_{\bar{x}_{st\_wx}}}{S^2_{\bar{y}_{st\_wx}}}\right\}$$

$$=S^2_{\bar{y}_{st\_wx}}\left\{1+\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}\right)^2-2\rho_{cwx}\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}\right\}$$

$$=S^2_{\bar{y}_{st\_wx}}\left\{1+\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2-\rho^2_{cwx}\right\}$$

$$=S^2_{\bar{y}_{st\_wx}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2+1-\rho^2_{cwx}\right\}$$

ในทำนองเดียวกัน $MSE\left(\hat{\bar{Y}}_{Rwy}\right)$ สามารถเขียนได้เป็น

$$MSE\left(\hat{\bar{Y}}_{Rwy}\right)=S^2_{\bar{y}_{st\_wy}}+R^2_{wy}S^2_{\bar{x}_{st\_wy}}-2R_{wy}S_{\bar{y}_{st\_wy}\bar{x}_{st\_wy}}$$

$$=S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wy}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwy}\right)^2+1-\rho^2_{cwy}\right\}$$

ถ้าให้ $MSE\left(\bar{Y}_{Rwx}\right) < MSE\left(\bar{Y}_{Rwy}\right)$ จะได้ว่า

$$S^2_{\bar{y}_{st\_wx}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2+\left(1-\rho^2_{cwy}\right)\right\}$$

$$< S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwy}\right)^2+1-\rho^2_{cwy}\right\}$$

ในทำนองเดียวกันอาจพิสูจน์ได้ว่า $MSE\left(\hat{\bar{Y}}_{Rwx}\right)$ น้อยกว่า $MSE\left(\hat{\bar{Y}}_{Rwyx}\right)$ ถ้า

$$MSE\left(\hat{\bar{Y}}_{Rwyx}\right) = S^2_{\bar{y}_{st\_wy}} + R^2_{wyx}S^2_{\bar{x}_{st\_wx}} - 2R_{wyx}S_{\bar{y}_{st\_wy},\bar{x}_{st\_wx}}$$

$$= S^2_{\bar{y}_{st\_wy}}\left\{1+R^2_{wyx}\frac{S^2_{\bar{x}_{st\_wx}}}{S^2_{\bar{y}_{st\_wy}}}-2R_{wyx}\rho_{cwyx}\frac{S_{\bar{y}_{st\_wy}}S_{\bar{x}_{st\_wx}}}{S^2_{\bar{y}_{st\_wy}}}\right\}$$

$$= S^2_{\bar{y}_{st\_wy}}\left\{1+\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}\right)^2-2R_{wyx}\rho_{cwyx}\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}\right\}$$

$$= S^2_{\bar{y}_{st\_wy}}\left\{1+\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwyx}\right)^2-\rho^2_{cwyx}\right\}$$

$$= S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwyx}\right)^2+1-\rho^2_{cwyx}\right\}$$

ดังนั้น

$$S^2_{\bar{y}_{st\_wx}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wx}}}-\rho_{cwx}\right)^2+1-\rho^2_{cwx}\right\}$$

$$< S^2_{\bar{y}_{st\_wy}}\left\{\left(\frac{C_{\bar{x}_{st\_wx}}}{C_{\bar{y}_{st\_wy}}}-\rho_{cwyx}\right)^2+1-\rho^2_{cwyx}\right\}$$

ช.ต.พ.

ทฤษฎีบทที่ 4.2 ตัวประมาณค่า $\hat{\bar{Y}}_{wy}$ มีค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง $MSE\left(\hat{\bar{Y}}_{wy}\right)$ น้อยกว่า $MSE\left(\hat{\bar{Y}}_{wx}\right)$ และ $MSE\left(\hat{\bar{Y}}_{wyx}\right)$ ถ้าเงื่อนไขสอดคล้องกับอสมการ (4.13)

ทฤษฎีบทที่ 4.3 ตัวประมาณค่า $\hat{\bar{Y}}_{wyx}$ มีค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง $MSE\left(\hat{\bar{Y}}_{wyx}\right)$ น้อยกว่า $MSE\left(\hat{\bar{Y}}_{wx}\right)$ และ $MSE\left(\hat{\bar{Y}}_{wy}\right)$ ถ้าเงื่อนไขสอดคล้องอสมการ (4.14)

การพิสูจน์ทฤษฎีบทที่ 4.2 และ 4.3 สามารถกระทำได้ในทำนองเดียวกันกับการพิสูจน์ทฤษฎีบทที่ 4.1

4.2 การเปรียบเทียบตัวประมาณค่า

ตัวประมาณค่าที่ใช้ในการเปรียบเทียบมี 4 ตัวคือ

4.2.1 ตัวประมาณของค่าเฉลี่ย $\bar{Y}_{st}$ ในการชักตัวอย่างแบบชั้นภูมิ (Cochran, 1977) ได้แก่

$$\hat{\bar{Y}}_{st} = \sum_{h=1}^{k} w_h \bar{y}_h \tag{4.15}$$

เมื่อ $w_h = \dfrac{N_h}{N}$, N_h คือ ขนาดของประชากรในชั้นภูมิที่ h, N คือ ขนาดของประชากรทั้งหมด และ $\bar{y}_h$ คือ ค่าเฉลี่ยของตัวอย่างในชั้นภูมิที่ h

ตัวประมาณค่า $\hat{\bar{Y}}_{st}$ เป็นตัวประมาณค่าไม่เอนเอียงมีค่าความแปรปรวน

$$V\left(\hat{\bar{Y}}_{st}\right) = \sum_{h=1}^{k} w_h^2 \gamma_h S_{yh}^2 \tag{4.16}$$

4.2.2 ตัวประมาณของค่าเฉลี่ย $\bar{Y}_{st}$ ที่ใช้ตัวแปรเสริมเข้ามาช่วยในกรณีที่ตัวแปรที่สนใจ y และตัวแปรช่วยx มีความสัมพันธ์กัน โดยใช้อัตราส่วนรวมในการช่วยประมาณค่าเฉลี่ยของประชากรมีสูตรดังนี้ (Cochran, 1977)

$$\hat{\bar{Y}}_{CR} = \frac{\sum_{h=1}^{k} w_h \bar{y}_h}{\sum_{h=1}^{k} w_h \bar{x}_h} \cdot \bar{X} \tag{4.17}$$

ตัวประมาณค่านี้มีค่าความเอนเอียงและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเท่ากับ

$$\mathrm{Bias}\left(\hat{\bar{Y}}_{CR}\right) = \frac{1}{\bar{X}} \sum_{h=1}^{k} w_h^2 \gamma_h \left(RS_{xh}^2 - S_{yxh}\right) \tag{4.18}$$

และ $$\mathrm{MSE}\left(\hat{\bar{Y}}_{CR}\right) = \sum_{h=1}^{k} w_h^2\gamma_h\left(S_{yh}^2 - 2RS_{yxh} + R^2S_{xh}^2\right) \tag{4.19}$$

เมื่อ $\gamma_h = \frac{1}{n_h} - \frac{1}{N_h}$, n_h คือ ขนาดของตัวอย่างในชั้นภูมิที่ h, S_{yh}^2 เป็นค่าความแปรปรวนของประชากรของตัวแปร y ในชั้นภูมิ h, S_{xh}^2 เป็นค่าความแปรปรวนของประชากรของตัวแปรช่วยx ในชั้นภูมิที่ h, S_{yxh} เป็นความแปรปรวนร่วมของตัวแปรที่สนใจ y และตัวแปรช่วยx ในชั้นภูมิที่ h และ R เป็นอัตราส่วนของ $\frac{\bar{Y}}{\bar{X}}$

4.2.3 ตัวประมาณค่าเฉลี่ยของคาดีลาร์และซินจิ

เป็นตัวประมาณค่าที่คาดีลาร์และซินกิเสนอขึ้นในปี 2005 ซึ่งตัวประมาณค่าของ $\hat{\bar{Y}}_{KC}$ จะมีค่าต่ำกว่าตัวประมาณค่าของ $\bar{Y}_{CR}$ เสมอ เนื่องจากค่า κ ที่นำมาใช้ปรับตัวประมาณค่ามีค่ามากกว่า 0 และน้อยกว่า 1 ในทำนองเดียวกันค่าของ $\mathrm{MSE}\left(\hat{\bar{Y}}_{KC}\right)$ ภายใต้สถานการณ์เดียวกัน จะมีค่าต่ำกว่า $\mathrm{MSE}\left(\hat{\bar{Y}}_{CR}\right)$ เสมอ

