ดังนั้นจึงกำหนดสัญลักษณ์ให้สอดคล้องกับสถานการณ์ข้างต้นได้ดังนี้ อย่างไรก็ตาม ก่อนที่จะติดตามศึกษาต่อไปขอให้ข้อแนะนำไว้ว่ากรณีนี้เป็นเพียงกรณีเฉพาะวิธีปฏิบัติที่ง่ายที่สุดก็คือเพียงแต่แทนที่ $\bar{N}$ และ $\bar{n}$ ลงใน N_i และ n_i ของสูตรทั่วไปเท่านั้นแต่ที่จะพัฒนาต่อไปนี้เป็นการพัฒนาสูตรไปสู่รูปที่เห็นว่าง่ายที่สุดและคล้ายสูตรของ SRS มากที่สุด หากนักศึกษาเห็นว่าเป็นภาระหนักเกินไปในการจดจำก็อาจใช้วิธีแทนที่ดังกล่าวข้างต้นก็ได้เพราะผลลัพธ์ตรงกัน

ให้ $N_i = \bar{N} = N/M$ [1]

และ $n_i = \bar{n} = n/m$

ดังนั้นจาก $\hat{T} = \frac{M}{m} \sum_i^m \frac{N_i}{n_i} \sum_j^{n_i} x_{ij}$

เมื่อ $N_i = \bar{N}$ และ $n_i = \bar{n}$ ดังนั้น

1. $\hat{T} = \frac{M}{m} \sum_i^m \frac{\bar{N}}{\bar{n}} \sum_j^{\bar{n}} x_{ij} = M\hat{\bar{T}}$

และจาก $\hat{\bar{\bar{X}}} = \hat{T}/N$

เมื่อ $N_i = \bar{N} = N/M$ ดังนั้น $N = M\bar{N}$

2. $\hat{\bar{\bar{X}}} = \frac{1}{Nm} M \sum_i^m \frac{\bar{N}}{\bar{n}} \sum_j^{\bar{n}} x_{ij} = \frac{M\bar{N}}{Nm} \sum_i^m \frac{1}{\bar{n}} \sum_j^{\bar{n}} x_{ij} = \frac{1}{m\bar{n}} \sum_i^m \sum_j^{\bar{n}} x_{ij}$

ในทำนองเดียวกัน

3. $\hat{P} = \frac{1}{m\bar{n}} \sum_i^m \sum_j^{\bar{n}} x_{ij}$ เมื่อ $x_{ij} = 0, 1$

---

1. มิได้หมายความว่าต้องทราบค่าของ N ก่อนจึงจะหาค่า $\bar{N}$ ได้แต่หมายความว่าค่าขนาด Cluster ที่ถือว่าเท่ากันหรือใกล้เคียงกันเท่ากับ $\bar{N}$ นั้นมีค่าเท่ากับ $\frac{N}{M}$ ด้วยเหตุนี้เมื่อเราสามารถตัดสินใจหรือทราบว่าขนาด Cluster ต่าง ๆ คงที่เท่ากันเท่ากับค่าคงที่ค่าหนึ่งคือ $\bar{N}$ ดังนั้นเราจึงสามารถทราบได้ทันทีว่าขนาดประชากรก็คือ $M\bar{N}$ ในทำนองเดียวกัน $n = m\bar{n}$

โดยอาศัยความรู้ในตอน 5.2.2

$$V(\hat{T}) = M^2 \frac{M-m}{M} \frac{S_b^2}{m} + \frac{M}{m} \sum_i^M \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

เมื่อ $N_i = \bar{N}$ และ $n_i = \bar{n}$

$$=> V(\hat{T}) = M^2 \frac{M-m}{M} \frac{S_b^2}{m} + \frac{M}{m} \sum_i^M \frac{\bar{N} - \bar{n}}{\bar{N}} \frac{\bar{N}^2 S_i^2}{\bar{n}}$$

4. $$\hat{V}(\hat{T}) = M^2 \frac{M-m}{M} \frac{s_b^2}{m} + \frac{M}{m} \sum_i^m \frac{\bar{N} - \bar{n}}{\bar{N}} \cdot \bar{N}^2 \frac{s_i^2}{\bar{n}}$$

$$V(\hat{\bar{\bar{X}}}) = V(\hat{T}/N) = \frac{1}{N^2} V(\hat{T})$$

$$= \frac{M^2}{\bar{N}^2} \cdot \frac{M-m}{M} \frac{S_b^2}{m} + \frac{1}{N^2} \frac{M}{m} \cdot \sum_i^M \frac{\bar{N} - \bar{n}}{\bar{N}} \frac{\bar{N}^2 S_i^2}{\bar{n}}$$

พิจารณาเทอมแรกจะพบว่า

$$\frac{M^2}{\bar{N}^2} \cdot \frac{M-m}{M} \cdot \frac{S_b^2}{m} = \frac{1}{(N/M)^2} \cdot \frac{M-m}{M} \cdot \frac{S_b^2}{m}$$

$$= \frac{1}{\bar{\bar{N}}^2} \frac{M-m}{M} \frac{S_b^2}{m}$$

$$= \frac{M-m}{M} \frac{1}{m} \frac{S_b^2}{\bar{\bar{N}}^2}$$

$$= \frac{M-m}{M} \frac{S_{1b}^2}{m}$$

เมื่อ $$S_{1b}^2 = \frac{S_b^2}{\bar{\bar{N}}^2} = \frac{1}{\bar{\bar{N}}^2} \cdot \frac{1}{M-1} \sum_i^M (T_i - \bar{T})^2 = \frac{1}{M-1} \sum_i^M \left(\frac{T_i}{\bar{\bar{N}}} - \frac{\bar{T}}{\bar{\bar{N}}} \right)^2$$

$$= \frac{1}{M-1} \sum_i^M (\bar{\bar{X}}_i - \bar{\bar{X}})^2$$

สำหรับเทอมที่สองจะพบว่า

$$\frac{1}{\bar{N}^2}\frac{M}{m}\sum_i^M \frac{\bar{N}-\bar{n}}{\bar{N}}\bar{N}^2\frac{S_i^2}{\bar{n}} = \frac{1}{\bar{N}^2}\frac{M}{m}\sum_i^M \frac{\bar{N}-\bar{n}}{\bar{N}}\left(\frac{N}{M}\right)^2\frac{S_i^2}{\bar{n}} = \frac{1}{m}\sum_i^m \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{1}{M}\frac{S_i^2}{\bar{n}}$$

$$= \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{1}{M}\sum_i^M \frac{S_i^2}{m\bar{n}} = \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{S_{2i}^2}{m\bar{n}}$$

เมื่อ $S_{2i}^2 = \frac{1}{M}\sum_i^M S_i^2 = \frac{1}{M}\sum_i^M \frac{1}{\bar{N}-1}\sum_j^{N_i}(x_{ij}-\bar{\bar{X}}_i)^2 = \frac{1}{M(\bar{N}-1)}\sum_i^M\sum_j^{N_i}(x_{ij}-\bar{\bar{X}}_i)^2$

ดังนั้น

5. $$V(\hat{\bar{\bar{X}}}) = \frac{M-m}{M}\frac{S_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{1}{M}\sum_i^M \frac{S_i^2}{m\bar{n}}$$

หรือ $$V(\hat{\bar{\bar{X}}}) = \frac{M-m}{M}\frac{S_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{S_{2i}^2}{m\bar{n}}$$

ในทำนองเดียวกันเราสามารถพิสูจน์ได้ว่า $\hat{V}(\hat{\bar{\bar{X}}})$ เป็นตัวประมาณค่าที่ปราศจากอคติของ $V(\hat{\bar{\bar{X}}})$ เมื่อ

6. $$\hat{V}(\hat{\bar{\bar{X}}}) = \frac{M-m}{M}\frac{s_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{1}{M}\sum_i^m \frac{s_i^2}{m\bar{n}}$$

หรือ $$\hat{V}(\hat{\bar{\bar{X}}}) = \frac{M-m}{M}\frac{s_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{m}{M}\frac{s_{2i}^2}{m\bar{n}}$$

