บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดและทฤษฎี

2.1.1 แผนการสุ่มตัวอย่างแบบกลุ่ม (Cluster Sampling)

การสุ่มตัวอย่างแบบกลุ่ม หมายถึง การสุ่มตัวอย่างที่หน่วยตัวอย่างเกิดจากหน่วยที่ให้ข้อมูลหลายหน่วยรวมกัน มีลักษณะรวมกันเป็นกลุ่ม โดยแต่ละกลุ่มนั้นอาจจะแบ่งได้หลายชั้นหรือหลายระดับ และจำนวนชั้นก็สามารถแบ่งออกได้อีก การสุ่มตัวอย่างสามารถสุ่มได้แล้วแต่จำนวนชั้น เช่น การสุ่มตัวอย่างกลุ่มหนึ่งชั้น (One-Stage Cluster Sampling) การสุ่มตัวอย่างกลุ่มสองชั้น (Two-Stage Cluster Sampling) หรือ การสุ่มตัวอย่างกลุ่มสามชั้น (Three-Stage Cluster Sampling) เป็นต้น การสุ่มตัวอย่างกลุ่มที่มากกว่าหนึ่งชั้น เรียกรวมๆ ว่าการสุ่มตัวกลุ่มหลายชั้น (Multi-Stage Cluster Sampling) หลักการสุ่มตัวอย่างกลุ่มให้เป็นตัวประมาณที่ดีของประชากร คือ หน่วยต่าง ๆ ภายในกลุ่มควรจะมีความหลากหลาย แต่ระหว่างกลุ่มควรจะมีความคล้ายคลึงกันมาก ในความเป็นจริงหลักการนี้เป็นไปได้ยาก เช่น จัดกลุ่มครัวเรือนตามท้องถิ่น จะเห็นได้ว่าครัวเรือนในกลุ่มท้องถิ่นย่อมมีความคล้ายคลึงกัน ในขณะที่ครัวเรือนต่างกลุ่มกับมีความหลากหลาย โดยตรงข้ามกับหลักการที่ดีของการสุ่มตัวอย่างกลุ่ม ส่งผลให้ตัวประมาณของประชากรอาจไม่แม่นยำเท่าที่ควร แต่วิธีการสุ่มตัวอย่างกลุ่มนี้ จะช่วยให้กรอบตัวอย่างเป็นพวกเดียวกัน หรือมีความใกล้เคียงกัน หน่วยตัวอย่างจะกระจุกตัวอยู่บริเวณใกล้เคียงกันทำให้สามารถควบคุมการจัดเก็บข้อมูลได้ เป็นการประหยัดค่าใช้จ่ายและสะดวกในการควบคุมงานสนาม เพราะผู้ควบคุมงานสนามก็จะควบคุมพนักงานสำรวจเฉพาะพื้นที่ที่กำลังทำการสำรวจ และง่ายในการติดต่อสื่อสาร

สำหรับวิธีการสุ่มแบบกลุ่ม ทำได้หลายแบบสามารถนำแบบการสุ่มตัวอย่างอื่น ๆ มาผสมผสานกัน เช่น การสุ่มตัวอย่างกลุ่มสองชั้นแบบมีระบบ (Stratified Two-Stage Cluster Sampling) หมายถึง การสุ่มตัวอย่างกลุ่มสองชั้นเมื่อแบ่งประชากรเป็นชั้นภูมิ หรือ การสุ่มตัวอย่างกลุ่มสามชั้นอย่างง่าย (Simple Three-Stage Cluster Sampling) หมายถึง การสุ่มตัวอย่างกลุ่มสามชั้นโดยสุ่มหน่วยตัวอย่างแบบง่าย

ในการสุ่มตัวอย่างแบบกลุ่ม ไม่ว่าจะเป็นการสุ่มตัวอย่างกลุ่มหนึ่งชั้น (One-Stage Cluster Sampling) การสุ่มตัวอย่างกลุ่มสองชั้น (Two-Stage Cluster Sampling) หรือ การสุ่มตัวอย่างกลุ่มสามชั้น (Three-Stage Cluster Sampling) โดยสามารถแยกพิจารณาได้เป็น 2 กรณี คือการสุ่มตัว

อย่างโดยใช้ความน่าจะเป็นของการสุ่มเท่า ๆ กัน (Equal Probability Sampling) และ การสุ่มตัวอย่างโดยใช้ความน่าจะเป็นของการสุ่มไม่เท่ากัน (Unequal Probability Sampling)

2.1.2 **แผนการสุ่มตัวอย่างแบบกลุ่มปรับ** (Adaptive Cluster Sampling)

แผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับ (Adaptive Cluster Sampling) ได้ถูกพัฒนามาจากการสุ่มตัวอย่างแบบกลุ่ม ตัวอย่างเช่น ต้องการสำรวจสัตว์ชนิดหนึ่งในพื้นที่หนึ่งขนาด 10×10 ตารางหน่วย โดยที่มีการสุ่มตัวอย่างเป็นแบบกลุ่มขนาด 5 หน่วย หากทั้ง 5 หน่วยพบสัตว์ที่สนใจเพียงหน่วยเดียว จะทำให้เกิดปัญหา โดยข้อมูลที่ได้ไม่เพียงพอที่จะคำนวณ เกิดความเอนเอียงมาก ดังนั้นเพื่อแก้ปัญหาดังกล่าว จึงนำแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับเข้ามาทำการสุ่มตัวอย่างแทน โดยในตอนเริ่มแรกจะนำหลักการสุ่มตัวอย่างแบบกลุ่มเข้ามาใช้ โดยจากตัวอย่างข้างต้นจะสุ่มตัวอย่างเป็นแบบกลุ่มขนาด 5 หน่วย สมมติให้พบสัตว์เพียงหน่วยเดียว จากนั้นหน่วยที่พบสัตว์ที่ต้องการก็ใช้หลักแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับ โดยจะทำการขยายหน่วยถัดไปทางซ้าย ขวา บน และล่าง หากพบสัตว์ที่สนใจก็จะเพิ่มอย่างนี้ไปเรื่อย ๆ จนไม่พบสัตว์ที่สนใจ ข้อมูลที่พบก็เพียงพอในการคำนวณ และตัวประมาณที่ได้จากแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับน่าจะมีประสิทธิภาพดีกว่าจากการสุ่มตัวอย่างแบบกลุ่ม

ในปกติการสุ่มหน่วยตัวอย่างจะไม่สามารถทำนายล่วงหน้าได้ก่อนการสำรวจข้อมูล แต่สำหรับแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับนี้ มีหลักการภายใต้แนวความคิดที่ว่า ลักษณะของข้อมูลที่เก็บรวบรวมได้น่าจะมีส่วนช่วยทำให้ตัวประมาณมีคุณภาพดีขึ้น โดยการปรับกระบวนการสุ่มตัวอย่างให้สามารถนำลักษณะของข้อมูลที่เก็บได้มากำหนดหน่วยตัวอย่างที่เก็บรวบรวมเพิ่มขึ้น ทั้งในเชิงจำนวนและลักษณะของหน่วยที่จะสุ่มเป็นตัวอย่างเพิ่มเติม เช่น ในกรณีของการประมาณจำนวนสัตว์และพบสัตว์ที่อยู่อาศัยในพื้นที่หนึ่งในลักษณะที่ค่อนข้างเกาะกลุ่มกัน หากสุ่มตัวอย่างและพบสัตว์ประเภทนี้อยู่ในหน่วยตัวอย่างใด โอกาสที่จะพบในหน่วยใกล้เคียงหรือพื้นที่ข้างเคียงจะมีมากขึ้น จึงควรสุ่มหน่วยตัวอย่างในบริเวณใกล้เคียงเพิ่มเติม[1] รูปแบบของการอยู่เป็นกลุ่มเช่นนี้ยังพบในสิ่งอื่น ๆ เช่น พืช แร่ธาตุ แหล่งของฟอสซิล และรูปแบบนี้ยังพบในการศึกษาเกี่ยวกับการระบาดของโรคด้วย เมื่อใดก็ตามที่

---

[1] สุชาดา กีระนันท์,ทฤษฎีและวิธีการสำรวจตัวอย่าง,พิมพ์ครั้งที่ 2 (กรุงเทพมหานคร : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย,2542),หน้า 352.

พบสิ่งที่ต้องการศึกษาโดยบังเอิญ การเพิ่มตัวอย่างในหน่วยที่ใกล้เคียง จะมีอัตราการพบสิ่งที่ศึกษาในอัตราที่ค่อนข้างสูง

แผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับสามารถเลือกทำการสุ่มตัวอย่างข้างต้น (Initial Sample) ได้โดยอาศัยกระบวนการสุ่มตัวอย่างด้วยความน่าจะเป็น เช่น วิธีการสุ่มตัวอย่างแบบง่าย , วิธีการสุ่มตัวอย่างแบบมีชั้นภูมิ และวิธีการสุ่มตัวอย่างแบบมีระบบ ซึ่งจะเลือกสุ่มตัวอย่างขั้นต้นด้วยวิธีใดจะขึ้นอยู่กับหลายปัจจัย ไม่ว่าจะเป็นลักษณะของประชากรที่ศึกษา ขนาดตัวอย่าง งบประมาณ รวมทั้งระยะเวลา เป็นต้น