คาดีลาร์และซินกิ (Kadilar และ Cingi, 2005) ได้ทำการหาตัวประมาณค่าที่ทำให้ MSE ของตัวประมาณค่า $\hat{\bar{Y}}_{KC}$ มีค่าต่ำสุด ตัวประมาณค่าของคาดีลาร์และซินจิ มีค่าเท่ากับ

$$\hat{\bar{Y}}_{KC} = \kappa\frac{\bar{y}_{st}}{\bar{x}_{st}}\bar{X} \tag{4.20}$$

ค่าพารามิเตอร์ที่เหมาะสมที่สุดของ κ ที่ทำให้ค่า MSE มีค่าน้อยที่สุด คือ

$$\kappa = \frac{\bar{Y}^2}{\bar{Y}^2 + \sum_{h=1}^{k} w_h^2\gamma_h\left(S_{yh}^2 - 2RS_{yxh} + R^2S_{xh}^2\right)} \tag{4.21}$$

ค่าความเอนเอียงและค่าความแปรปรวนของตัวประมาณค่าของ $\hat{\bar{Y}}_{KC}$ ได้แก่

$$\mathrm{Bias}\left(\hat{\bar{Y}}_{KC}\right) = (\kappa - 1)\bar{Y} + \frac{1}{\bar{X}}\sum_{h=1}^{k} w_h^2\gamma_h\left(RS_{xh}^2 - \kappa S_{yxh}\right) \tag{4.22}$$

และ $$\mathrm{MSE}\left(\hat{\bar{Y}}_{KC}\right) = \kappa^2\sum_{h=1}^{k} w_h^2\gamma_h\left(S_{yh}^2 - 2RS_{yxh} + R^2S_{xh}^2\right) + (\kappa - 1)^2\,\bar{Y}^2 \tag{4.23}$$

4.2.4 ตัวประมาณค่าของซิสโซเดียและดวายดิที่ปรับปรุงใหม่ (Kadilar and Cingi, 2003)

คาดีลาร์และซินจิได้ปรับตัวประมาณโดยใช้อัตราส่วนรวมในการประมาณตัวประมาณค่าในการสุ่มแบบอย่างง่าย ตัวประมาณค่าของซิสโซเดียและดวายดิ (Sisodia and Dwivedi, 1981) และของยูเพดยายาและสิงห์ (Upadhyaya and Singh, 1999) นำมาใช้ในการประมาณค่าของประชากรในการชักตัวอย่างแบบชั้นภูมิ ตัวประมาณค่าของยูเพดยายาและสิงห์ไม่ได้นำมาเป็นตัวเปรียบเทียบในกรณีนี้ เนื่องจากตัวประมาณที่สร้างขึ้นนั้น ต้องมีการประมาณค่าความโด่ง (kurtosis) ของ x การประมาณค่าของความโด่งเป็นการประมาณค่าของโมเมนต์ที่ 4 ซึ่งจะถูกกระทบและเปลี่ยนค่าอย่างรวดเร็วเมื่อมีการเปลี่ยนแปลงเล็กน้อยในข้อมูล และต้องการขนาดตัวอย่างที่ใหญ่ เพื่อที่จะให้ได้ตัวประมาณค่าที่มีความถูกต้องใกล้เคียงกับการประมาณค่าของโมเมนต์ที่ 2 ดังนั้น ตัวประมาณค่าของซิสโซเดียกับของดวายดิที่จะถูกนำมาเปรียบเทียบผลกับตัวประมาณค่าที่ได้พัฒนาขึ้นใหม่ ซึ่งตัวประมาณค่าเฉลี่ยของประชากร ความเอนเอียง และค่ากำลังสองเฉลี่ยของความคลาดเคลื่อน มีค่าดังนี้

$$\hat{\bar{Y}}_{SD} = \bar{y}_{st}\frac{\bar{x}_{SD}}{\bar{X}_{SD}} \tag{4.24}$$

$$\mathrm{Bias}\left(\hat{\bar{Y}}_{SD}\right) = \frac{1}{\bar{X}_{SD}}\left\{\sum_{h=1}^{k} w_h^2\gamma_h\left(R_{SD}S_{xh}^2 - S_{yxh}\right)\right\} \tag{4.25}$$

$$\mathrm{MSE}\left(\hat{\bar{Y}}_{SD}\right) = \sum_{h=1}^{k} w_h^2\gamma_h\left(2S_{yh}^2 - 2R_{SD}S_{yxh} + R_{SD}^2S_{xh}^2\right) \tag{4.26}$$

เมื่อ $\bar{x}_{SD} = \sum_{h=1}^{k} w_h\left(\bar{x}_h + C_{xh}\right)$, $\bar{X}_{SD} = \sum_{h=1}^{k} w_h\left(\bar{X}_h + C_{xh}\right)$, $R_{SD} = \dfrac{\bar{y}_{st}}{\bar{X}_{SD}}$ and C_{xh} คือ ค่าสัมประสิทธิ์การกระจายของประชากร (population coefficient of variation) ของตัวแปรช่วย x ในชั้นภูมิที่ h

4.3 ผลการเปรียบเทียบค่าความไม่เอนเอียงและค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนของตัวประมาณค่า

ในการเปรียบเทียบตัวประมาณค่าที่พัฒนาขึ้นมาทั้ง 4 ตัวเปรียบเทียบกับตัวประมาณค่าที่มีอยู่แล้ว 4 ตัวคือ ตัวประมาณค่าเฉลี่ยของชั้นภูมิโดยใช้น้ำหนัก $\dfrac{N_h}{N}$ และไม่ได้ใช้ตัวแปรเสริมมาช่วยในการประมาณค่าตามสูตร (4.15) ในหัวข้อ 4.2.1 ตัวที่สองคือ ตัวประมาณค่าที่ใช้ตัวแปรเสริมมาช่วยในการประมาณค่าโดยวิธีอัตราส่วนรวม ซึ่งเป็นตัวประมาณค่าที่ใช้กันอยู่ทั่วไปในการชักตัวอย่างแบบชั้นภูมิ (Cochran, 1977) ตามสูตร (4.17) ในหัวข้อ

4.2.2 ตัวที่สามคือ ตัวประมาณค่าเฉลี่ยของ คาดีลาร์ และซินจิ $\hat{\hat{Y}}_{KC}$ ตามที่ได้กล่าวแล้วในหัวข้อ 4.2.3 ตามสูตร (4.20) และตัวที่ 4 คือ ตัวประมาณค่าของ ซิสโซเดีย และดวายดิ ในหัวข้อ 4.2.4 ตามสูตร (4.24)

การเปรียบเทียบตัวประมาณค่ากระทำใน 2 รูปแบบ รูปแบบที่ 1 ใช้วิธีการจำลองข้อมูลแสดงในตัวอย่างที่ 1 ซึ่งผลการจำลองข้อมูลปรากฏอยู่ในตารางที่ 4.1-4.4 รูปแบบที่ 2 ได้แก่ ตัวอย่างที่ 2-4 ใช้เปรียบเทียบกับข้อมูลจริงที่นำมาจากหนังสือของครอกแครน (Cochran, 1977) และสิงห์ และแมนแกต (Singh and Mangat, 1996) ข้อมูลในตารางที่ 4.5 เป็นข้อมูลของตัวอย่างที่ 2 ซึ่งตัวแปร y แทนจำนวนพื้นที่เป็นเอเคอร์ในการปลูกข้าวโพด และ x แทนจำนวนพื้นที่เป็นเอเคอร์ของฟาร์มทั้งหมด ตัวอย่างนี้มีการแบ่งประชากรออกเป็น 2 ชั้นภูมิ ชั้นภูมิที่ 1 มีจำนวน 70 ฟาร์ม ในชั้นภูมิที่ 2 มีจำนวน 30 ฟาร์ม ซึ่งการจัดสรรขนาดตัวอย่างเป็นการจัดสรรแบบเหมาะสมที่สุดตามวิธีของเนย์แมน (Neyman sample allocation) ข้อมูลชุดนี้มาจากหนังสือของครอกแครน (Cochran, 1977 : 168) ตัวอย่างที่ 3 เป็นข้อมูลที่มาจากหนังสือของ สิงห์และแมนแกต (Singh และ Mangat, 1996 : 220) ซึ่งแสดงไว้ในตารางที่ 4.6 ตัวแปร y แทนจำนวนของวัวนมในปี 1993 และตัวแปรช่วยx เป็นจำนวนวัวนมในปี 1990 ในตัวอย่างที่ 4 เป็นข้อมูลที่ได้จากการทำสำมะโนของสวนพืชที่ทำในเชิงธุรกิจในนอร์ทคาโรไลนา ในปี 1946 (North Carolina; 1946) ซึ่งเป็นข้อมูลที่นำมาจากหนังสือของครอกแครน (Cochran, 1977 : 173) โดยตัวแปรช่วยx แทนจำนวนต้นพืชในสวน และตัวแปรที่สนใจ y แทนผลผลิตของพืชเป็นบุชเชล (bushels) ข้อมูลแสดงอยู่ในตารางที่ 4.7