เมื่อ $$s_{1b}^2 = \frac{1}{m-1}\sum_i^m(\bar{\bar{x}}_i-\bar{\bar{x}})^2 \; ; \; \bar{\bar{x}}_i = \frac{1}{\bar{n}}\sum_j^{\bar{n}} x_{ij}, \; \bar{\bar{x}} = \sum_i^m\sum_j^{\bar{n}} x_{ij}$$

$$s_i^2 = \frac{1}{\bar{n}-1}\sum_j^{\bar{n}}(x_{ij}-\bar{\bar{x}}_i)^2$$

$$s_{2i}^2 = \frac{1}{m}\sum_i^m s_i^2$$

7. เนื่องจาก p_i เป็นกรณีเฉพาะของ $\bar{\bar{x}}_i$ และ p เป็นกรณีเฉพาะของ $\bar{\bar{x}}$ ดังนี้

$$\hat{V}(\hat{P}) = \frac{M-m}{M}\frac{s_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{m}{M}\sum_i^m \frac{\bar{n}_i p_i q_i}{m\bar{n}(\bar{n}-1)}$$

ตัวอย่าง 5.6 การศึกษาประชาบาลต้องการทราบว่าเด็กนักเรียนชั้นประถมปีที่ 5 ของโรงเรียนประถมศึกษาในภาคตะวันออกเฉียงเหนือเสียเงินค่าอาหารกลางวันสัปดาห์ละเท่าไร

จากการศึกษาพบว่าในภาคตะวันออกเฉียงเหนือมีนักเรียนชั้นประถมปีที่ 5 รวม 400 ชั้นเรียน แต่ละชั้นเรียนมีนักเรียนห้องละ 40 คน สุ่มตัวอย่างชั้นเรียนมาเป็นตัวอย่าง 10 ชั้นเรียนแต่ละชั้นที่เป็นตัวอย่างสุ่มนักเรียนมาเป็นตัวอย่างชั้นละ 10 คน แล้วสอบถามเกี่ยวกับเงินที่ใช้จ่ายเป็นค่าอาหารกลางวันในสัปดาห์ที่ 2 ของเดือน ผลการบันทึกข้อมูลปรากฏดังนี้

ชั้นเรียนที่	Σx_{ij}	ชั้นเรียนที่	Σx_{ij}
1	25	6	18
2	30	7	24
3	20	8	27
4	19	9	20
5	13	10	21

ก. จงกะประมาณค่าใช้จ่ายเฉลี่ยสำหรับอาหารกลางวันใน 1 สัปดาห์ของนักเรียนแต่ละคน

ข. คำนวณช่วงเชื่อมั่น 95% ของการประมาณค่าในข้อ ก. ทั้งนี้ขอให้กำหนดข้อจำกัดตามความเหมาะสม

วิธีทำ ก. $M = 400, m = 10, \bar{N} = 40, \bar{n} = 10$

วิธีที่ 1

$$\hat{\bar{\bar{X}}} = \frac{1}{m\bar{n}} \sum_{i}^{m} \sum_{j}^{\bar{n}} x_{ij}$$

$$= \frac{1}{10 \times 10} (25 + 30 + 20 + \ldots + 20 + 21) = 227/100$$

$$= 2.27$$

นั่นคือ นักเรียนคนหนึ่งจะใช้ค่าใช้จ่ายเป็นค่าอาหารกลางวันต่อสัปดาห์คนละ 2.27 บาท

วิธีที่ 2 $M = 400, m = 10, N_1 = N_2 = \ldots = N_{400} = 40$

$n_1 = n_2 = \ldots = n_{10} = 10$

$$\hat{\bar{\bar{X}}} = \frac{1}{N} \frac{M}{m} \sum_{i}^{m} \frac{N_i}{n_i} \sum_{j}^{n_i} x_{ij} = \frac{1}{M\bar{N}} \cdot \frac{M}{M} \sum_{i}^{m} \frac{N_i}{n_i} \sum_{j}^{n_i} x_{ij}$$

$$= \frac{1}{400 \times 40} (400/10) \left(\frac{40}{10} \times 25 + \frac{40}{10} \times 30 + \ldots + \frac{40}{10} \times 20 + \frac{40}{10} \times 21 \right)$$

$$= \frac{1}{400} \times 908 = 2.27$$

ข. $\hat{V}(\hat{\bar{\bar{X}}}) = \frac{M-m}{M} \frac{s_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}} \frac{m}{M} \frac{s_{2i}^2}{m\bar{n}}$

แต่เมื่อพิจารณาข้อมูล จะเห็นว่ามิได้เสนอค่า x_{ij} ไว้ให้แต่เสนอเพียง $\sum_{j}^{n_i} x_{ij}$ ดังนั้น จึงไม่อาจคำนวณหา s_i^2 ได้เพราะ $s_i^2 = \frac{1}{n_i - 1} \sum_{j}^{n_i} (x_{ij} - \bar{x}_i)^2$ ต้องการข้อมูล x_{ij} ด้วย เมื่อพิจารณารูป $\hat{V}(\hat{\bar{\bar{X}}})$ จะพบว่า ถ้า $M >> m$ และ $\bar{N} >> \bar{n}$ แล้วสูตร $\hat{V}(\hat{\bar{\bar{X}}}) = \frac{M-m}{M} \frac{s_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}} \frac{m}{M} \frac{s_i^2}{m\bar{n}}$ จะลดรูปลงเหลือเพียง $\hat{V}(\bar{\bar{x}}) \simeq \frac{s_{1b}^2}{m}$

ในที่นี้ $M = 400$ และ $m = 10$ แสดงว่า $M >> m$ จะเห็นว่า $\frac{M-m}{M} = \frac{390}{400} = .975 \cong 1$

ขณะเดียวกัน $\frac{\bar{N}-\bar{n}}{\bar{N}} \frac{m}{M} \frac{s_{2i}^2}{m\bar{n}} = \frac{30}{40} \frac{10}{400} \frac{s_{2i}^2}{10 \times 10} = .00018\, s_{2i}^2 \simeq 0$

ดังนั้น $\hat{V}(\hat{\bar{\bar{X}}}) \simeq s_{1b}^2/m$

$$s_{1b}^2 = \frac{1}{m-1}\left\{\sum_i^m (\bar{x}_i - \bar{\bar{x}})^2\right\} = \frac{1}{m-1}\left\{\sum_i^m \bar{x}_i^2 - \frac{(\sum^m \bar{x}_i)^2}{m}\right\}$$

$$= \frac{1}{9}\{(2.5^2+3^2+2^2+1.9^2+\ldots+2^2+2.1^2)$$

$$-(2.5+3+2+1.9+\ldots+2+2.1)^2/10\}$$

$$= \frac{1}{9}(49.25 - \frac{21.7^2}{10}) = .240$$

ดังนั้น $\hat{V}(\hat{\bar{\bar{X}}}) \cong s_{1b}^2/m = .240/10 = .0240$

$$s_{\hat{\bar{\bar{x}}}} = .155$$

ดังนั้นช่วงเชื่อมั่น 95% ที่คาดว่าค่าจริงของ $\bar{\bar{X}}$ จะปรากฏอยู่คือ

$$2.27 - (1.96)(.024) < \bar{\bar{X}} < 2.27 + (1.96)(.024)$$

$$2.22 < \bar{\bar{X}} < 2.32$$

หรือสามารถกล่าวได้ด้วยความเชื่อมั่น 95% ได้ว่านักเรียนชั้นประถมศึกษาปีที่ 5 ในภาคตะวันออกเฉียงเหนือคนหนึ่ง ๆ จะใช้จ่ายเป็นค่าอาหารต่อสัปดาห์ประมาณ 2.22 ถึง 2.32 บาท

5.2.7 การจัดสรรขนาดตัวอย่าง

ปัญหาในทางปฏิบัติอีกประการหนึ่งของงานสำรวจด้วยตัวอย่างโดยใช้แผน CS คือการกำหนดจำนวนตัวอย่าง psu และจำนวนตัวอย่าง ssu บ่อยครั้งที่เราจำเป็นต้องเลี่ยงความยุ่งยากเรื่องนี้ไปโดยกำหนดตัวอย่างให้เป็นสัดส่วนกับขนาดประชากรโดยสุ่มตัวอย่าง psu และ ssu มาเป็นตัวอย่างระหว่าง 5-60% ของขนาดประชากร เช่นการสำรวจปัญหาความประพฤติของนักเรียนอาชีวศึกษาในเขตกทม. การวิจัยดำเนินการโดยสุ่มตัวอย่างโรงเรียนอาชีวศึกษามาเป็นตัวอย่าง 30% แต่ละโรงเรียนที่เป็นตัวอย่างสุ่มตัวอย่างนักเรียน