2.1.3 แผนการสุ่มตัวอย่างแบบกลุ่มปรับ (Adaptive Cluster Sampling) ที่ทำการสุ่มตัวอย่างขั้นต้นอย่างง่าย[2]

แผนการสุ่มตัวอย่างแบบกลุ่มปรับ (Adaptive Cluster Sampling) ที่ทำการสุ่มตัวอย่างขั้นต้นอย่างง่าย หมายถึง แผนแบบการสุ่มตัวอย่างแบบที่ปรับกระบวนการสุ่มตัวอย่างให้สามารถนำลักษณะของข้อมูลที่เก็บได้มากำหนดหน่วยตัวอย่าง โดยการสุ่มตัวอย่างเบื้องต้นจะอาศัยหลักการสุ่มตัวอย่างแบบง่าย เช่น กำหนดพื้นที่เบื้องต้นซึ่งอาศัยหลักการสุ่มตัวอย่างแบบง่ายในการสำรวจสัตว์ หากพื้นที่เบื้องต้นใดพบสัตว์ที่เป็นไปตามเกณฑ์ก็จะเพิ่มพื้นที่บริเวณใกล้เคียงเป็นตัวอย่างเพิ่มเติม แต่ถ้าพื้นที่ใดไม่พบสัตว์ที่เป็นไปตามเกณฑ์พื้นที่นั้นก็ไม่ต้องสนใจ

ในการสุ่มหน่วยตัวอย่างที่เพิ่มขึ้นนั้น หน่วยตัวอย่างขั้นต้นจะต้องเป็นไปตามเงื่อนไขที่เรากำหนด เช่น พื้นที่หน่วยตัวอย่างขั้นต้นพบสิ่งที่เราสนใจ ทั้งนี้หน่วยตัวอย่างขั้นต้นจะประกอบไปด้วยกลุ่ม (Cluster) ซึ่งหมายถึง Network รวมกับ หน่วยขอบ (Edge) โดยที่ Network จะเป็นหน่วยย่อยที่เป็นไปตามเงื่อนไขที่ต้องการ ส่วนหน่วยขอบจะไม่ใช้ในการคำนวณ แต่เป็นตัวบอกขอบเขตสิ้นสุดที่ไม่มีการขยายหน่วยถัดไป บางครั้งการขยายการสุ่มหน่วยตัวอย่างในแต่ละกลุ่ม เมื่อทำการขยายบางหน่วยตัวอย่าง อาจทำให้เกิดการรวมกันเป็น Network เดียวกัน ในขณะที่หน่วยตัวอย่างขั้นต้นที่ไม่เป็นไปตามเงื่อนไขที่เรากำหนดจะไม่มีการขยายหน่วยตัวอย่างถัดไป โดยจะหมายถึง Network 1 หน่วย

ความคิดพื้นฐานของแผนแบบสามารถอธิบายได้ ดังภาพที่ 2 เป็นการประมาณจำนวนค่าเฉลี่ยของจุดที่สุ่ม โดยจุดดังกล่าวอาจเป็นบริเวณของสัตว์หรือแหล่งแร่ ได้แบ่งบริเวณที่ศึกษาออกเป็น 400

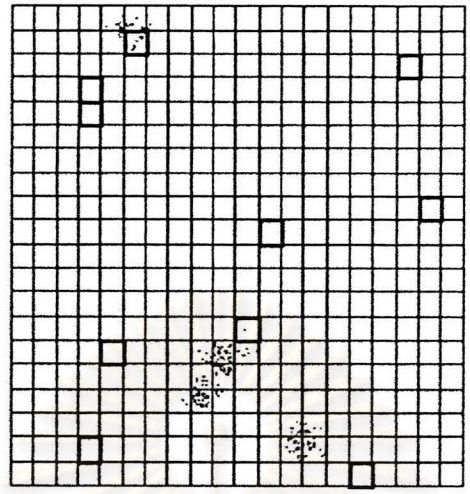
---

[2] Thompson,S.K(1990), "Adaptive Cluster Sampling," Jour.Amer. Stat. Assoc 85, No.412 : 1050-1059.