4.3.1 ผลการศึกษาจากข้อมูลจำลอง

การศึกษาเพื่อเปรียบเทียบค่าความเอนเอียงและประสิทธิภาพของตัวประมาณค่าในตัวอย่างที่ 1 ใช้ข้อมูลที่สร้างขึ้นจากการจำลอง 2 ประชากรที่มีการแจกแจงแบบปกติ ตัวแปรที่สนใจ y มีค่าเฉลี่ยเท่ากับ 150 และความแปรปรวนเท่ากับ 500 และตัวแปรช่วยx มีค่าเฉลี่ยเท่ากับ 50 และความแปรปรวนเท่ากับ 45 ตัวแปรที่สนใจ y และตัวแปรช่วยx ในการสร้างข้อมูลสำหรับการจำลองได้กำหนดค่าสัมประสิทธิ์สหสัมพันธ์ไว้ 4 ค่า คือ เท่ากับ 0.3, 0.5, 0.7 และ 0.9 ข้อมูลจำลองแต่ละกรณีมีจำนวน 1 ล้านหน่วย หลังจากนั้นทำการแบ่งประชากรทุกชุดออกเป็น 3 ชั้นภูมิ การแบ่งชั้นภูมินี้ได้ใช้วิธีการของแดลเนียสและฮอจส์ (Dalenius และ Hodges, 1959) ด้วยการปรับทำซ้ำขอบเขตของชั้นภูมิ เพื่อให้สัมประสิทธิ์การกระจาย (coefficient of variation) ของ x_h ในแต่ละชั้นภูมิมีค่าใกล้เคียงกัน เมื่อแบ่งประชากรเรียบร้อยแล้วได้คำนวณค่าสถิติต่าง ๆ ของชั้นภูมิ ดังแสดงในตารางที่ 4.1 การเปรียบเทียบจะใช้การจัดสรรขนาดตัวอย่าง 2 วิธี คือ การจัดสรรขนาดของตัวอย่างตามวิธีการของเนย์แมน

และการจัดสรรขนาดตัวอย่างเท่ากันทุกชั้นภูมิ โดยกำหนดขนาดของตัวอย่างเท่ากับ 30 และ 120 ดังนั้น กรณีที่ทำการศึกษาจะมีทั้งหมดจำนวน 16 กรณี

ตารางที่ 4.1 พารามิเตอร์ของประชากรเมื่อมีการชักตัวอย่างแบบชั้นภูมิและการจัดสรรตัวอย่างแบบเนย์แมนตามค่าสัมประสิทธิ์สหสัมพันธ์ของประชากร

ρ	ชั้นภูมิ	พารามิเตอร์ของชั้นภูมิ							การจัดสรรแบบเนย์แมน	
		N_h	$\bar{X}_h$	$\bar{Y}_h$	S^2_{yh}	S_{xyh}	S^2_{xh}	$CV_{\bar{x}_h}$	n_h	n_h
0.3	1	297,684	57.80	157.79	466.90	12.06	11.83	0.06	9	37
	2	612,553	47.99	148.05	465.60	11.35	11.44	0.07	19	74
	3	89,763	37.90	137.79	460.45	7.36	7.46	0.07	2	9
	รวม	1,000,000	50.00						30	120
0.5	1	333,998	57.30	162.18	409.24	21.13	12.60	0.06	11	45
	2	547,058	47.99	146.69	399.01	14.47	8.76	0.06	16	62
	3	118,944	38.80	131.25	394.58	13.37	8.14	0.07	3	13
	รวม	1,000,000	50.00						30	120
0.7	1	376,040	56.77	165.82	327.91	31.48	13.50	0.06	12	50
	2	551,268	47.08	143.17	309.73	23.67	10.16	0.07	16	63
	3	72,692	37.25	120.28	292.30	16.10	6.94	0.07	2	7
	รวม	1,000,000	50.01						30	120
0.9	1	320,983	57.49	172.45	206.35	37.14	12.39	0.06	10	42
	2	578,763	47.89	143.70	185.40	30.23	10.08	0.07	17	68
	3	100,254	38.26	114.75	163.76	22.90	7.60	0.07	3	10
	รวม	1,000,000	50.01						30	120

พารามิเตอร์ของประชากรที่ได้จากการจำลอง และการจัดสรรตัวอย่างของเนย์แมน สรุปไว้ในตารางที่ 4.1 โดยการจำแนกตามสัมประสิทธิ์สหสัมพันธ์ของประชากร จะเห็นได้ว่า ขนาดตัวอย่างโดยใช้การจัดสรรตัวอย่างของเนย์แมน มีขนาดใกล้เคียงกับขนาดตัวอย่าง โดยใช้การจัดสรรตามขนาดของชั้นภูมิ (proportional to size) น้ำหนักของชั้นภูมิ w_h w_{yh} และ w_{xh} ที่ใช้พารามิเตอร์ของประชากรของ y และ x สำหรับการจัดสรรตัวอย่างทั้ง 2 วิธี ในกรณีที่ n = 30 และ 120 ได้แสดงไว้ในตารางที่ 4.2 ซึ่งจะเห็นได้ว่าน้ำหนักของชั้นภูมิ w_h, $\hat{w}^*_{xh}$ และ $\hat{w}^*_y$ มีความแตกต่างไม่มากนักในการจัดสรรตัวอย่างแบบเนย์แมน แต่จะมีความแตกต่างกันอย่างชัดเจนในการจัดสรรตัวอย่างที่มีขนาดเท่ากัน ความแตกต่างระหว่าง $\hat{w}^*_{xh}$ และ $\hat{w}^*_{yh}$ จะลดลงเมื่อค่า ρ มีค่าเข้าใกล้ 1 จากตารางที่ 4.2 ยังมีข้อสังเกตอีกข้อหนึ่งว่า การเปลี่ยนแปลงของน้ำหนัก $\hat{w}^*_{xh}$ และ $\hat{w}^*_{yh}$ เมื่อใช้วิธีการจัดสรรตัวอย่างที่มีขนาดเท่ากันจะน้อยกว่าในกรณีการจัดสรรขนาดตัวอย่างแบบเนย์แมน เมื่อขนาดตัวอย่างเพิ่มมากขึ้นจาก 30 เป็น 120

สัมประสิทธิ์สหสัมพันธ์ส่วนรวม และค่าที่อยู่ทางด้านขวาของอสมการ (3.2) ได้ประมาณค่าแสดงไว้ในตารางที่ 4.3 ในกรณี $\rho = 0.3$ และ 0.5 โดยใช้น้ำหนัก w_h, $\hat{w}^*_{yh}$ และ $\hat{w}^*_{xh}$ จากการจัดสรรตัวอย่างแบบเนย์แมนและการจัดสรรตัวอย่างที่เท่ากัน ทั้งขนาดตัวอย่าง n

ตารางที่ 4.2 น้ำหนักของชั้นภูมิจำแนกตามสัมประสิทธิ์สหสัมพันธ์ของประชากรและการจัดสรรขนาดตัวอย่าง

ρ	ชั้นภูมิ	w_h	การจัดสรรขนาดตัวอย่างแบบเนย์แมน			
			n = 30		n = 120	
			$\hat{w}^*_{yh}$	$\hat{w}^*_{xh}$	$\hat{w}^*_{yh}$	$\hat{w}^*_{xh}$
0.3	1	0.2977	0.2871	0.2990	0.2920	0.3089
	2	0.6126	0.6409	0.6098	0.6258	0.5904
	3	0.0898	0.0719	0.0913	0.0822	0.1007
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000
0.5	1	0.3340	0.3350	0.3164	0.3411	0.3251
	2	0.5471	0.5509	0.5819	0.5345	0.5649
	3	0.1189	0.1141	0.1016	0.1244	0.1100
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000
0.7	1	0.3760	0.3784	0.3788	0.3734	0.3735
	2	0.5513	0.5473	0.5450	0.5565	0.5560
	3	0.0727	0.0743	0.0762	0.0701	0.0705
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000
0.9	1	0.3210	0.3250	0.3262	0.3190	0.3214
	2	0.5788	0.5701	0.5676	0.5827	0.5779
	3	0.1003	0.1049	0.1062	0.0983	0.1007
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000
ρ	ชั้นภูมิ	w_h	การจัดสรรขนาดตัวอย่างเท่ากัน			
			n = 30		n = 120	
			$\hat{w}^*_{yh}$	$\hat{w}^*_{xh}$	$\hat{w}^*_{yh}$	$\hat{w}^*_{xh}$
0.3	1	0.2977	0.4176	0.4490	0.4323	0.4497
	2	0.6126	0.3343	0.3122	0.3346	0.3123
	3	0.0898	0.2482	0.2388	0.2331	0.2381
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000
0.5	1	0.3340	0.4309	0.4121	0.4368	0.4128
	2	0.5471	0.3360	0.3877	0.3361	0.3878
	3	0.1189	0.2331	0.2002	0.2272	0.1994
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000
0.7	1	0.3760	0.4779	0.4668	0.4812	0.4677
	2	0.5513	0.3399	0.3683	0.3400	0.3685
	3	0.0727	0.1823	0.1649	0.1788	0.1638
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000
0.9	1	0.3210	0.4382	0.4351	0.4394	0.4358
	2	0.5788	0.3416	0.3490	0.3417	0.3491
	3	0.1003	0.2202	0.2159	0.2190	0.2151
	รวม	1.0000	1.0000	1.0000	1.0000	1.0000