มาเป็นตัวอย่างโรงเรียนละ 30% ดังนี้เป็นต้น และมีบ่อยครั้งที่เราสุ่มตัวอย่างจำนวนนักเรียนมาเป็นตัวอย่างโรงเรียนละ เท่า ๆ กัน เมื่อพิจารณาเห็นว่าจำนวนนักเรียนในแต่ละโรงเรียนมิได้แตกต่างกันนัก[1]

อย่างไรก็ตามเพื่อให้การกำหนดตัวอย่างเป็นไปโดยชัดเจนมิได้อ้างอิงหลักการเชิงอัตตนิยมมากนัก เราพึงเสาะหาวิธีกำหนดตัวอย่างเป็นไปอย่างเหมาะสมกับงบประมาณที่มีอยู่อย่างจำกัดด้วย วิธีจัดสรรดังกล่าวคือ Optimum Allocation และเพื่อความสะดวกการจัดสรรโดยนัยดังกล่าวจะอาศัยสูตรความแปรปรวนเฉพาะจากตอน 5.2.6 นักศึกษาที่สนใจและเห็นว่าจะเป็นการศึกษาเฉพาะเจาะจงในกรณีของ Cluster มีขนาดเดียวกันเท่านั้นให้พัฒนาสูตรเองโดยอาศัยสูตรความแปรปรวนทั่วไป วิธีพัฒนาให้เลียนแบบวิธีการที่จะดำเนินการในลำดับต่อไปนี้ อนึ่ง ในที่นี้จะพัฒนาเฉพาะวิธีกำหนดค่า m และ n_i สำหรับประมาณค่าเฉลี่ยประชากรเท่านั้น การกำหนดขนาดตัวอย่าง m และ n_i สำหรับประมาณยอดรวม สัดส่วน และอัตราส่วนจะขอเว้นไว้เป็นแบบฝึกหัด

ให้ C_1 และ C_2 เป็นค่าใช้จ่ายในการเลือก psu และ ssu ตามลำดับ[2] และให้ C คือ ค่าใช้จ่ายรวม ขณะที่ C_0 คือค่าใช้จ่ายคงที่ ดังนั้น งบประมาณรวมเพื่อการวิจัยจึงเสนอเป็นสมการเชิงเส้นได้ดังนี้

$$C = C_0 + C_1 m + C_2 m\bar{n}$$

หรือเมื่อหักค่าใช้จ่ายคงที่ C_0 ออกจากงบประมาณ C ทั้งหมดจะได้สมการที่เกี่ยวข้องกับงานสำรวจ (Variable Cost) ดังนี้คือ

$$C' = C_1 m + C_2 m\bar{n} \quad \text{เมื่อ } C' = C - C_0$$

---

1. กล่าวแล้วในตอน 5.2.6

2. C_2 คือค่าใช้จ่ายในการเลือก ssu ต่อหน่วย รายละเอียดให้อ่านวิธีพิจารณาค่าใช้จ่ายในการสำรวจในบทที่ 3 ตอน 3.3.3 ส่วน C_1 โดยทั่วไปมักหมายถึงค่าใช้จ่ายในการเดินทางไปสู่ psu psu ที่อยู่ห่างไกลศูนย์ปฏิบัติงานมากกว่าจะทำให้เสียค่าใช้จ่ายสูงกว่า

เมื่อขนาดตัวอย่างภายใต้ข้อจำกัดของงบประมาณ C' และมีผลให้ความแปรปรวนของการประมาณค่าต่ำสุดจึงอาศัยวิธีของ Lagrange ดังนี้

$$F = \hat{V}(\hat{\bar{\bar{X}}}) + \lambda(C_1 m + C_2 m\bar{n} - C')$$ [1]

$$=> F = \frac{M-m}{M}\frac{S_{1b}^2}{m} + \frac{\bar{N}-\bar{n}}{\bar{N}}\frac{S_{2i}^2}{m\bar{n}} + \lambda(C_1 m + C_2 m\bar{n} - C')$$

$$= \frac{S_{1b}^2}{m} - \frac{S_{1b}^2}{M} + \frac{S_{2i}^2}{m\bar{n}} - \frac{S_{2i}^2}{m\bar{N}} + \lambda C_1 m + \lambda C_2 m\bar{n} - \lambda C'$$

$$\frac{\partial F}{\partial m} = 0 => -\frac{S_{1b}^2}{m^2} - \frac{S_{2i}^2}{\bar{n}m^2} + \frac{S_{2i}^2}{\bar{N}m^2} + \lambda C_1 + \lambda C_2\bar{n} = 0 \quad \text{..........(1)}$$

$$\frac{\partial F}{\partial \bar{n}} = 0 => (-S_{2i}^2/m\bar{n}^2) + \lambda C_2 m = 0 \quad \text{..........(2)}$$

$$(1) \times m \quad \lambda C_1 m + \lambda C_2 m\bar{n} = \frac{S_{1b}^2}{m} + \frac{S_{2i}^2}{m\bar{n}} - \frac{S_{2i}^2}{\bar{N}m} \quad \text{..........(3)}$$

$$(2) \times \bar{n} \quad \lambda C_2 m\bar{n} = S_{2i}^2/m\bar{n} \quad \text{..........(4)}$$

$$(3) - (4) \quad \lambda C_1 m = \frac{S_{1b}^2}{m} - \frac{S_{2i}^2}{\bar{N}m}$$

$$=> \quad \lambda m^2 = \frac{S_{1b}^2 - (S_{2i}^2/\bar{N})}{C_1} \quad \text{..........(5)}$$

จาก (2) $\quad \lambda m^2 = S_{2i}^2/C_2\bar{n}^2 \quad$ ....(6)

$$(6) \div (5) \quad 1 = \frac{S_{2i}^2}{C_2\bar{n}^2}\frac{C_1}{S_{1b}^2 - (S_{2i}^2/\bar{N})}$$

$$\bar{n}^2 = \frac{C_1}{C_2} \cdot \frac{S_{2i}^2}{S_{1b}^2 - (S_{2i}^2/\bar{N})} \quad \text{..........(7)}$$

$$\bar{n} = \sqrt{\frac{C_1}{C_2} \cdot \frac{S_{2i}^2}{S_{1b}^2 - (S_{2i}^2/\bar{N})}}$$

---

[1] $C_1 m + C_2 m\bar{n} - C = 0$ คือ Constraint Condition

$$\cong \sqrt{\frac{C_1}{C_2} \cdot \frac{s_{2i}^2}{s_{1b}^2 - (s_{2i}^2/\overline{N})}}$$

แทนค่า $\bar{n} = \sqrt{\frac{C_1}{C_2} \cdot \frac{S_{2i}^2}{S_{1b}^2 - (S_{2i}^2/\overline{N})}}$ ลงใน Cost Function $C' = C_1m + C_2m\bar{n}$

ดังนั้น

$$\begin{aligned}
C' &= m(C_1 + C_2\bar{n}) \\
&= m(C_1 + C_2\sqrt{\frac{C_1}{C_2} \cdot \frac{S_{2i}^2}{S_{1b}^2 - (S_{2i}^2/\overline{N})}}\,) \\
&= m(C_1 + \sqrt{\frac{C_1C_2S_{2i}^2}{S_{1b}^2 - (S_{2i}^2/\overline{N})}}\,)
\end{aligned}$$

$$m = \frac{C'}{C_1 + \sqrt{\frac{C_1C_2s_{2i}^2}{s_{1b}^2 - (s_{2i}^2/\overline{N})}}}$$

หมายเหตุ

1. $s_{1b}^2 = \frac{1}{m-1}\sum_{i}^{m}(\bar{\bar{x}}_i - \bar{\bar{x}})^2$ เป็นตัวประมาณค่าที่ปราศจากอคติของ S_{1b}^2

โดยที่ $S_{1b}^2 = \frac{1}{M-1}\sum_{i}^{M}(\bar{\bar{X}}_i - \bar{\bar{X}})^2$

2. $s_{2i}^2 = \frac{1}{m}\sum_{i}^{m} s_{2i}^2$ เป็นตัวประมาณค่าที่ปราศจากอคติของ S_{2i}^2