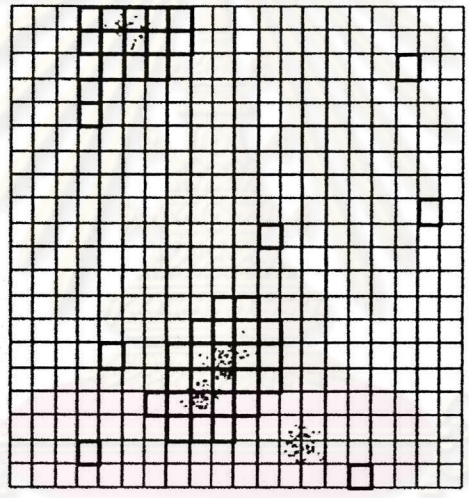
หน่วยตัวอย่าง ทำการสุ่มตัวอย่างขั้นต้นขนาด 10 หน่วย แสดงให้เห็นในภาพที่ 2 (a) เมื่อสังเกตพบสิ่งที่สนใจจำนวนหนึ่งหรือมากกว่าในหน่วยที่สุ่ม หน่วยใกล้เคียงถัดไปทางซ้าย ขวา บน ล่าง จะถูกเพิ่มเป็นตัวอย่าง เมื่อกระบวนการนี้เสร็จสมบูรณ์ ตัวอย่างจะประกอบด้วย 45 หน่วย แสดงให้เห็นในภาพที่ 2 (b)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(a)



(b)



ภาพที่ 2 : ภาพแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับ (Adaptive Cluster Sampling) ในการประมาณจำนวนจุดของค่าสังเกตในบริเวณที่ศึกษา 400 หน่วย

(a) การสุ่มตัวอย่างขั้นต้น 10 หน่วย

(b) การเพิ่มหน่วยที่ใกล้เคียงเนื่องจากการสุ่มพบ 1 หรือ มากกว่า 1 ค่าสังเกตของประชากรในหน่วยที่สุ่ม โดยจะทำการขยายหน่วยถัดไปทางซ้าย ขวา บน และล่าง แต่ถ้าไม่พบสิ่งที่สนใจก็ไม่ต้องเพิ่มหน่วยถัดไป

ที่มา : Thompson,S.K.(1990). Adaptive Cluster Sampling. Jour. Amer. Stat. Assoc., 85, 1051.

ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz –Thompson

ตัวประมาณค่าเฉลี่ยแบบง่าย ถึงแม้จะไม่เอนเอียงสำหรับค่าเฉลี่ยของประชากรเมื่อใช้กับแผนแบบการสุ่มตัวอย่างแบบที่ยังไม่ได้ปรับ แต่ถ้าใช้กับแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับ จะเกิดความเอนเอียง ดังนั้น Thompson (1990) ได้เสนอตัวประมาณค่าเฉลี่ยที่ไม่เอนเอียงสำหรับแผนแบบนี้ โดยการสร้างตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-Thompson

ตัวประมาณที่ไม่เอนเอียงของค่าเฉลี่ย คือ

$$\bar{y}_{SRS1} = \frac{1}{N}\sum_{k=1}^{\nu}\frac{y_k}{\alpha_k}$$

เมื่อ $\alpha_k = 1 - \binom{N-x_k}{n} \Big/ \binom{N}{n}$

โดย $\bar{y}_{SRS1}$ เป็นตัวประมาณที่ไม่เอนเอียงของค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-Thompson เมื่อตัวอย่างขั้นต้นสุ่มแบบง่าย

α_k เป็นความน่าจะเป็นที่หน่วยที่ k รวมอยู่ในตัวอย่างที่เลือกมาโดยวิธี Simple Random Sampling

y_k เป็นจำนวนหน่วยตัวอย่างของสิ่งที่เราสนใจ ในหน่วยที่ k

N เป็นจำนวนหน่วยของกลุ่มประชากร

n เป็นจำนวนหน่วยของกลุ่มตัวอย่าง

x_k เป็นจำนวน network ที่พบของสิ่งที่เราสนใจ ในหน่วยที่ k

ν เป็นจำนวนหน่วยของกลุ่มตัวอย่างหลังจากการสุ่มตัวอย่างแบบกลุ่มปรับ

ค่าประมาณของความแปรปรวนที่ไม่เอนเอียง คือ

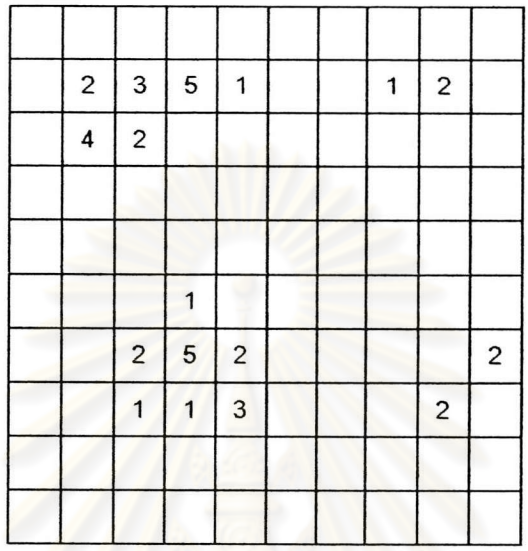
$$\hat{var}(\bar{y}_{SRS1}) = \frac{1}{N^2}\sum_{k=1}^{\nu}\sum_{l=1}^{\nu}\frac{y_k y_l}{\alpha_{kl}}\left(\frac{\alpha_{kl}}{\alpha_k\alpha_l}-1\right)$$