= 30 และ 120 ปรากฏว่า เมื่อค่า $\rho = 0.3$ เงื่อนไขในอสมการ (3.2) ไม่สอดคล้องในการจัดสรรตัวอย่างทั้งสองวิธีและขนาดตัวอย่างทั้งสอง แต่เมื่อค่า $\rho = 0.5$ เงื่อนไขเป็นไปตามอสมการ (3.2) ในทุกกรณี ดังนั้น จึงสรุปได้ว่า เมื่อ ρ มีค่าเท่ากับ 0.3 ตัวประมาณค่าจะใช้

$\hat{\hat{Y}}_{wy}$ สำหรับตัวประมาณค่าที่เสนอใหม่ ซึ่งจะเปรียบเทียบกับตัวประมาณค่าที่ใช้อยู่ในปัจจุบัน $\hat{\hat{Y}}_{st}$ และเมื่อ ρ มีค่าตั้งแต่ 0.5 ขึ้นไป ตัวประมาณค่าที่เสนอใหม่จะเป็น $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{Rwx}$ และ $\hat{\hat{Y}}_{Rwyx}$ โดยจะเปรียบเทียบกับ $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$

ตารางที่ 4.3 สัมประสิทธิ์สหสัมพันธ์ส่วนรวมของประชากรและค่าทางด้านขวามือของสมการที่ (3.2) จำแนกตามค่า ρ ขนาดตัวอย่าง และน้ำหนักของชั้นภูมิ

ρ	ขนาดตัวอย่าง	การจัดสรรขนาดตัวอย่างแบบเนย์แมน					
		w_h		$\hat{w}^*_{yh}$		$\hat{w}^*_{xh}$	
		$\hat{\rho}_c$	RHS(3.2)	$\hat{\rho}_c$	RHS(3.2)	$\hat{\rho}_c$	RHS(3.2)
0.3	30	0.1543	0.2314	0.1553	0.2328	0.1542	0.2313
	120	0.1546	0.2319	0.1550	0.2325	0.1540	0.2309
0.5	30	0.2598	0.2348	0.2600	0.2348	0.2584	0.2336
	120	0.2595	0.2345	0.2601	0.2350	0.2588	0.2340

ρ	ขนาดตัวอย่าง	การจัดสรรขนาดตัวอย่างเท่ากัน					
		w_h		$\hat{w}^*_{yh}$		$\hat{w}^*_{yh}$	
		$\hat{\rho}_c$	RHS(3.2)	$\hat{\rho}_c$	RHS(3.2)	$\hat{\rho}_c$	RHS(3.2)
0.3	30	0.1564	0.2351	0.1542	0.2351	0.1551	0.2308
	120	0.1564	0.2351	0.1550	0.2313	0.1552	0.2308
0.5	30	0.2586	0.2337	0.2711	0.2443	0.2687	0.2421
	120	0.2586	0.2337	0.2717	0.2446	0.2687	0.2421

ตารางที่ 4.4 แสดงประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าที่เสนอใหม่โดยเปรียบเทียบกับตัวประมาณค่า $\hat{\bar{y}}_{st}$ จากวิธีการจัดสรรขนาดตัวอย่างแบบเนย์แมน เมื่อ $\rho = 0.3$ และเปรียบเทียบกับตัวประมาณค่า $\hat{\hat{Y}}_{CR}$ จากวิธีการจัดสรรขนาดตัวอย่างแบบเนย์แมน เมื่อ ρ มีค่าตั้งแต่ 0.5 ขึ้นไป

ประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าตัวใหม่ $\hat{\hat{Y}}_{st\_wy}$ เป็น 1.01 และ 1.00 สำหรับ n = 30 และ 120 ตามลำดับ เมื่อ $\rho = 0.3$ โดยมีการจัดสรรขนาดตัวอย่างแบบเนย์แมน แต่เมื่อมีการจัดสรรขนาดตัวอย่างแบบเท่ากัน ประสิทธิภาพสัมพัทธ์ของ $\hat{\hat{Y}}_{st}$ ลดลงมาเหลือเท่ากับ 0.71 หรือลดลงร้อยละ 29 จากกรณีการจัดสรรตัวอย่างแบบเนย์แมน ทั้งขนาดตัวอย่างเท่ากับ 30 และ 120 ส่วนประสิทธิภาพสัมพัทธ์ของ $\hat{\hat{Y}}_{st\_wy}$ ลดลงเหลือเพียง 0.97 หรือร้อยละ 4 เมื่อ n = 30 และ 0.95 หรือร้อยละ 5 เมื่อ n = 120 ซึ่งแสดงให้เห็นว่า $\hat{\hat{Y}}_{st\_wy}$ มีประสิทธิภาพสูงกว่า $\hat{\hat{Y}}_{st}$ เมื่อมีการจัดสรรตัวอย่างที่มีขนาดเท่ากัน และประสิทธิภาพของ $\hat{\hat{Y}}_{st\_wy}$ เปลี่ยนแปลงน้อยกว่า $\hat{\hat{Y}}_{st}$ มาก เมื่อการจัดสรรขนาดตัวอย่างเปลี่ยนจากการจัดสรรแบบเนย์แมนเป็นการจัดสรรขนาดตัวอย่างแบบเท่ากัน

ตารางที่ 4.4 ร้อยละของความเอนเอียงและประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าจำแนกตาม ρ ขนาดตัวอย่าง และวิธีการจัดสรรขนาดตัวอย่าง

การจัดสรรขนาดตัวอย่างแบบเนย์แมน										
ρ	ขนาดตัวอย่าง	ประสิทธิภาพ	$\overline{y}_{st}$	$\hat{\bar{Y}}_{CR}$	$\hat{\bar{Y}}_{KC}$	$\hat{\bar{Y}}_{SD}$	$\hat{\bar{Y}}_{Rwy}$	$\hat{\bar{Y}}_{Rwx}$	$\hat{\bar{Y}}_{Rwyx}$	$\hat{\bar{Y}}_{st\_wy}$
0.3	30	ความเอนเอียง,%		0.01	0.06	0.01	0.01	0.01	0.01	0.01
		ประสิทธิภาพสัมพัทธ์	1.00	0.93	0.93	0.93	0.94	0.93	0.94	1.01
	120	ความเอนเอียง,%		0.00	0.02	0.00	0.00	0.00	0.00	0.02
		ประสิทธิภาพสัมพัทธ์	1.00	0.93	0.93	0.93	0.93	0.93	0.93	1.00
0.5	30	ความเอนเอียง,%		0.01	0.05	0.01	0.01	0.01	0.01	
		ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.00	1.00	1.00	1.00	
	120	ความเอนเอียง,%		0.00	0.01	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.00	1.00	1.00	1.00	
0.7	30	ความเอนเอียง,%		0.00	0.04	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.00	1.00	1.00	1.00	
	120	ความเอนเอียง,%		0.00	0.01	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.00	1.00	1.00	1.00	
0.9	30	ความเอนเอียง,%		0.00	0.01	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.00	1.00	1.00	1.00	
	120	ความเอนเอียง,%		0.00	0.00	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.00	1.00	1.00	1.00	

ตารางที่ 4.4 ร้อยละของความเอนเอียงและประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าจำแนกตาม ρ ขนาดตัวอย่าง และวิธีการจัดสรรขนาดตัวอย่าง (ต่อ)