เมื่อ $S_{2i}^2 = \frac{1}{M}\sum_{i}^{m} S_i^2$ แต่ s_i^2 เป็นตัวประมาณค่าที่ปราศจากอคติของ S_i^2

โดยที่ $s_i^2 = \frac{1}{n_i - 1}\sum_{j}^{n_i}(x_{ij} - \bar{\bar{x}}_i)^2$ และ $S_i^2 = \frac{1}{N_i - 1}\sum_{j}^{N_i}(x_{ij} - \bar{\bar{X}}_i)^2$

ตัวอย่าง 5.7 จากตัวอย่าง 5.6 ถ้าถือว่าข้อมูลดังกล่าวเป็นผลการสำรวจเบื้องต้น จงกำหนดขนาดตัวอย่าง m และ $\bar{n}$ ถ้า $C' = 12{,}000$ บาท $C_1 = 55$ บาท $C_2 = 180$ บาท และสมมุติว่า $s_{2i}^2 = 9.14$

วิธีทำ จากตัวอย่าง 5.6 พบว่า $s_{1b}^2 = .240$

ดังนั้น
$$m \simeq \frac{C'}{C_1 + \sqrt{\dfrac{C_1 C_2 s_{2i}^2}{s_{1b}^2 - (s_{2i}^2/\bar{N})}}}$$

$$\simeq \frac{12000}{45 + \sqrt{\dfrac{55 \times 180 \times 9.14}{.240 - (9.14/40)}}}$$

$$\simeq 4.21 \simeq 4$$

$$\bar{n} \simeq \sqrt{\frac{C_1}{C_2} \cdot \frac{s_{2i}^2}{s_{1b}^2 - (s_{2i}^2/\bar{N})}}$$

$$\simeq \sqrt{\frac{55}{180} \frac{9.14}{.240 - (9.14/40)}}$$

$$\simeq 15.58 \simeq 16$$

5.3 3CS

แผนสำรวจแบบ Three Stage Cluster Sampling (3CS) เป็นแผนสำรวจที่ขยายออกไปจาก 2 CS อีก 1 ชั้น หมายความว่าการสำรวจจะดำเนินการเป็น 3 ชั้นคือ

ชั้นที่ 1 สุ่มตัวอย่าง psu มาเป็นตัวอย่างจำนวนหนึ่ง

ชั้นที่ 2 สุ่มตัวอย่าง ssu จากแต่ละ psu ที่เป็นตัวอย่าง

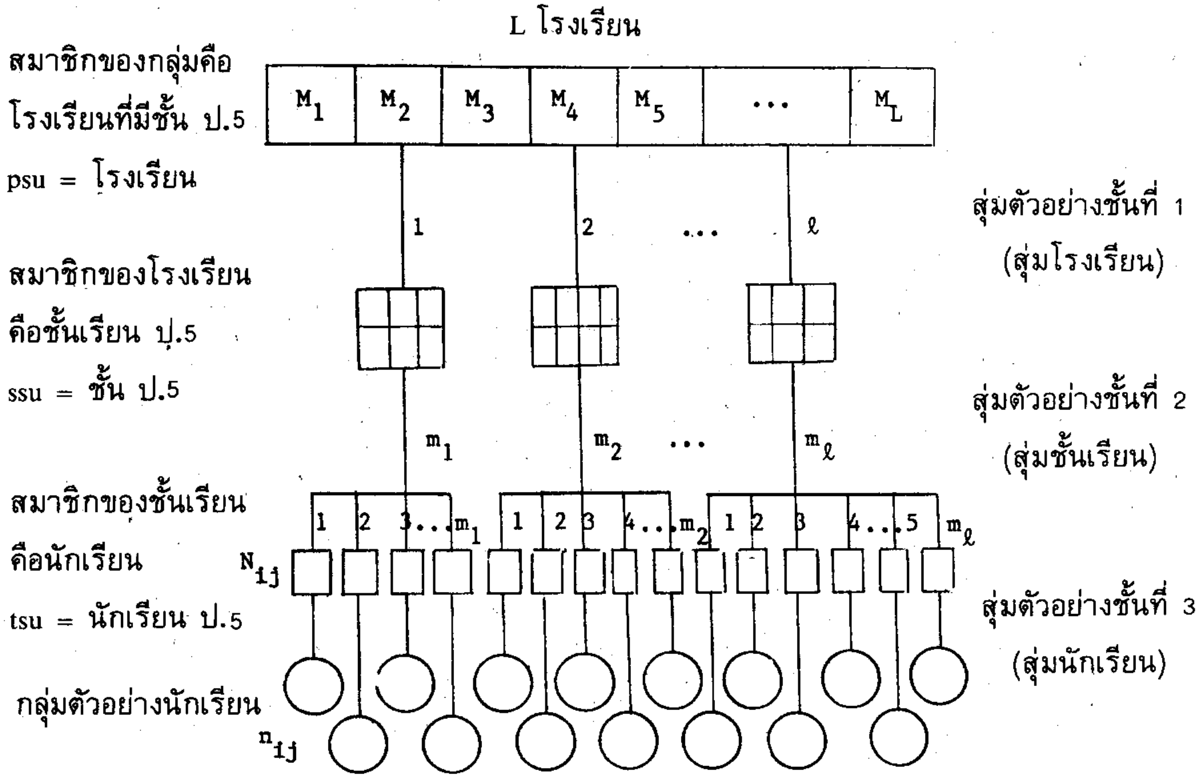
ชั้นที่ 3 สุ่มตัวอย่าง tsu จากแต่ละ ssu ที่เป็นตัวอย่าง

ตัวอย่างเช่น ในการสำรวจค่าใช้จ่ายสำหรับอาหารกลางวันของนักเรียนชั้นประถมศึกษาปีที่ 5 ในภาคตะวันออกเฉียงเหนือ โรงเรียนจะทำหน้าที่เป็น psu ชั้นเรียนทำหน้าที่เป็น ssu (แต่ละโรงเรียนมีชั้นประถมศึกษาปีที่ 5 มากกว่า 1 ชั้น) ตัวนักเรียนทำหน้าที่เป็น tsu ดังนั้นการสำรวจเราจะสุ่มโรงเรียนมาเป็นตัวอย่าง แต่ละโรงเรียนที่เป็นตัวอย่างให้สุ่มชั้นเรียนประถมศึกษาปีที่ 5 มาเป็นตัวอย่างแต่ละชั้นเรียนที่เป็นตัวอย่าง

ให้สุ่มตัวอย่างนักเรียนมาเป็นตัวอย่าง แล้วสอบถามค่าใช้จ่ายสำหรับอาหารกลางวันจากกลุ่มตัวอย่าง

หรือการสำรวจทัศนคติในด้านการวางแผนครอบครัวของประชาชนในเขตอำเภอโพธาราม จังหวัดราชบุรี วิธีสำรวจให้สุ่มตำบล (psu) มาเป็นตัวอย่าง แต่ละตำบลที่เป็นตัวอย่างให้สุ่มหมู่บ้าน (ssu) มาเป็นตัวอย่าง แต่ละหมู่บ้านที่เป็นตัวอย่างให้สุ่มครัวเรือนมาเป็นตัวอย่างแล้วสัมภาษณ์คู่สมรสเกี่ยวกับทัศนคติเกี่ยวกับการวางแผนครอบครัว

ดังนั้นเราจึงสามารถแสดงไดอะแกรมของแผนสำรวจ 3 CS ได้ดังนี้ และเพื่อให้เข้าใจง่ายขอแสดงไดอะแกรมของตัวอย่างที่ 1 ส่วนไดอะแกรมของตัวอย่างที่ 2 ได้แสดงไว้แล้วในตอน 5.1