โดยที่ $\alpha_{kl} = 1 - \left\{\binom{N-x_k}{n} + \binom{N-x_l}{n} - \binom{N-x_k-x_l}{n}\right\} \Big/ \binom{N}{n}$

ให้ α_{kl} เป็นความน่าจะเป็นที่หน่วยที่ k และ l รวมอยู่ในตัวอย่าง

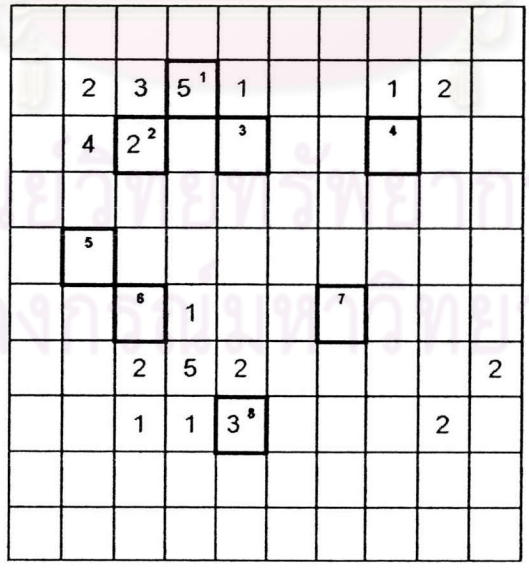
ตัวอย่างแผนการสุ่มตัวอย่างแบบกลุ่มปรับ (Adaptive Cluster Sampling) **ที่ทำการสุ่มตัวอย่างขั้นต้นอย่างง่าย**

ป่าแห่งหนึ่งมีพื้นที่ 100 ตร.หน่วย โดยมีจำนวนกวาง แสดงดังภาพที่ 3



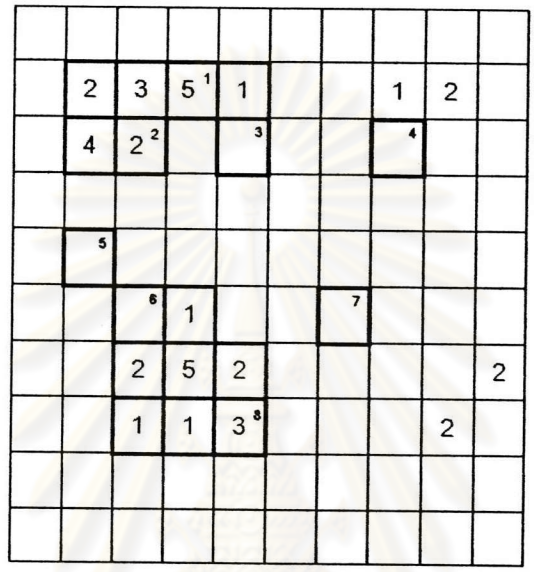
ภาพที่ 3 : ภาพแสดงที่ตั้งและจำนวนของสิ่งที่เราสนใจ

หากเราอาศัยหลักการของ Adaptive Cluster Sampling ซึ่งให้มีการสุ่มตัวอย่างขั้นต้นขนาด 8 หน่วย (n=8) โดยสุ่มพื้นที่ได้ตรงตำแหน่งดังภาพที่ 4



ภาพที่ 4 : ภาพแสดงการสุ่มตัวอย่างขั้นต้นขนาด 8 หน่วย (n=8)

จากภาพที่ 4 จะเห็นได้ว่า เมื่อเราสุ่มตัวอย่างขั้นต้นขนาด 8 หน่วย (n=8) ซึ่งพบว่ามีตัวอย่างขั้นต้น หรือพื้นที่จำนวน 3 หน่วยที่พบกวาง ได้แก่ พื้นที่ที่ 1,2 และ 8 ดังนั้นจึงขยายพื้นที่ทั้ง 3 หน่วยเพิ่มเติม โดยขยายพื้นที่ในลักษณะ ขึ้น-ลง-ซ้าย-ขวา และขยายไปเรื่อย ๆ จนจะไม่พบกวาง โดยสามารถอธิบายได้ดังภาพที่ 5



ภาพที่ 5 : ภาพแสดงการขยายพื้นที่ในลักษณะ ขึ้น-ลง-ซ้าย-ขวา และขยายไปเรื่อย ๆ จนจะไม่พบสิ่งที่เราสนใจ