การจัดสรรขนาดตัวอย่างเท่ากัน										
ρ	ขนาดตัวอย่าง	ประสิทธิภาพ	$\overline{y}_{st}$	$\hat{\hat{Y}}_{CR}$	$\hat{\hat{Y}}_{KC}$	$\hat{\hat{Y}}_{SD}$	$\hat{\hat{Y}}_{Rwy}$	$\hat{\hat{Y}}_{Rwx}$	$\hat{\hat{Y}}_{Rwyx}$	$\hat{\hat{Y}}_{st\_wy}$
0.3	30	ความเอนเอียง,%		0.01	0.09	0.01	0.01	0.01	0.01	0.22
		ประสิทธิภาพสัมพัทธ์	0.71	0.66	0.66	0.66	0.90	0.88	0.90	0.97
	120	ความเอนเอียง,%		0.00	0.02	0.00	0.00	0.00	0.00	0.07
		ประสิทธิภาพสัมพัทธ์	0.71	0.66	0.66	0.66	0.88	0.88	0.88	0.95
0.5	30	ความเอนเอียง,%		0.01	0.07	0.01	0.01	0.01	0.01	
		ประสิทธิภาพสัมพัทธ์		0.79	0.79	0.79	0.95	0.93	0.95	
	120	ความเอนเอียง,%		0.00	0.02	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		0.79	0.79	0.79	0.94	0.93	0.94	
0.7	30	ความเอนเอียง,%		0.00	0.05	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		0.74	0.74	0.74	0.89	0.88	0.88	
	120	ความเอนเอียง,%		0.00	0.01	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		0.75	0.75	0.75	0.89	0.88	0.88	
0.9	30	ความเอนเอียง,%		0.00	0.02	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		0.74	0.74	0.74	0.93	0.93	0.93	
	120	ความเอนเอียง,%		0.00	0.00	0.00	0.00	0.00	0.00	
		ประสิทธิภาพสัมพัทธ์		0.75	0.75	0.75	0.94	0.94	0.94	

เมื่อ ρ มีค่าเท่ากับ 0.5 0.7 และ 0.9 ตัวประมาณค่าอัตราส่วนรวม $\hat{\hat{Y}}_{Rwy}$ $\hat{\hat{Y}}_{wx}$ และ $\hat{\hat{Y}}_{Rwyx}$ มีประสิทธิภาพใกล้เคียงกับ $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ ในกรณีการจัดสรรขนาดตัวอย่างแบบเนย์แมนทั้ง 2 ขนาดตัวอย่าง แต่เมื่อใช้การจัดสรรขนาดตัวอย่างแบบเท่ากัน $\hat{\hat{Y}}_{Rwy}$ มีประสิทธิภาพสูงสุดและค่าประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าอัตราส่วนรวมที่เสนอใหม่ทั้ง 3 ตัว มีค่าสูงกว่าตัวประมาณค่า $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ เมื่อการจัดสรรขนาดตัวอย่างมีจำนวนเท่ากันทุกชั้นภูมิ ประสิทธิภาพสัมพัทธ์ของตัวประมาณค่า $\hat{\hat{Y}}_{Rwy}$ อยู่ระหว่าง 0.89-0.95 สำหรับ n = 30 และ 0.88-0.94 สำหรับ n = 120 ประสิทธิภาพสัมพัทธ์ของ $\hat{\hat{Y}}_{Rwx}$

อยู่ระหว่าง 0.88-0.93 สำหรับ n = 30 และ n = 120 ประสิทธิภาพสัมพัทธ์ของ $\hat{\hat{Y}}_{Rwyx}$ อยู่ระหว่าง 0.88-0.95 สำหรับ n = 30 และ 0.88-0.94 สำหรับ n = 120 ส่วนตัวประมาณค่า $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ มีค่าอยู่ระหว่าง 0.74-0.79 และ 0.75-0.79 สำหรับ n = 30 และ 120 ตามลำดับ หรืออาจกล่าวอีกนัยหนึ่งก็คือ ประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าที่เสนอใหม่ $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{wx}$ และ $\hat{\hat{Y}}_{Rwyx}$ ที่ใช้วิธีการจัดสรรขนาดตัวอย่างแบบเท่ากันมีค่าลดลงไม่มากกว่าร้อยละ 12 เมื่อเปรียบเทียบกับประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าที่ได้จากวิธีการจัดสรรขนาดตัวอย่างแบบเนย์แมน แต่ประสิทธิภาพสัมพัทธ์ของตัวประมาณค่า $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ ลดลงถึงร้อยละ 26 จึงสรุปจากผลการจำลองได้ว่า ประสิทธิภาพของตัวประมาณค่า $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{Rwx}$ และ $\hat{\hat{Y}}_{Rwxy}$ อย่างน้อยเท่ากับประสิทธิภาพของ $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ การเปลี่ยนแปลงของประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าอัตราส่วนรวมทั้งหมดแทบจะไม่มีการเปลี่ยนแปลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 30 เป็น 120 ของการจำลองข้อมูลทุกกรณี ประสิทธิภาพของตัวประมาณที่พัฒนาขึ้นใหม่ทั้ง 4 ตัวที่ใช้ค่าของน้ำหนักขึ้นอยู่กับค่าสถิติของชั้นภูมินั้น อาจสรุปได้ว่ามีประสิทธิภาพอย่างน้อยเท่ากับตัวประมาณค่าที่มีอยู่แล้ว เพื่อจัดสรรขนาดตัวอย่างแบบเนย์แมนและแบบเท่ากัน ส่วนความเอนเอียงของตัวประมาณค่านั้น พบว่าความเอนเอียงสัมบูรณ์ของตัวประมาณค่าอัตราส่วนรวมทุกตัวที่เปรียบเทียบกันมีค่าอยู่ระหว่างร้อยละ 0.00-0.01 ยกเว้นตัวประมาณค่า $\hat{\hat{Y}}_{KC}$ ที่มีค่าความเอนเอียงสัมบูรณ์ โตกว่าตัวประมาณค่าตัวอื่น ๆ ซึ่งอยู่ระหว่างร้อยละ 0.00-0.09 แต่ความเอนเอียงสัมบูรณ์ของ $\hat{\hat{Y}}_{st\_wy}$ มีค่าอยู่ระหว่าง 0.01-0.02 เมื่อมีการจัดสรรขนาดตัวอย่างแบบเนย์แมนและอยู่ระหว่าง 0.07-0.22 เมื่อมีการจัดสรรขนาดตัวอย่างแบบเท่ากัน การที่ร้อยละความเอนเอียงสัมบูรณ์ของ $\hat{\hat{Y}}_{st\_wy}$ มีค่าเพิ่มเล็กน้อยเป็น 0.02 คงเป็นผลจากความคลาดเคลื่อนในการจำลอง

ตารางที่ 4.5 ข้อมูลของตัวอย่างที่ 2

ชั้นภูมิ	N_h	S^2_{yh}	S_{yxh}	S^2_{xh}	$\bar{Y}_h$	$\bar{X}_h$	n_h
1	1580	312	494	2055	82.56	19.40	70
2	430	922	858	7357	244.85	51.63	30
ประชากร	2010	620	1453	7619	117.28	26.30	100

แหล่งที่มา : Cochran, 1977 : 168

ตารางที่ 4.6 ข้อมูลของตัวอย่างที่ 3

ชั้นภูมิ	N_h	S^2_{yh}	S_{yxh}	S^2_{xh}	$\bar{Y}_h$	$\bar{X}_h$	n_h
1	1260	17.62	14.69	20.91	17.43	15.29	7
2	2400	18.39	5.25	30.21	20.42	17.25	12
3	1150	13.30	5.90	10.70	20.60	17.80	5
ประชากร	4810						24

แหล่งที่มา : Singh and Mangat, 1996 : 220

ตารางที่ **4.7** ข้อมูลของตัวอย่างที่ 4

ชั้นภูมิ	S^2_{yh}	S_{yxh}	S^2_{xh}	$\bar{Y}_h$	$\bar{X}_h$
1	8699	6462	5186	69.48	53.80
2	4614	3100	2367	43.64	31.07
3	7311	4817	4877	66.39	56.97
ประชากร	6409	4434	3898	56.47	44.45

ชั้นภูมิ	N_h	n_h เป็นสัดส่วนกับ			
		N_h	$N_h S_{yh}$	$N_h\sqrt{\bar{X}_h}$	$N_h\bar{X}_h$
1	47	18	22	20	22
2	118	46	40	39	32
3	91	36	38	41	46
ประชากร	256	100	100	100	100