ดังนั้นเราสามารถกำหนดนิยามและสัญลักษณ์เพิ่มเติมได้ดังนี้

5.3.1 นิยามและสัญลักษณ์

L = จำนวน psu ในกลุ่มประชากรของ psu

ℓ = จำนวน psu ที่เลือกเป็นตัวอย่าง

M_i = ขนาดของ psu ที่ i หรือ M_i คือจำนวน ssu ใน psu ที่ i;
$i = 1,2,\ldots,L$

m_i = จำนวน ssu ที่เลือกมาเป็นตัวอย่างจาก psu ที่ i;
$i = 1,2,\ldots,\ell$

N_{ij} = จำนวน tsu ใน ssu ที่ j ที่สุ่มมาจาก psu ที่ i;
$i = 1,2,\ldots,\ell$; $j = 1,2,\ldots,M_i$

n_{ij} = ขนาดตัวอย่างของหน่วย tsu ที่สุ่มจาก ssu ที่ j ของ psu ที่ i;
$i = 1,2,\ldots,\ell$; $j = 1,2,\ldots,m_i$

x_{ijk} = ค่าสังเกตของหน่วยสำรวจ (tsu) ที่ k ที่สุ่มจาก ssu ที่ j ของ psu
i; $k = 1,2,\ldots,n_{ij}$; $j = 1,2,\ldots,m_i$; $i = 1,2,\ldots,\ell$

เพื่อให้เข้าใจสัญลักษณ์เหล่านี้ขอยกตัวอย่างประกอบดังนี้

สมมุติในเขตบางกะปิมีโรงเรียนประถมศึกษา 5 โรงเรียน แต่ละโรงเรียนมีชั้น ป.5 อยู่ 6,7,4,9,10 ชั้นเรียนตามลำดับ แต่ละชั้นประกอบด้วยนักเรียนชั้นละ 35 คน เท่า ๆ กัน

การสำรวจค่าใช้จ่ายสำหรับอาหารกลางวันดำเนินการโดยใช้แผนสำรวจแบบ 3CS โดยสุ่มตัวอย่าง psu ssu และ tsu มา 40% ของที่มีอยู่ นั่นคือสุ่มโรงเรียน (psu) มาเป็นตัวอย่าง 2 โรงเรียน สมมุติได้โรงเรียนที่ 2 และโรงเรียนที่ 5 ซึ่งมีชั้นเรียนป.5 จำนวน 7 และ 10 ห้องตามลำดับ ดังนั้น ชั้นเรียน (ssu) ตัวอย่างคือ 3 ห้องจากโรงเรียนที่ 2 และ 4 ห้องจากโรงเรียนที่ 5 แต่ละห้องสุ่มนักเรียนมาเป็นตัวอย่าง 40% คือห้องละ 14 คน ดังนั้นไดอะแกรมสำหรับแผนสำรวจนี้จึงปรากฏดังภาพหน้าต่อไป

ตามไดอะแกรมนี้นักศึกษาจะพบว่า

L = จำนวนโรงเรียน = 5

ℓ = จำนวนโรงเรียน (psu) ที่เลือกมาเป็นตัวอย่าง = 2

M_i = จำนวนห้องเรียนชั้นป.5 (ssu) ในโรงเรียนที่ $i; i=1,2,\ldots,5$
นั่นคือ $M_1=6,\ M_2=7,\ M_3=4,\ M_4=9,\ M_5=10$

m_i = จำนวนห้องเรียนชั้นป.5 (ssu) ที่เลือกมาเป็นตัวอย่างจากโรงเรียน (psu) ที่ i $i=1,2,\ldots,\ell$ นั่นคือ $m_1=3,\ m_2=m_\ell=4$

N_{ij} = จำนวนนักเรียน (tsu) ในชั้นเรียนป.5 (ssu) ที่ j ที่สุ่มจากโรงเรียน (psu) ที่ i; $i=1,2,\ldots,\ell;\ j=1,2,\ldots,m_i$; ในที่นี้ $\ell=2,\ m_1=3,\ m_2=4$

นั่นคือ N_{11} = จำนวนนักเรียนในชั้นที่ 1 ของโรงเรียนตัวอย่างที่ 1 = 35

N_{12} = จำนวนนักเรียนในชั้นที่ 2 ของโรงเรียนตัวอย่างที่ 1 = 35

N_{13} = จำนวนนักเรียนในชั้นที่ 3 ของโรงเรียนตัวอย่างที่ 1 = 35

N_{21} = จำนวนนักเรียนในชั้นที่ 1 ของโรงเรียนตัวอย่างที่ 2 = 35

N_{22} = จำนวนนักเรียนในชั้นที่ 2 ของโรงเรียนตัวอย่างที่ 2 = 35

N_{23} = จำนวนนักเรียนในชั้นที่ 3 ของโรงเรียนตัวอย่างที่ 2 = 35

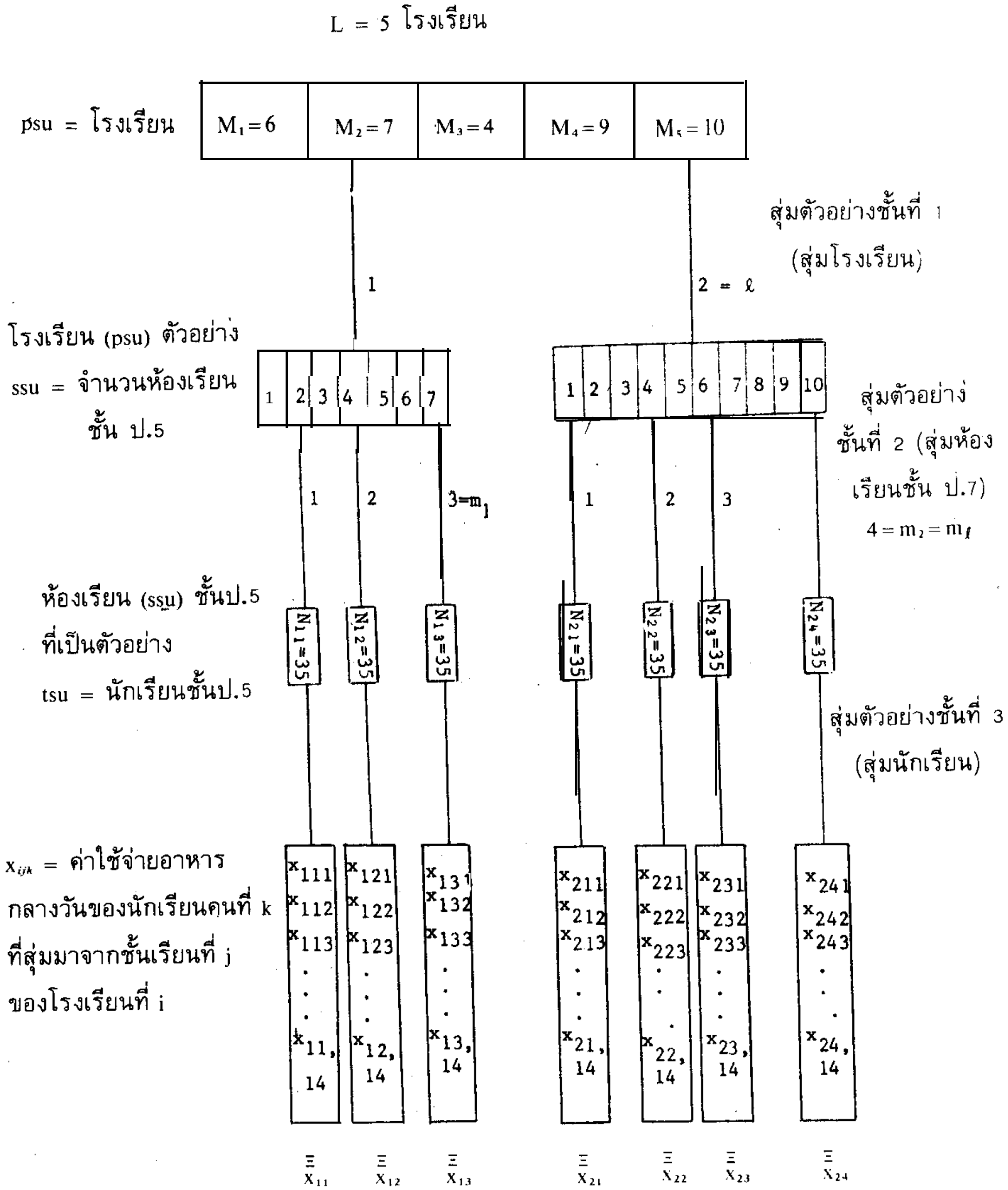
N_{24} = จำนวนนักเรียนในชั้นที่ 4 ของโรงเรียนตัวอย่างที่ 2 = 35

n_{ij} = จำนวนนักเรียน (tsu) ที่สุ่มมาเป็นตัวอย่างจากห้องเรียนที่ j จากโรงเรียนที่ i ในที่นี้ $n_{11}=n_{12}=n_{13}=n_{21}=n_{22}=n_{23}=n_{24}=14$

x_{ijk} = ค่าใช้จ่ายค่าอาหารกลางวันของนักเรียนคนที่ k ที่สุ่มจากชั้นเรียนที่ j ของโรงเรียนที่ i; $k=1,2,\ldots,n_{ij};\ j=1,2,\ldots,m_i;\ i=1,2,\ldots,\ell$