จากภาพที่ 5 สามารถแทนในสูตรได้ ดังนี้

N = 100

n = 8

ν = 7 เนื่องจากหน่วยตัวอย่างขั้นต้นหรือพื้นที่ที่ 1 กับ 2 รวมเป็น Network เดียวกัน โดยจะไม่รวมพื้นที่ที่ 3 เข้าเป็น Network เดียวกันกับพื้นที่ 1 กับ 2 เพราะ พื้นที่ที่ 3 ไม่มีกวางอยู่

การคำนวณหาค่า $\bar{y}_{srs1}$ **ซึ่งเป็นตัวประมาณที่ไม่เอนเอียง** $\bar{Y}_{srs1}$

จาก $\bar{y}_{SRS\,1} = \frac{1}{N}\sum_{k=1}^{\nu}\frac{y_k}{\alpha_k}$

เมื่อ $\alpha_k = 1 - \binom{N-x_k}{n} \Big/ \binom{N}{n}$

โดยที่ $y_{1\&2} = 17$

$y_3 = y_4 = y_5 = y_6 = y_7 = 0$

$y_8 = 15$

$x_{1\&2} = 6$

$x_3 = x_4 = x_5 = x_6 = x_7 = 1$

$x_8 = 7$

$$\alpha_{1\&2} = 1 - \binom{N-x_{1\&2}}{n} \Big/ \binom{N}{n} = 1 - \binom{100-6}{8} \Big/ \binom{100}{8} = 0.4018$$

$$\alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = 1 - \binom{N-x_{3=\ldots=7}}{n} \Big/ \binom{N}{n}$$

$$= 1 - \binom{100-1}{8} \Big/ \binom{100}{8} = 0.08$$

$$\alpha_8 = 1 - \binom{N-x_8}{n} \Big/ \binom{N}{n} = \binom{100-7}{8} \Big/ \binom{100}{8} = 0.4527$$

$$\therefore \quad \bar{y}_{SRS\,1} = \frac{1}{100}\sum_{k=1}^{7}\frac{y_k}{\alpha_k} = \frac{1}{100}\left[\frac{17}{0.4018} + (5)\frac{0}{0.08} + \frac{15}{0.4527}\right] = 0.7544$$

การคำนวณหาค่า $\hat{var}(\bar{y}_{SRS1})$ **ซึ่งเป็นตัวประมาณที่ไม่เอนเอียงของ** $Var(\bar{y}_{SRS1})$

จาก $$\hat{var}(\bar{y}_{SRS1}) = \frac{1}{N^2}\sum_{k=1}^{v}\sum_{l=1}^{v}\frac{y_k y_l}{\alpha_{kl}}\left(\frac{\alpha_{kl}}{\alpha_k\alpha_l} - 1\right)$$

$$= \frac{1}{N^2}\left[\sum_{k=1}^{v}\frac{(y_k)^2}{\alpha_k}\left(\frac{1}{\alpha_k} - 1\right) + 2\sum_{k=1}^{v-1}\sum_{l=k+1}^{v}\frac{y_k y_l}{\alpha_{kl}}\left(\frac{\alpha_{kl}}{\alpha_k\alpha_l} - 1\right)\right]$$

$$\alpha_{kl} = 1 - \left\{\binom{N-x_k}{n} + \binom{N-x_l}{n} - \binom{N-x_k-x_l}{n}\right\} \Big/ \binom{N}{n}$$

ข้อสังเกตุ หากหน่วยตัวอย่างขั้นต้นใดไม่พบกับสิ่งที่เราสนใจ หรือ $y_k=0$ เราไม่จำเป็นต้องคำนวณค่าของ α_k หรือ α_{kl} นั้นๆ ตัวอย่างเช่น หาก $y_2=0$ ก็ทำให้

$\bar{y}_{SRS\,1} = \frac{1}{N}\sum_{k=1}^{2}\left(\frac{y_1}{\alpha_1} + \left(\frac{0}{\alpha_2} = 0\right)\right)$ หรือ จาก

$$\hat{var}(\bar{y}_{SRS1}) = \frac{1}{N^2}\left[\sum_{k=1}^{v}\frac{(y_k)^2}{\alpha_k}\left(\frac{1}{\alpha_k}-1\right)+2\sum_{k=1}^{1}\sum_{l=k+1}^{2}\frac{(y_k.0=0)}{\alpha_{kl}}\left(\frac{\alpha_{kl}}{\alpha_k\alpha_l}-1\right)\right]$$