แหล่งที่มา : Cochran, 1977 : 173

ตารางที่ 4.8 น้ำหนักของชั้นภูมิของตัวอย่างที่ 2-4 ตามวิธีการจัดสรรขนาดตัวอย่าง

ตัวอย่าง	วิธีการจัดสรรตัวอย่าง	ชั้นภูมิ	w_h	$\hat{w}^*_{yh}$	$\hat{w}^*_{xh}$
ตัวอย่างที่ 2	เนย์แมน	1	0.7861	0.7896	0.7871
		2	0.2139	0.2104	0.2129
	เท่ากัน	1	0.7861	0.8026	0.8042
		2	0.2139	0.1974	0.1958
ตัวอย่างที่ 3	เนย์แมน	1	0.2620	0.2824	0.2909
		2	0.4990	0.4556	0.3285
		3	0.2391	0.2620	0.3806
	เท่ากัน	1	0.2620	0.3038	0.3094
		2	0.4990	0.2916	0.1859
		3	0.2391	0.4047	0.5047
วิธีการจัดสรรขนาดตัวอย่าง ในตัวอย่างที่ 4					
n_h เป็นสัดส่วนกับ N_h		1	0.1836	0.1554	0.1556
		2	0.4609	0.4834	0.4855
		3	0.3555	0.3612	0.3590
n_h เป็นสัดส่วนกับ $N_h S_{yh}$		1	0.1836	0.1887	0.1939
		2	0.4609	0.4620	0.4698
		3	0.3555	0.3494	0.3363
n_h เป็นสัดส่วนกับ $N_h \sqrt{\bar{X}_h}$		1	0.1836	0.1520	0.1583
		2	0.4609	0.4622	0.4715
		3	0.3555	0.3859	0.3702
n_h เป็นสัดส่วนกับ $N_h \bar{X}_h$		1	0.1836	0.1465	0.1635
		2	0.4609	0.4382	0.4564
		3	0.3555	0.4152	0.3801
ขนาดตัวอย่างเท่ากัน		1	0.1836	0.3687	0.3854
		2	0.4609	0.4154	0.4292
		3	0.3555	0.2159	0.1854

ตารางที่ 4.9 การเปรียบเทียบความเอนเอียงและประสิทธิภาพสัมพัทธ์ของค่าประมาณของค่าเฉลี่ยของประชากรในตัวอย่างที่ 2-4

ตัวอย่าง: วิธีการจัดสรรตัวอย่าง	ประสิทธิภาพ	$\overline{y}_{st}$	$\hat{\overline{Y}}_{CR}$	$\hat{\overline{Y}}_{KC}$	$\hat{\overline{Y}}_{SD}$	$\hat{\overline{Y}}_{Rwy}$	$\hat{\overline{Y}}_{Rwx}$	$\hat{\overline{Y}}_{Rwyx}$	$\hat{\overline{Y}}_{st\_w}$
ตัวอย่างที่ 2: เนย์แมน	ความเอนเอียง,%		0.03	0.39	0.03	0.03	0.03	0.03	
	ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.00	1.01	1.00	1.01	
เท่ากัน	ความเอนเอียง,%		0.01	0.43	0.01	0.01	0.00	0.01	
	ประสิทธิภาพสัมพัทธ์		0.96	0.96	0.96	0.96	0.96	0.96	
ตัวอย่างที่ 3: เนย์แมน	ความเอนเอียง,%		0.24	0.08	0.23	0.17	0.21	0.29	0.01
	ประสิทธิภาพสัมพัทธ์	1.00	0.56	0.56	0.58	0.72	0.61	0.70	1.01
เท่ากัน	ความเอนเอียง,%		0.32	0.11	0.31	0.19	0.17	0.19	0.07
	ประสิทธิภาพสัมพัทธ์	0.86	0.42	0.42	0.43	0.69	0.75	0.70	1.05
วิธีการจัดสรรขนาดตัวอย่าง ในตัวอย่างที่ 4			$\hat{\overline{Y}}_{CR}$	$\hat{\overline{Y}}_{KC}$	$\hat{\overline{Y}}_{SD}$	$\hat{\overline{Y}}_{Rwy}$	$\hat{\overline{Y}}_{Rwx}$	$\hat{\overline{Y}}_{Rwyx}$	
n_h เป็นสัดส่วนกับ N_h	ความเอนเอียง,%		0.11	0.15	0.15	0.11	0.11	0.11	
	ประสิทธิภาพสัมพัทธ์		0.96	0.96	0.98	0.94	0.95	0.94	
n_h เป็นสัดส่วนกับ $N_h S_{yh}$	ความเอนเอียง,%		0.10	0.16	0.16	0.09	0.09	0.09	
	ประสิทธิภาพสัมพัทธ์		1.00	1.00	1.02	1.02	1.06	1.04	
n_h เป็นสัดส่วนกับ $N_h \sqrt{\overline{X}_h}$	ความเอนเอียง,%		0.08	0.15	0.15	0.10	0.09	0.09	
	ประสิทธิภาพสัมพัทธ์		1.07	1.07	1.09	0.98	1.02	1.01	
n_h เป็นสัดส่วนกับ $N_h \overline{X}_h$	ความเอนเอียง,%		0.06	0.16	0.16	0.08	0.07	0.09	
	ประสิทธิภาพสัมพัทธ์		1.13	1.14	1.15	1.00	1.07	1.04	
ขนาดตัวอย่างเท่ากัน	ความเอนเอียง,%		0.12	0.19	0.19	0.05	0.06	0.06	
	ประสิทธิภาพสัมพัทธ์		0.82	0.82	0.84	1.46	1.60	1.53	

4.3.2 ผลการศึกษาจากข้อมูลจริง

ตามที่ได้กล่าวมาแล้ว ข้อมูลที่นำมาวิเคราะห์ นำมาจากตำราเกี่ยวกับการชักตัวอย่างของครอกแครน และของสิงห์ และแมนแกต (Cochran, 1977; Singh และMangat 1996) ข้อมูลในตัวอย่างที่ 2 มาจากตำราของครอกแครนปี 1977 (Cochran, 1977) ซึ่งปรากฏอยู่ในตารางที่ 4.5 โดย y เป็นจำนวนพื้นที่มีหน่วยเป็นเอเคอร์ในการปลูกข้าวโพด และตัวแปรช่วยx คือ จำนวนพื้นที่มีหน่วยเป็นเอเคอร์ในการทำฟาร์ม ขนาดตัวอย่างในชั้นภูมิที่ 1 มี 70 ฟาร์ม และในชั้นภูมิที่ 2 มี 30 ฟาร์ม การจัดสรรขนาดตัวอย่างในข้อมูลชุดนี้เป็นแบบการจัดสรรแบบอุตตมะโดยประมาณ ในตัวอย่างที่ 3 ข้อมูลสรุปอยู่ในตารางที่ 4.6 ซึ่งเป็นตัวอย่างใน

หนังสือของสิงห์และแมนแกต (Singh และ Mangat, 1966) ตัวแปร y แทนจำนวนวัวนมในปี 1993 และตัวแปรช่วยx แทนจำนวนวัวนมในปี 1990

สำหรับข้อมูลชุดสุดท้ายในตัวอย่างที่ 4 ปรากฏอยู่ในตารางที่ 4.7 ซึ่งเป็นข้อมูลที่ทำสำมะโนของสวนพืชในการทำเชิงธุรกิจในรัฐนอร์ทคาร์โรไลนาในปี 1946 (Cochran, 1977) ซึ่งแสดงข้อมูลในตารางที่ 4.7 ตัวแปรช่วยx แทนจำนวนต้นพืชในสวน และ y แทนผลผลิตของพืชมีหน่วยเป็นบุชเชล (Bushels) ในตัวอย่างที่ 4 จะใช้วิธีการจัดสรรขนาดตัวอย่าง 5 วิธี เพิ่มจากในหนังสือของครอกแครนอีก 1 วิธี คือ การจัดสรรขนาดตัวอย่างแบบเท่ากัน

ก) การจัดสรรขนาดตัวอย่างให้ n_h เป็นสัดส่วนกับ N_h

ข) การจัดสรรขนาดตัวอย่างให้ n_h เป็นสัดส่วนกับ $N_h S_{yh}$

ค) การจัดสรรขนาดตัวอย่างให้ n_h เป็นสัดส่วนกับ $N_h\sqrt{\bar{X}_h}$

ง) การจัดสรรขนาดตัวอย่างให้ n_h เป็นสัดส่วนกับ $N_h\bar{X}_h$

จ) การจัดสรรขนาดตัวอย่างแบบเท่ากัน $n_h = \dfrac{n}{3}$ เมื่อ $h = 1, 2, 3$ และ n เป็นขนาดตัวอย่างทั้งหมดที่ทำการศึกษา ซึ่งมีค่าเท่ากับ 99

ตามที่ได้กล่าวมาแล้ว ค่าประมาณของค่าเฉลี่ย $\bar{Y}$ เพื่อใช้ในการคำนวณหาน้ำหนักของ w^*_{yh} จะใช้ $\hat{\bar{Y}}_{Rwx}$ โดยใช้น้ำหนัก $\hat{w}^*_x$ ซึ่งคำนวณจากค่าสถิติของตัวแปรช่วยx ค่าน้ำหนักของชั้นภูมิในตัวอย่าง 2-4 สรุปไว้ในตารางที่ 4.8 ค่าน้ำหนักของ $w_h = \dfrac{N_h}{N}$ จะแตกต่างจากค่าน้ำหนักของ w^*_{yh} และ w^*_{xh} แต่น้ำหนักของทั้ง 2 ตัวหลังไม่ค่อยแตกต่างกันจากพารามิเตอร์ของประชากรในชั้นภูมิในตารางที่ 4.5 และน้ำหนักของชั้นภูมิในตารางที่ 4.8 พบว่าค่าของสัมประสิทธิ์สหสัมพัทธ์ส่วนรวม ρ_c ในตัวอย่างที่ 2 มีค่าเท่ากับ 0.51, 0.52 และ 0.52 ในกรณีของการจัดสรรขนาดตัวอย่างของเนย์แมน และมีค่าใกล้เคียงกับ 0.56 ในกรณีของการจัดสรรขนาดตัวอย่างเท่ากันเมื่อใช้ w_h , w^*_{yh} และ w^*_{xh} ตามลำดับ ค่าของ $\dfrac{C_{\bar{x}_{st}}}{2C_{\bar{y}_{st}}}$ มีค่าประมาณเท่ากับ 0.30 ในทุกกรณี และมีค่าน้อยกว่าค่าของ ρ_c ดังนั้น ตัวประมาณค่าแบบอัตราส่วนรวมถูกใช้ในการประมาณค่าเฉลี่ยของประชากรในตัวอย่างที่ 2 ประสิทธิภาพสัมพัทธ์ของ $\hat{\bar{Y}}_{Rwy}, \hat{\bar{Y}}_{Rwx}$ และ $\hat{\bar{Y}}_{Rwyx}$ เมื่อเทียบกับ $\hat{\bar{Y}}_{CR}$ ที่ใช้การจัดสรรขนาดตัวอย่างแบบเนย์แมน มีค่าสูงสุดเท่ากับ 1.01 ประสิทธิภาพสัมพัทธ์ในทุกกรณีที่ใช้ตัวประมาณค่าแบบอัตราส่วนรวมลดลงเป็น 0.96 เมื่อมีการจัดสรรขนาดตัวอย่างเท่ากัน ร้อยละความเอนเอียงสัมพัทธ์ของตัวประมาณค่าแบบอัตราส่วนรวมอยู่ในระดับเดียวกัน ยกเว้นร้อยละความเอนเอียงสัมพัทธ์ของตัวประมาณค่า $\hat{\bar{Y}}_{KC}$ ที่มีค่าในลำดับที่สูงกว่า ในตัวอย่างที่ 2 ตัวประมาณค่าแบบอัตราส่วนรวม $\hat{\bar{Y}}_{Rwy}$ และ $\hat{\bar{Y}}_{Rwyx}$ มีประสิทธิภาพสูงกว่าตัวประมาณค่า $\hat{\bar{Y}}_{CR}$, $\hat{\bar{Y}}_{KC}$ และ $\hat{\bar{Y}}_{SD}$

เล็กน้อยเมื่อใช้การจัดสรรขนาดตัวอย่างแบบเนย์แมน ตัวประมาณค่าแบบอัตราส่วนรวมที่พัฒนาขึ้นใหม่ทั้ง 3 ตัว มีประสิทธิภาพในระดับเดียวกันกับตัวประมาณค่าแบบอัตราส่วนรวมในปัจจุบัน เมื่อใช้วิธีการจัดสรรขนาดตัวอย่างแบบเท่ากัน

ในตัวอย่างที่ 3 น้ำหนักของชั้นภูมิ w_h , w^*_{yh} และ w^*_{xh} ในตารางที่ 4.8 มีความแตกต่างกัน ในกรณีของการจัดสรรขนาดตัวอย่างแบบเนย์แมน สัมประสิทธิ์สหสัมพันธ์ส่วนรวม ρ_c เท่ากับ 0.39 0.49 และ 0.44 และค่า $\frac{C_{\bar{x}_{st}}}{2C_{\bar{y}_{st}}}$ เท่ากับ 0.68, 0.61 และ 0.66 เมื่อใช้น้ำหนักของชั้นภูมิ w_h , w^*_{yh} และ w^*_{xh} ตามลำดับ ดังนั้น เงื่อนไขไม่เป็นไปตามอสมการ (3.2) ในกรณีของการจัดสรรขนาดตัวอย่างแบบเนย์แมน และเมื่อใช้การจัดสรรขนาดตัวอย่างแบบเท่ากัน เงื่อนไขก็ไม่เป็นไปตามอสมการ (3.2) เช่นกัน ดังนั้น ค่าเฉลี่ยตัวอย่างของชั้นภูมิก็จะใช้เป็นตัวประมาณค่าเฉลี่ยของประชากร $\hat{\bar{Y}}_{st\_wy}$ และ $\hat{\bar{Y}}_{st}$ ประสิทธิภาพสัมพัทธ์ของ $\hat{\bar{Y}}_{st\_wy}$ เมื่อเทียบกับตัวประมาณค่า $\hat{\bar{Y}}_{st}$ ที่ใช้การจัดสรรขนาดตัวอย่างแบบเนย์แมนมีค่าเท่ากับ 1.01 และเพิ่มขึ้นเป็น 1.05 เมื่อใช้วิธีการจัดสรรขนาดตัวอย่างแบบเท่ากัน แต่ประสิทธิภาพสัมพัทธ์ของ $\hat{\bar{Y}}_{st}$ เมื่อใช้การจัดสรรขนาดตัวอย่างแบบเท่ากันจะลดลงเหลือ 0.86 ตามที่แสดงไว้ในตารางที่ 4.9 ดังนั้น ตัวประมาณค่าตัวใหม่ $\hat{\bar{Y}}_{st\_wy}$ ในทั้ง 2 วิธีของการจัดสรรขนาดตัวอย่าง มีประสิทธิภาพสูงกว่าตัวประมาณแบบเดิม $\bar{y}_{st}$

ในตัวอย่างที่ 4 ค่าน้ำหนักของชั้นภูมิ w_h , w^*_{yh} และ w^*_{xh} มีค่าอยู่ในระดับใกล้เคียงกันเมื่อใช้การจัดสรรขนาดตัวอย่างแบบเนย์แมน แต่มีความแตกต่างจาก 4 วิธีของการจัดสรรขนาดตัวอย่างตามที่ได้นำเสนอไว้ในตารางที่ 4.8 เนื่องจากสัมประสิทธิ์สหสัมพัทธ์ระหว่างผลผลิตของพืชและจำนวนต้นพืชมีค่าสูงเท่ากับ 0.86 ค่าน้ำหนักของชั้นภูมิ w^*_{yh} และ w^*_{xh} มีค่าใกล้เคียงกันในทุกวิธีของการจัดสรรขนาดของตัวอย่างในชั้นภูมิ

ค่าสัมประสิทธิ์สหสัมพัทธ์ส่วนรวม ρ_c มีค่ามากกว่า $\frac{C_{\bar{x}_{st}}}{2C_{\bar{y}_{st}}}$ ในทุกกรณีของตัวอย่างที่ 4 ตัวประมาณค่าแบบอัตราส่วนรวมจึงนำมาใช้ในการประมาณค่าเฉลี่ยของประชากร ประสิทธิภาพสัมพัทธ์ในตารางที่ 4.9 ของตัวอย่างที่ 4 ทุกกรณีจะเปรียบเทียบกับประสิทธิภาพของ $\hat{\bar{Y}}_{CR}$ ที่ใช้การจัดสรรขนาดตัวอย่างแบบเนย์แมน จะเห็นได้ว่า ในกรณีที่ขนาดตัวอย่างในชั้นภูมิเป็นสัดส่วนกับจำนวนหน่วยประชากรในชั้นภูมิ ประสิทธิภาพสัมพัทธ์ของ $\hat{\bar{Y}}_{Rwy}$, $\hat{\bar{Y}}_{Rwx}$ และ $\hat{\bar{Y}}_{Rwyx}$ มีค่าระหว่าง 0.94-0.95 ต่ำกว่าประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าแบบอัตราส่วนรวม $\hat{\bar{Y}}_{CR}$, $\hat{\bar{Y}}_{KC}$ และ $\hat{\bar{Y}}_{SD}$ เล็กน้อย ซึ่งมีค่าอยู่ระหว่าง 0.96-0.98 ในกรณีที่ขนาดตัวอย่าง n_h เป็นสัดส่วนกับ N_hS_{yh} หรือขนาดตัวอย่างเป็นไปตามการจัดสรรแบบเนย์แมน ประสิทธิภาพสัมพัทธ์ของ $\hat{\bar{Y}}_{Rwy}$, $\hat{\bar{Y}}_{Rwx}$ และ $\hat{\bar{Y}}_{Rwyx}$ อยู่ระหว่าง 1.02-1.06 สูงกว่ากรณีของ