สมมุติว่า $x_{136}=26$ บาท หมายความว่านักเรียนคนที่ 6 ที่สุ่มมาจากชั้นเรียนตัวอย่างที่ 3 ของโรงเรียนตัวอย่างที่ 1 จ่ายค่าอาหารกลางวันอาทิตย์ละ 26 บาท

เมื่อนักศึกษาเข้าใจได้ตามนี้แล้วจึงขอเพิ่มนิยามและสัญลักษณ์ต่อไปนี้ อนึ่งเพื่อความเข้าใจที่ง่ายและชัดเจนขอให้สังเกตว่านิยามและสัญลักษณ์ต่าง ๆ นี้ยึดถือแนวทางของการประมาณค่าในลักษณะ "จากล่างขึ้นบน" ดังที่ได้ปฏิบัติมาแล้ว ใน 2CS ถ้าหากนักศึกษาไม่ปฏิบัติตามแนวนี้จะเกิดความยุ่งยากในการศึกษาเป็นอย่างมากและเสียเวลาและสิ้นเปลืองสติปัญญาในการจดจำทั้ง ๆ ที่ไม่มีความจำเป็น


L = 5 โรงเรียน
psu = โรงเรียน
M₁=6 M₂=7 M₃=4 M₄=9 M₅=10
สุ่มตัวอย่างชั้นที่ 1
(สุ่มโรงเรียน)
1
2 = ℓ
โรงเรียน (psu) ตัวอย่าง
ssu = จำนวนห้องเรียน
ชั้น ป.5
1 2 3 4 5 6 7
1 2 3 4 5 6 7 8 9 10
สุ่มตัวอย่าง
ชั้นที่ 2 (สุ่มห้อง
เรียนชั้น ป.7)
4 = m₂ = mℓ
1 2 3=m₁
1 2 3
ห้องเรียน (ssu) ชั้นป.5
ที่เป็นตัวอย่าง
tsu = นักเรียนชั้นป.5
N₁₁=35 N₁₂=35 N₁₃=35 N₂₁=35 N₂₂=35 N₂₃=35 N₂₄=35
สุ่มตัวอย่างชั้นที่ 3
(สุ่มนักเรียน)
xᵢⱼₖ = ค่าใช้จ่ายอาหาร
กลางวันของนักเรียนคนที่ k
ที่สุ่มมาจากชั้นเรียนที่ j
ของโรงเรียนที่ i
x111 x112 x113 . . . x11,14
x121 x122 x123 . . . x12,14
x131 x132 x133 . . . x13,14
x211 x212 x213 . . . x21,14
x221 x222 x223 . . . x22,14
x231 x232 x233 . . . x23,14
x241 x242 x243 . . . x24,14
X̄11 X̄12 X̄13 X̄21 X̄22 X̄23 X̄24

$\bar{\bar{\bar{X}}}_{ij} = \frac{1}{N_{ij}} \sum_{k}^{N_{ij}} x_{ijk}$ = ค่าเฉลี่ยต่อ tsu ใน ssu ที่ j ของ psu ที่ i

$\bar{\bar{\bar{x}}}_{ij} = \frac{1}{n_{ij}} \sum_{k}^{n_{ij}} x_{ij}$ = ค่าเฉลี่ยของตัวอย่างต่อ tsu ของ ssu ที่ j จาก psu ที่ i

เช่น $\bar{\bar{\bar{x}}}_{13}$ = 18.50 บาท หมายความว่านักเรียนในห้องที่ 3 ที่สุ่มมาจากโรงเรียนตัวอย่างที่ 1 เสียค่าใช้จ่ายอาหารกลางวันต่อสัปดาห์เฉลี่ยคนละ 18.50 บาท

$T_{ij} = N_{ij}\bar{\bar{\bar{X}}}_{ij}$ = ค่ายอดรวมของคุณลักษณะที่สนใจของ ssu ที่ j จาก psu ที่ i

$\hat{T}_{ij} = N_{ij}\bar{\bar{\bar{x}}}_{ij}$ = ค่าประมาณยอดรวมของคุณลักษณะที่น่าสนใจของ ssu ที่ j จาก psu ที่ i

$\bar{\bar{T}}_{i} = \frac{1}{M_i} \sum_{j}^{M_i} T_{ij}; i = 1,2,\ldots,L$ = ยอดรวมเฉลี่ยต่อ ssu ของ psu ที่ i

$\hat{\bar{\bar{T}}}_{i} = \frac{1}{m_i} \sum_{j}^{m_i} \hat{T}_{ij}; i = 1,2,\ldots,\ell$ = ค่าประมาณยอดรวมเฉลี่ยต่อ ssu ของ psu ที่ i

$T_i = M_i\bar{\bar{T}}_i$ = ยอดรวมของคุณลักษณะที่สนใจของ psu ที่ i; $i = 1,2,\ldots,L$

$\hat{T}_i = M_i\hat{\bar{\bar{T}}}_i$ = ค่าประมาณยอดรวมของคุณลักษณะที่สนใจของ psu ที่ i; $i = 1,2,\ldots,\ell$

$\bar{T} = \frac{1}{L} \sum_{i}^{L} T_i$ = ยอดรวมเฉลี่ยต่อ psu

$\hat{\bar{T}} = \frac{1}{\ell} \sum_{i}^{\ell} \hat{T}_i$ = ค่าประมาณยอดรวมเฉลี่ยต่อ psu

$L\bar{T}$ = ยอดรวมของคุณลักษณะที่สนใจในกลุ่มประชากร

$L\hat{\bar{T}}$ = ค่าประมาณยอดรวมของคุณลักษณะที่สนใจในกลุ่มประชากร

5.3.2 การประมาณค่ายอดรวมประชากร

ทฤษฎี 5.2 เมื่อดำเนินการสุ่มตัวอย่างโดยใช้แผนสำรวจแบบ 3 CS โดยที่ทำการสุ่ม

ตัวอย่าง psu ssu และ tsu โดยวิธี SRS จะพบว่า

ก. $\hat{T} = L\hat{\bar{T}} = \frac{L}{\ell} \sum_{j}^{\ell} \frac{M_i}{m_i} \sum_{j}^{m_i} \frac{N_{ij}}{n_{ij}} \sum_{k}^{n_{ij}} x_{ijk}$ เป็นตัวประมาณค่าที่ปราศจากอคติ

ของยอดรวมประชากร T เมื่อ $T = \sum_{i}^{L} \sum_{j}^{M_i} \sum_{k}^{N_{ij}} x_{ijk}$

ข. $$V(\hat{T}) = L^2 \frac{L-\ell}{L} \frac{S_b^2}{\ell} + \frac{L}{\ell} \sum_{i}^{L} M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i}$$

$$+ \frac{L}{\ell} \sum_{i}^{L} \frac{M_i}{m_i} \sum_{j}^{M_i} N_{ij}^2 \frac{N_{ij} - n_{ij}}{N_{ij}} \frac{S_{ij}^2}{n_{ij}}$$

โดยที่ $$S_b^2 = \frac{1}{L-1} \sum_{i}^{L} (T_i - \bar{T})^2, \; S_i^2 = \frac{1}{M_i - 1} \sum_{j}^{M_i} (T_i - \bar{\bar{T}}_i)^2 ; \; i = 1,2,\ldots,L$$

$$S_{ij}^2 = \frac{1}{N_{ij} - 1} \sum_{k}^{N_{ij}} (x_{ijk} - \bar{\bar{\bar{X}}}_{ij})^2$$

$$\hat{V}(\hat{T}) = L^2 \frac{L-\ell}{L} \frac{s_b^2}{\ell} + \frac{L}{\ell} \sum_{i}^{\ell} M_i^2 \frac{M_i - m_i}{M_i} \frac{s_i^2}{m_i}$$

$$+ \frac{L}{\ell} \sum_{i}^{\ell} \frac{M_i}{m_i} \sum_{j}^{m_i} N_{ij}^2 \frac{N_{ij} - n_{ij}}{N_{ij}} \frac{s_{ij}^2}{n_{ij}}$$

$$s_b^2 = \frac{1}{\ell - 1} \sum_{i}^{\ell} (\hat{T}_i - \hat{\bar{T}})^2 ; \; \hat{T}_i = M_i \hat{\bar{\bar{T}}}, \; \hat{\bar{T}}_i = \frac{1}{\ell} \sum_{i}^{\ell} \hat{T}_i$$

$$s_i^2 = \frac{1}{m_i - 1} \sum_{j}^{m_i} (\hat{T}_{ij} - \hat{\bar{\bar{T}}}_i)^2$$

$$; \; \hat{T}_{ij} = N_{ij}\bar{\bar{x}}_{ij}, \; \hat{\bar{\bar{T}}} = \frac{1}{m_i} \sum_{j}^{m_i} \hat{T}_{ij} ; \; i = 1,2,\ldots,\ell$$

$$s_{ij}^2 = \frac{1}{n_{ij} - 1} \sum_{k}^{n_{ij}} (x_{ijk} - \bar{\bar{\bar{x}}}_{ij})^2 ; \; \bar{\bar{\bar{x}}}_{ij} = \frac{1}{n_{ij}} \sum_{k}^{n_{ij}} x_{ijk} ; \; i = 1,2,\ldots,\ell$$

$$j = 1,2,\ldots,m_i$$

พิสูจน์

ก. $E(\hat{T}) = T$

การพิสูจน์ว่า $E(\hat{T}) = T$ เป็นเรื่องที่ไม่ยุ่งยากแต่ประการใดจึงขอเว้นไว้เป็นแบบฝึกหัด

ในที่นี้จะแสดงให้เห็นพัฒนาหรือความเป็นมาของสูตรตัวประมาณค่าของ T คือ $\hat{T}$ ว่าพัฒนาอย่างไร

การพัฒนาให้ย้อนจากล่างขึ้นบน (ดูไดอะแกรม) เช่นเดียวกับกรณีของ 2CS ดังนี้

$$\bar{\bar{x}}_{ij} = \frac{1}{n_{ij}} \sum_{k}^{n_{ij}} x_{ijk} = \text{ค่าเฉลี่ยตัวอย่าง tsu ใน ssu ที่ j ของ psu ที่ i}$$

ดังนั้น
$$\hat{T}_{ij} = N_{ij}\bar{\bar{x}}_{ij} = \text{ยอดรวมต่อ ssu ที่ j ของ psu ที่ i}$$

แต่เราสุ่มตัวอย่าง ssu จาก psu มา m_i หน่วย ดังนั้นค่าเฉลี่ยของยอดรวมต่อ 1 ssu ใน psu ที่ i คือ

$$\hat{\bar{\bar{T}}}_i = \frac{1}{m_i} \sum_{j}^{m_i} \hat{T}_{ij}\,;\ i = 1,2,\ldots,\ell$$

แต่ในแต่ละ psu ประกอบด้วย ssu ทั้งสิ้น M_i หน่วย ดังนั้นยอดรวมต่อ psu ที่ i คือ

$$\hat{T}_i = M_i\hat{\bar{\bar{T}}}_i = \frac{M_i}{m_i} \sum_{j}^{m_i} \hat{T}_{ij} = \frac{M_i}{m_i} \sum_{j}^{m_i} N_{ij}\bar{\bar{x}}_{ij} = \frac{M_i}{m_i} \sum_{j}^{m_i} \frac{N_{ij}}{n_{ij}} \sum_{k}^{n_{ij}} x_{ijk}$$

แต่เนื่องจากเราสุ่มตัวอย่าง psu มาเพียง ℓ หน่วย ดังนั้นค่าเฉลี่ยของยอดรวมต่อ 1 psu คือ

$$\hat{\bar{T}} = \frac{1}{\ell} \sum_{i}^{\ell} \hat{T}_i = \frac{1}{\ell} \sum_{i}^{\ell} \frac{M_i}{m_i} \sum_{j}^{m_i} \frac{N_{ij}}{n_{ij}} \sum_{k}^{n_{ij}} x_{ijk}$$

และเนื่องจากในกลุ่มประชากรประกอบไปด้วย psu ทั้งสิ้น L หน่วย ดังนั้นยอดรวมทั้งสิ้นของกลุ่มประชากรคือ

$$\hat{T} = L\hat{\bar{T}} = \frac{L}{\ell} \sum_{i}^{\ell} \frac{M_i}{m_i} \sum_{j}^{m_i} \frac{N_{ij}}{n_{ij}} \sum_{k}^{n_{ij}} x_{ijk}$$

ขอให้สังเกตว่าวิธีประมาณค่ายอดรวมสำหรับ 3CS มิใช่สิ่งที่ยากเลย การประมาณค่าให้ดำเนินการจากชั้นล่างสุดของสายการสุ่มตัวอย่างของแต่ละสาย โดยประมาณค่าเฉลี่ยแล้วประมาณยอดรวม แล้วเฉลี่ยแล้วประมาณยอดรวม กระทำดังนี้สลับกันไปจนถึงสุดของสายการสุ่มตัวอย่างโดยยึดถือเทคนิคการประมาณยอดรวมจากแนวทางในบทที่ 2 คือ

$\hat{T} = N\bar{x}$ สิ่งหนึ่งที่ขอให้ระวังไว้ก็คือการตีความหมาย คำว่ายอดรวมเฉลี่ยต่อ ssu หรือยอดรวมเฉลี่ยต่อ psu นั้น มิได้หมายถึงค่าเฉลี่ยต่อหน่วยสำรวจเล็กที่สุด แต่มีความหมายตามคำพูดนั้น เช่น psu = ตำบล ssu = หมู่บ้าน tsu = ครัวเรือน ยอดรวมเฉลี่ยต่อ ssu หมายถึง โดยเฉลี่ยแล้วหมู่บ้านหนึ่ง ๆ ของตำบลที่ i จะมียอดรวมโดยถัวเฉลี่ยเท่ากับค่าหนึ่ง สมมุติการสำรวจเรื่องจำนวนของคนว่างงานถัวเฉลี่ยต่อหมู่บ้านของตำบลที่ i ส่วนยอดรวมถัวเฉลี่ยต่อ psu หมายถึงยอดรวมของคนว่างงานถัวเฉลี่ยต่อ 1 ตำบลในอำเภอที่สำรวจ

ข. การพัฒนาสูตร $V(\hat{T})$ ของกรณี 3CS ให้ยึดถือแนวทางการพัฒนาสูตร $V(\hat{T})$ ของกรณี 2CS โดยให้ถือเสมือนว่าเรากำลังดำเนินการสำรวจ 2CS 2 ครั้ง คือระหว่าง psu กับ ssu ครั้งหนึ่งและ ssu กับ tsu อีกครั้งหนึ่ง จากนั้นจึงนำสูตร $V(\hat{T})$ ทั้ง 2 ครั้งนั้นมารวมกันก็จะได้ $V(\hat{T})$ ของกรณี 3CS ตามต้องการ โครงสร้างของสูตรยึดถือโครงสร้างของ 2CS เปลี่ยนแปลงเพียงสัญลักษณ์และ subscript เฉพาะชั้นของการสำรวจเท่านั้น โครงสร้างของสูตร $V(\hat{T})$ สำหรับ 2CS คือ

$$V(\hat{T}) \;=\; M^2\,\frac{M-m}{M}\,\frac{S_b^2}{m} + \frac{M}{m}\sum_{i}^{M}\frac{N_i-n_i}{N_i}\,\frac{N_i^2S_i^2}{n_i}$$

การพัฒนาจะเป็นไปได้โดยง่ายดังนี้คือ

ครั้งที่ 1 ระหว่าง psu กับ ssu

โดยเลียนแบบโครงสร้าง $V(\hat{T})$ ของ 2CS จะพบว่า

$$V(\hat{T}) \;=\; L^2\,\frac{L-\ell}{L}\,\frac{S_b^2}{\ell} + \frac{L}{\ell}\sum_{i}^{L}\frac{M_i-m_i}{M_i}\,\frac{M_i^2S_i^2}{m_i} \;=\; A + \frac{L}{\ell}\sum_{i}^{L}B_i \quad (1)$$

ทั้งนี้ $S_b^2 = \dfrac{1}{L-1}\sum_{i}^{L}(T_i-\bar{T})^2$ และ $S_i^2 = \dfrac{1}{M_i-1}\sum_{j}^{M_i}(T_{ij}-\bar{\bar{T}}_i)^2$

ครั้งที่ 2 ระหว่าง ssu กับ tsu

โดยเลียนแบบโครงสร้าง $V(\hat{T})$ ของ 2 CS จะพบว่า

$$V(\hat{T}) \;=\; M_i^2\,\frac{M_i-m_i}{M_i}\,\frac{S_i^2}{m_i} + \frac{M_i}{m_i}\sum_{j}^{M_i}\frac{N_{ij}-n_{ij}}{N_{ij}}\,\frac{N_{ij}^2S_{ij}^2}{n_{ij}} = C \quad (2)$$

โดยที่

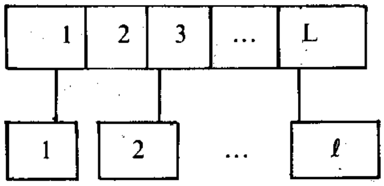
$$S_{ij}^2 \;=\; \frac{1}{N_{ij}-1}\sum_{k}^{N_{ij}}(x_{ijk}-\bar{\bar{\bar{X}}}_{ij})^2$$

การนำสมการ (1) และ (2) มารวมกันขอให้พิจารณาด้วยความระมัดระวังและขอให้นักศึกษาพยายามเปรียบเทียบไดอะแกรมไว้ตลอดเวลา การพิจารณาให้เริ่มดังนี้

จากสมการที่ (1)

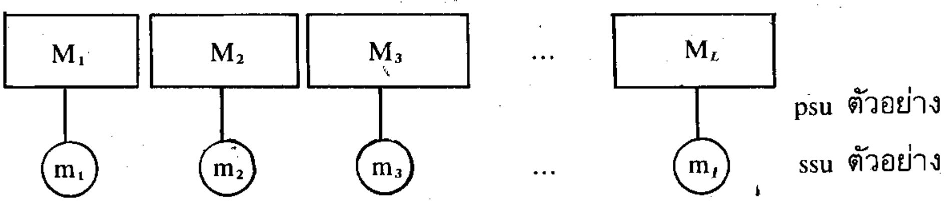
$$V(\hat{T}) = L^2 \frac{L-\ell}{L} \frac{S_b^2}{\ell} + \frac{L}{\ell} \sum_i^L M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i} = A + \frac{L}{\ell} \sum_i^L B_i$$

1. จะเห็นว่าส่วนแรกคือ $\frac{L-\ell}{L} \frac{S_b^2}{\ell}$ เป็นความผันแปรระหว่าง psu ซึ่งเกิดจากการเลือก psu ที่มีอยู่ L หน่วย มาเป็นตัวอย่าง ℓ หน่วยโดยสุ่ม (SRS) สูตรโครงสร้าง A = $L^2 \frac{L-\ell}{L} \frac{S_b^2}{\ell}$ จึงเป็นสูตรประมาณยอดรวมระหว่าง psu ตามวิธี SRS [1] ดังภาพ



2. เมื่อพิจารณาในส่วนที่ 2 คือ $\frac{L}{\ell} \sum_i^L M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i}$

จะพบว่า $\sum_i^L M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i}$ คือความผันแปรระหว่าง ssu ในระหว่างชั้นภูมิต่าง ๆ สูตรโครงสร้าง $\sum_i^L M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i}$ เป็นสูตรประมาณยอดรวมระหว่าง ssu [2] ตามวิธีแบ่งชั้นภูมิที่มีอยู่ทั้งสิ้น L ชั้นภูมิ แต่ละชั้นภูมิสุ่มตัวอย่าง ssu มาโดยวิธี SRS สังเกตได้จากโครงสร้าง $B_i = M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i}$ ดังภาพ



3. แต่จะเห็นว่าการสุ่มตัวอย่างจากแต่ละ psu ที่เป็นตัวอย่างเรามิได้ใช้แผนสำรวจแบบ SRS แต่ใช้วิธีสำรวจแบบ 2 CS ดังนั้นในส่วนที่ 2 ของสูตรสมการที่ (1) จึงต้องแทนที่ B_i ด้วย C เมื่อ

---

1. ใน SRS จะพบว่า $\hat{T} = N\bar{x}, V(\hat{T}) = N^2V(\bar{x}) = N^2 \frac{N-n}{N} \frac{S^2}{n}$

2. ใน Str จะพบว่า $\hat{T} = \sum_i^L N_h\bar{x}_h, V(\hat{T}) = \sum_h^L N_h^2 \frac{N_h - n_h}{N_h} \frac{S_h^2}{n_h}$

$$B_i = M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i}$$

และ $$C = M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i} + \frac{M_i}{m_i} \sum_j^{M_i} N_{ij}^2 \frac{N_{ij} - n_{ij}}{N_{ij}} \frac{S_{ij}^2}{n_{ij}}$$

ดังนั้น $V(\hat{T})$ สำหรับ 3 CS จึงปรากฏดังนี้

$$V(\hat{T}) = L^2 \frac{L-\ell}{L} \frac{S_b^2}{\ell} + \frac{L}{\ell} \sum_i^L B_i = L^2 \frac{L-\ell}{L} \frac{S_b^2}{\ell} + \frac{L}{\ell} \sum_i^L C$$

$$= L^2 \frac{L-\ell}{L} \frac{S_b^2}{\ell} + \frac{L}{\ell} \sum_i^L \{ M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m^i}$$

$$+ \frac{M_i}{m_i} \sum_j^{M_i} N_{ij}^2 \frac{N_{ij} - n_{ij}}{N_{ij}} \frac{S_{ij}^2}{n_{ij}} \}$$

$$=> V(\hat{T}) = L^2 \frac{L-\ell}{L} \frac{S_b^2}{\ell} + \frac{L}{\ell} \sum_i^L M_i^2 \frac{M_i - m_i}{M_i} \frac{S_i^2}{m_i} +$$

$$\frac{L}{\ell} \sum_i^L \frac{M_i}{m_i} \sum_j^{M_i} N_{ij}^2 \frac{N_{ij} - n_{ij}}{N_{ij}} \frac{S_{ij}^2}{n_{ij}}$$

สำหรับตัวประมาณค่าที่ปราศจากอคติของ V(T) คือ $\hat{V}(\hat{T})$ ก็พิจารณาได้ทำนองเดียวกันนี้

$$=> \hat{V}(\hat{T}) = L^2 \frac{L-\ell}{L} \frac{s_b^2}{\ell} + \frac{L}{\ell} \sum_i^{\ell} M_i^2 \frac{M_i - m_i}{M_i} \frac{s_i^2}{m_i}$$

$$+ \frac{L}{\ell} \sum_i^{\ell} \frac{M_i}{m_i} \sum_j^{m_i} N_{ij}^2 \frac{N_{ij} - n_{ij}}{N_{ij}} \frac{s_{ij}^2}{n_{ij}}$$

หมายเหตุ จากสูตร $V(\hat{T})$ นักศึกษาจะพบว่า $V(\hat{T})$ ประกอบไปด้วยความแปรปรวนถึง 3 ส่วน คือความผันแปรระหว่าง psu คือ S_b^2 ความผันแปรระหว่าง ssu คือ S_i^2 และความผันแปรระหว่าง tsu คือ S_{ij}^2 และจากโครงสร้างจะสังเกตได้ว่า S_b^2 ถูกควบคุมด้วยจำนวนตัวอย่าง psu คือ ℓ นั่นคือ $\frac{S_b^2}{\ell}$ ขณะที่ S_i^2 ถูกควบคุมด้วยจำนวนตัวอย่าง ssu คือ m_i นั่นคือ $\frac{S_i^2}{m_i}$ และ S_{ij}^2 ถูกควบคุมด้วยจำนวนตัวอย่าง tsu คือ n_{ij} นั่นคือ $\frac{S_{ij}^2}{n_{ij}}$ ซึ่งถ้าเราสุ่มตัวอย่าง