ดังนั้น $\alpha_{(1\&2)(8)} = 1-\left\{\binom{100-6}{8}+\binom{100-7}{8}-\binom{100-6-7}{8}\right\}/\binom{100}{8} = 0.1685$

$$\therefore \quad \hat{var}(\bar{y}_{SRS1}) = \frac{1}{100^2}\left[\begin{array}{l}\left(\frac{(y_{1\&2})^2}{\alpha_{1\&2}}\left(\frac{1}{\alpha_{1\&2}}-1\right)+5.\frac{0}{\alpha_8}\left(\frac{1}{\alpha_8}-1\right)+\frac{(y_8)^2}{\alpha_8}\left(\frac{1}{\alpha_8}-1\right)\right)\\ +2\left(\frac{y_{1\&2}y_8}{\alpha_{(1\&2)(8)}}\left(\frac{\alpha_{(1\&2)(8)}}{\alpha_{1\&2}\alpha_8}-1\right)+0\right)\end{array}\right]$$

$$\hat{var}(\bar{y}_{SRS1}) = \frac{1}{100^2}\left[\begin{array}{l}\left(\frac{17^2}{0.4018}\left(\frac{1}{0.4018}-1\right)+\frac{15^2}{0.4527}\left(\frac{1}{0.4527}-1\right)\right)\\ +2\left(\frac{(17)(15)}{0.1685}\left(\frac{0.1685}{(0.4527)(0.4018)}-1\right)\right)\end{array}\right]$$

$$= 0.1448$$

2.1.4 **แผนแบบการสุ่มตัวอย่างแบบง่ายแบบที่ไม่ปรับ** (Simple Random Sampling)

เป็นวิธีการที่ง่ายที่สุดของการสุ่มตัวอย่างด้วยความน่าจะเป็น โดยโอกาสของทุกหน่วยในประชากรมีโอกาสถูกสุ่มเท่า ๆ กัน เช่น ถ้าสุ่มตัวอย่าง n จากประชากรขนาด N แบบไม่ใส่คืน จำนวนตัวอย่างที่เป็นไปได้ทั้งหมดจากการสุ่มตัวอย่าง คือ ${}^{N}c_{n}$ ดังนั้นโอกาสที่แต่ละชุดตัวอย่างจะมีโอกาสถูกสุ่มมาด้วยโอกาสเท่ากับ 1 / ${}^{N}c_{n}$ ซึ่งในการสุ่มตัวอย่างแบบง่ายนี้จะต้องทราบจำนวนประชากรขนาด N และลักษณะประชากรขนาด N จะต้องมีลักษณะคล้ายกันมากที่สุด

ตัวประมาณที่ไม่เอนเอียงของค่าเฉลี่ย คือ

$$\bar{y}_{SRS2} = \sum_{i=1}^{n}\frac{y_i}{n}$$

เมื่อ $\bar{y}_{SRS2}$ เป็นตัวประมาณที่ไม่เอนเอียงเมื่อเป็นการสุ่มตัวอย่างสุ่มแบบง่าย

ความแปรปรวนของตัวประมาณเมื่อสุ่มตัวอย่างแบบไม่คืนที่ คือ

$$Var(\bar{y}_{SRS2}) = \frac{S^2}{n}(1-\frac{n}{N})$$

ในทางปฏิบัติ $\text{var}(\bar{y}_{SRS2})$ มักเป็นตัวเลขที่ไม่ทราบเพราะติดค่า $S^2 = \sum_{i=1}^{N} \frac{(Y_i - \bar{Y})^2}{N-1}$ เป็นความแปรปรวนของประชากร ดังนั้นเมื่อทำการสำรวจ และได้ข้อมูลจากหน่วยที่ตกเป็นตัวอย่างในการสำรวจ n หน่วย ผู้สำรวจต้องอาศัยข้อมูลที่ได้จากตัวอย่างขนาด n เพื่อประมาณค่า $\hat{\text{var}}(\bar{y}_{SRS2})$ โดยการประมาณค่า S^2 ด้วย s^2 โดยที่ $s^2 = \sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n-1}$ เป็นตัวประมาณค่าที่ไม่เอนเอียงของ S^2

ดังนั้น $\hat{\text{var}}(\bar{y}_{SRS2}) = \frac{s^2}{n}\left(1 - \frac{n}{N}\right)$

โดยที่ y_i คือ ข้อมูลที่ได้จากหน่วยที่ i ในตัวอย่างขนาด n

$\bar{y}$ คือ ค่าเฉลี่ยตัวอย่างของตัวอย่างขนาด n ที่สุ่มมาจากประชากรขนาด N โดยแผนการสุ่มตัวอย่างสุ่มอย่างง่าย