$\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ ซึ่งอยู่ระหว่าง 1.00-1.02 ในกรณีที่ขนาดตัวอย่าง n_h เป็นสัดส่วนกับ $N_h\sqrt{\bar{X}_h}$ ประสิทธิภาพสัมพัทธ์ของ $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{Rwx}$ และ $\hat{\hat{Y}}_{Rwyx}$ อยู่ระหว่าง 0.98-1.02 ต่ำกว่ากรณีของ $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ ซึ่งอยู่ระหว่าง 1.07-1.09 เช่นเดียวกันกับกรณีที่ขนาดตัวอย่าง n_h เป็นสัดส่วนกับ $N_h\bar{X}_h$ ประสิทธิภาพสัมพัทธ์ของ $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{Rwx}$ และ $\hat{\hat{Y}}_{Rwyx}$ อยู่ระหว่าง 1.00-1.07 ต่ำกว่ากรณี $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ ซึ่งอยู่ระหว่าง 1.13-1.15 แต่ในกรณีที่ขนาดตัวอย่างเท่ากันทุกชั้นภูมิ ประสิทธิภาพของ $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{Rwx}$ และ $\hat{\hat{Y}}_{Rwyx}$ อยู่ระหว่าง 1.46-1.60 สูงกว่ากรณี $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ มาก ซึ่งอยู่ระหว่าง 0.82-0.84 เท่านั้น

ในตัวอย่างที่ 4 $\hat{\hat{Y}}_{Rwx}$ มีประสิทธิภาพสูงสุดในบรรดาตัวประมาณค่าที่ได้พัฒนาขึ้นใหม่

4.4 สรุป

ตัวประมาณค่าที่ได้พัฒนาขึ้นในการชักตัวอย่างแบบชั้นภูมิ ซึ่งได้แก่ $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{Rwx}$, $\hat{\hat{Y}}_{Rwyx}$ และ $\hat{\hat{Y}}_{st\_wy}$ เป็นตัวประมาณค่าถ่วงน้ำหนักที่ได้พัฒนามาจากค่าสถิติของชั้นภูมิ ซึ่งแตกต่างจากตัวประมาณค่าที่มีอยู่เดิมที่ถ่วงน้ำหนักด้วยจำนวนหน่วยตัวอย่างในชั้นภูมิ N_h หารด้วยจำนวนหน่วยตัวอย่างในประชากร N

หากเงื่อนไขไม่เป็นไปตามอสมการ (3.2) ตัวประมาณค่าที่นำเสนอ $\hat{\hat{Y}}_{st\_wy}$ มีประสิทธิภาพสูงกว่าตัวประมาณค่า $\bar{y}_{st}$ ในการจัดสรรขนาดของตัวอย่างทั้ง 2 วิธี ดังในข้อมูลจากการจำลองและในข้อมูลจริงในตัวอย่างที่ 3

เมื่อใช้การจัดสรรขนาดตัวอย่างแบบเนย์แมน ผลการจำลองสรุปได้ว่า ประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนรวมที่ได้พัฒนาขึ้นไม่แตกต่างจากประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนรวมที่ใช้อยู่ในปัจจุบัน $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ แต่ตัวประมาณค่าแบบอัตราส่วนรวมที่ได้พัฒนาขึ้น มีประสิทธิภาพสูงกว่าตัวประมาณค่าปัจจุบัน $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ ในข้อมูลจริง ดังในตัวอย่างที่ 2 และ 4

เมื่อใช้การจัดสรรขนาดตัวอย่างเท่ากัน ประสิทธิภาพของ $\hat{\hat{Y}}_{Rwy}$, $\hat{\hat{Y}}_{Rwx}$ และ $\hat{\hat{Y}}_{Rwyx}$ สูงกว่าประสิทธิภาพของ $\hat{\hat{Y}}_{CR}$, $\hat{\hat{Y}}_{KC}$ และ $\hat{\hat{Y}}_{SD}$ มาก ทั้งในข้อมูลจากการจำลอง และข้อมูลจริงในตัวอย่างที่ 4 แต่ไม่พบความแตกต่างของประสิทธิภาพของตัวประมาณค่าแบบอัตราส่วนรวมที่ได้พัฒนาขึ้นกับตัวประมาณค่าแบบอัตราส่วนรวมที่ใช้อยู่ในปัจจุบันทั้ง 3 ตัว ในข้อมูลจริงในตัวอย่างที่ 2

อย่างไรก็ตาม การเปรียบเทียบประสิทธิภาพตามที่ได้เสนอมานั้น มิได้รวมความคลาดเคลื่อนในการประมาณค่าเฉลี่ยในชั้นภูมิ ซึ่งหากพิจารณาประเด็นนี้ อาจกล่าวได้ว่า ในกรณีที่ขนาดตัวอย่างเล็ก เช่น $n = 30$ ความคลาดเคลื่อนในการประมาณค่าเฉลี่ยในชั้นภูมิ เมื่อใช้การจัดสรรขนาดตัวอย่างแบบเนย์แมน จะสูงกว่าในกรณีใช้การจัดสรรขนาดตัวอย่างเท่ากัน เพราะในชั้นภูมิที่มีขนาดเล็ก มีตัวอย่างเพียง 2-3 ตัวอย่างเท่านั้นที่ใช้ในการประมาณค่า ดังในตารางที่ 4.1 ดังนั้น หากนำประเด็นความคลาดเคลื่อนที่เกิดจากการที่มีจำนวนตัวอย่างน้อยสำหรับประมาณค่า ข้อสรุปเกี่ยวกับประสิทธิภาพดังที่ได้กล่าวมาแล้ว ในกรณีการจัดสรรขนาดตัวอย่างแบบเนย์แมนอาจคลาดเคลื่อนและควรจะใช้ตัวประมาณค่าแบบอัตราส่วนรวม $\hat{\bar{Y}}_{Rwy}$, $\hat{\bar{Y}}_{Rwx}$ และ $\hat{\bar{Y}}_{Rwyx}$ มีประสิทธิภาพสูงกว่าตัวประมาณค่าแบบอัตราส่วนรวมที่ใช้อยู่ในปัจจุบันทั้ง 3 ตัว ทั้งในข้อมูลจากการจำลองและข้อมูลจริงในตัวอย่างที่ 2 และ 4

ตัวประมาณค่าที่ใช้น้ำหนักจากค่าสถิติในระดับชั้นภูมิ จึงเป็นทางเลือกอีกทางหนึ่งที่มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ของประชากร เพิ่มเติมจากตัวประมาณค่าที่ใช้น้ำหนักจากขนาดประชากรในชั้นภูมิ

บรรณานุกรม

Cingi, H (1994). *Sampling Theory*. Ankara, Turkey: Hacettepe University Press.

Cochran, W. G. (1961). Comparison of methods for determining stratum boundaries. *Bulletin of the International Statistical Institute*, 38:(2): 345-358

Cochran, W. G. (1977). *Sampling Techniques*. New York: John Wiley and Sons.

Dalenius, T. (1950). The problem of optimum stratification. *Skandinavisk Aktuarietidskrift* 33:203-213.

Dalenius, T and Hodges, J.L. (1959). Minimum variance stratification. *Journal of the American Statistical Association* 54(285):88-101.

Ekman, G. (1959). An approximation useful in univariate stratification. *Ann. Math. Stat*., 30:219-229.

Goodman, L.A. (1960). On the exact variance of products. *Journal of the American Statistical Association* 55(292):708-713.

Gunning, P. and Horgan, J.M. (2004). A new algorithm for the construction of stratum boundaries in skewed populations. *Survey Methodology* 2(30): 159-166.

Hedlin, D. (2000). A procedure for stratification by an extended Ekman rule. *Journal of Official Statistics*, 16: 15-29.

Henderson, H.V. and Searle, S.R. (1981). On deriving the inverse of a sum of matrices. *SIAM Review* 23(1): 53-60.

Kadilar, C. and Cingi, H. (2003). Ratio estimators in stratified random sampling. *Biometrical Journal* 45(2):218-225.

Kadilar, C. and Cingi, H. (2005). A new ratio estimator in stratified random sampling. *Communications in Statistics: Theory and Methods* 34: 597-602.

Lavellee, P. and Hidiroglou, M. (1988). On the stratification of the skewed populations. *Survey Methodology*, 14:33-43.

Neyman, J. (1934). On the two different aspects of the respective method: The method of stratified sampling and the method of purposive selection. *Jour. Roy. Stat. Soc*., 97: 558-606

Singh, R. and Mangat, N.S. (1996). *Elements of Survey Sampling*. London: Kluwer Academic Publishers.

Singh, H. and Vishwakarma, G.K. (2008). A family of estimators of population mean using auxiliary information in stratified sampling. *Communications in Statistics: Theory and Methods* 37: 1038-1050.

Sisodia, B. V. S. and Dwivedi, V. K. (1981). A modified ratio estimator using known coefficient of variation of auxiliary variable. *Journal of Indian Society of Agricultural Statistics* 33:13-18.

Upadhyaya, L. N. and Singh, H. P. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical Journal* 41(5):627-636.