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

ศิริประภา มโนมัธย์ (2539) ทำการศึกษาเรื่อง ประสิทธิภาพของแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับเมื่อตัวอย่างข้างต้นใช้วิธีการสุ่มอย่างง่าย แบบมีชั้นภูมิและแบบมีระบบ ภายใต้แบบจำลองประชากรที่กำหนด โดยใช้ร้อยละของอัตราส่วนความแปรปรวนของค่าเฉลี่ย (%eff) ระหว่างการสุ่มตัวอย่างขั้นต้นแบบมีชั้นภูมิและแบบมีระบบเทียบกับการสุ่มตัวอย่างขั้นต้นแบบง่ายของ S.K.Thompson (1990,1991a,1991b) สำหรับตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz และ Horvitz-Thompson การดำเนินงานวิจัยประกอบด้วยการจำลองแบบประชากรที่มีลักษณะหายากสร้างขึ้นทั้งหมด 5 ประชากร แต่ละกรณีเป็นประชากรที่มีลักษณะต่างกันโดยใช้กระบวนการพัวซองคลัสเตอร์ (Poisson Cluster Process) ขนาดตัวอย่างที่ใช้ในแต่ละวิธี คือ 4 8 16 32 และ 64 หน่วย ผลการวิจัยพบว่า

ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz เมื่อตัวอย่างมีขนาด 4 8 16 32 และ 64 วิธีการสุ่มตัวอย่างขั้นต้นแบบมีระบบมีประสิทธิภาพในแง่ของความแม่นยำดีที่สุด โดย %eff เท่ากับ 50.154 , 50.156 , 50.158 , 50.166 และ 50.070 ตามลำดับ รองลงมา คือ การสุ่มตัวอย่างขั้นต้นแบบมีชั้นภูมิ โดย %eff เท่ากับ 77.936 , 77.934 , 77.934 , 79.142 และ 79.082 ตามลำดับ ซึ่งพบว่า เมื่อตัวอย่างมีขนาดใหญ่ขึ้นแนวโน้มตัวประมาณค่าเฉลี่ยจะมีประสิทธิภาพในแง่ของความแม่นยำสูงขึ้นทั้งการสุ่มตัวอย่างขั้นต้นแบบมีระบบ และแบบมีชั้นภูมิ

ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-Thompson พบว่าเมื่อตัวอย่างมีขนาด 4 8 16 หน่วย วิธีการสุ่มตัวอย่างขั้นต้นแบบมีระบบมีค่า %eff เท่ากับ 63.364 , 56.024 และ 45.062 ตามลำดับ รองลงมา คือการสุ่มขั้นต้นแบบมีชั้นภูมิ โดย %eff เท่ากับ 84.334 , 77.930 และ 68.386 ตามลำดับชี้ให้เห็นประสิทธิภาพเพิ่มขึ้นตามขนาดตัวอย่างที่เพิ่มขึ้นด้วย ในขณะที่หากตัวอย่างใหญ่ขึ้นเป็น 32 และ 64 หน่วย ประสิทธิภาพของการสุ่มขั้นต้นแบบมีระบบและแบบมีชั้นภูมิกับลดลงและมีประสิทธิภาพใกล้เคียงกัน

เมื่อเปรียบเทียบความแปรปรวนของตัวประมาณค่าเฉลี่ยของแผนแบบกลุ่มปรับ เมื่อการสุ่มตัวอย่างขั้นต้นเป็นแบบง่าย แบบมีชั้นภูมิ และแบบมีระบบ พบว่า ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-Thompson จะให้ค่าความแปรปรวนน้อยกว่า ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz มีเพียงกรณีเดียว เมื่อตัวอย่างมีขนาดเล็ก (n=4) การสุ่มตัวอย่างขั้นต้นเป็นแบบมีระบบ ค่าความแปรปรวนของค่าเฉลี่ยที่ดัดแปลงมาจากทั้ง 2 วิธีจะมีค่าเท่ากัน แสดงให้เห็นว่า ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-Thompson มีประสิทธิภาพมากกว่า ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz

สรุปได้ว่า แผนแบบกลุ่มปรับ สามารถใช้ได้กับประชากรที่มีลักษณะหายากและอยู่รวมกันเป็นกลุ่ม ความมีประสิทธิภาพของแผนแบบจะขึ้นอยู่กับลักษณะการกระจายของประชากรที่ทำการศึกษา วิธีการสุ่มตัวอย่างขั้นต้น ตัวประมาณที่ใช้ และขนาดตัวอย่างที่ทำการสุ่มในขั้นต้น และพบว่าประสิทธิภาพของแผนแบบกลุ่มปรับที่มีการสุ่มตัวอย่างขั้นต้นแบบมีระบบใช้ได้ดีกว่าแบบอื่น ๆ รองลงมา คือ แบบมีชั้นภูมิและแบบง่ายตามลำดับ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย