$$= \frac{M}{m} \sum_{i}^{M} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i} + 0 + 0 + M^2 \frac{M-m}{M} \frac{S_b^2}{m}$$

$$= M^2 \frac{M-m}{M} \frac{S_b^2}{m} + \frac{M}{m} \sum_{i}^{M} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

โดยที่ $S_b^2 = \frac{1}{M-1} \sum_{i}^{M} (T_i - \bar{T})^2$ และ $S_i^2 = \frac{1}{N_i - 1} \sum_{j}^{N_i} (x_{ij} - \bar{X}_i)^2$ [1]

พิสูจน์ $\hat{V}(\hat{T}) = M^2 \frac{M-m}{M} \frac{s_b^2}{m} + \frac{M}{m} \sum_{i}^{m} \frac{N_i - n_i}{N_i} \frac{N_i^2 s_i^2}{n_i}$ เป็นตัวประมาณค่าที่ปราศจากอคติของ $V(\hat{T})$ หรือ $E(\hat{V}(\hat{T})) = V(\hat{T})$

1. เราสามารถพิสูจน์ได้เช่นเดียวกันกับในบทที่ 2 ว่า $E(s_i^2) = S_i^2$; $i = 1,2,\ldots,m$ ทั้งนี้เพราะถือว่าแต่ละ psu ตัวอย่างคือกลุ่มประชากรย่อยของ ssu และเราสุ่ม ssu มาโดยวิธี SRS

2. $E(s_b^2) \neq S_b^2$

$$\because \; s_b^2 = \frac{1}{m-1} \sum_{i}^{m} (\hat{T}_i - \bar{\hat{T}})^2 \qquad \bar{\hat{T}} = \frac{1}{m} \sum_{i}^{m} \hat{T}_i$$

$$= \frac{1}{m-1} \sum_{i}^{m} (\hat{T}_i - \frac{M}{M} \frac{1}{m} \sum_{i}^{m} \hat{T}_i)^2 \text{ : คูณด้วย } \frac{m}{m} \text{ และ } \frac{M}{M}$$

เพื่อประโยชน์ในการจัดรูป

$$= \frac{1}{m-1} \frac{m}{m} \sum_{i}^{m} (\hat{T}_i - \frac{\hat{T}}{M})^2 \quad \because \hat{T} = M\bar{\hat{T}} = M \cdot \frac{1}{m} \sum_{i}^{m} \hat{T}_i$$

$$=> \frac{m-1}{m} s_b^2 = \frac{1}{m} \sum_{i}^{m} (\hat{T}_i - \frac{\hat{T}}{M})^2$$

$$= \frac{1}{m} \sum_{i}^{m} \{\hat{T}_i^2 - 2\hat{T}_i \frac{\hat{T}}{M} + (\frac{\hat{T}}{M})^2\}$$

---

[1] เกี่ยวกับ $\hat{V}(\hat{T})$ มีข้อที่น่าสังเกตก็คือ โดยปกติเท่าที่ผ่านมาเราจะหา $\hat{V}(\hat{T})$, $\hat{V}(\bar{x})$, $\hat{V}(\hat{P})$ หรือ $\hat{V}(\hat{R})$ ได้โดยง่ายเพียงแต่แทนที่ S^2 ด้วย s^2 โดยไม่เปลี่ยนเงื่อนไขโครงสร้างอื่น ๆ ของสูตรแต่สำหรับ 2CS ขอให้สังเกตว่านอกจากแทนที่ S_i^2 ด้วย s_i^2 และแทนที่ S_b^2 ด้วย s_b^2 แล้วยังเปลี่ยนโครงสร้างของสูตร ขอให้สังเกตเทอมขวามือใน $V(\hat{T})$ เราจะ sum ถึง M และใน $\hat{V}(\hat{T})$ เราจะ sum ถึง m ที่เป็นเช่นนี้เพราะ $E(s_b^2) \neq S_b^2$

$$= \frac{1}{m}\left\{\sum_{i}^{m}\hat{T}_i - 2m\frac{\hat{T}}{M}\cdot\frac{1}{m}\sum_{i}^{m}\hat{T}_i + m\left(\frac{\hat{T}}{M}\right)^2\right\}$$

$$= \frac{1}{m}\left\{\sum_{i}^{m}\hat{T}_i^2 - 2m\frac{\hat{T}}{M}\hat{\bar{T}} + m\left(\frac{\hat{T}}{M}\right)^2\right\}$$

$$= \frac{1}{m}\left\{\sum_{i}^{m}\hat{T}_i^2 - 2m\frac{\hat{T}}{M}\cdot\frac{M}{M}\hat{\bar{T}} + m\left(\frac{\hat{T}}{M}\right)^2\right\}$$

$$= \frac{1}{m}\left\{\sum_{i}^{m}\hat{T}_i^2 - 2m\left(\frac{\hat{T}}{M}\right)^2 + m\left(\frac{\hat{T}}{M}\right)^2\right\}$$

$$=> \frac{m-1}{m}s_b^2 = \frac{1}{m}\left\{\sum_{i}^{m}\hat{T}_i^2 - m(\hat{T}/M)^2\right\} = \frac{1}{m}\sum_{i}^{m}\hat{T}_i^2 - (\hat{T}/M)^2$$

$$\therefore E\left(\frac{m-1}{m}s_b^2\right) = E_iE_j\left(\frac{1}{m}\sum_{i}^{m}\hat{T}_i^2\right) - E_iE_j(\hat{T}/M)^2$$

$$= E_i\left(\frac{1}{m}\sum_{i}^{m}E_j\hat{T}_i^2\right) - \frac{E(\hat{T})^2}{M^2}$$

$$\because \hat{T}_i = \hat{T}_i - T_i + T_i = (\hat{T}_i - T_i) + T_i$$

ดังนั้น $\hat{T}_i^2 = (\hat{T}_i - T_i)^2 + T_i^2 + 2(\hat{T}_i - T_i)T_i$

$$=> E_j(\hat{T}_i^2) = E_j(\hat{T}_i - T_i)^2 + E_jT_i^2 + 2T_iE_j(\hat{T}_i - T_i)$$

$$= V(\hat{T}_i) + T_i^2 = \frac{N_i - n_i}{N_i}\,\frac{N_i^2S_i^2}{n_i} + T_i^2$$

$$\because E_j(\hat{T}_i - T_i) = E_j(\hat{T}_i) - T_i$$

$$= T_i - T_i = 0$$

และ $\hat{T} = \hat{T} - T + T = (\hat{T} - T) + T$

$$=> \hat{T}^2 = (\hat{T} - T)^2 + T^2 + 2T(\hat{T} - T)$$

$$E(\hat{T}^2) = E(\hat{T} - T)^2 + ET^2 + 2TE(\hat{T} - T)$$

$$= V(\hat{T}) + T^2 \because E(\hat{T} - T) = E(\hat{T}) - T = T - T = 0$$

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

ดังนั้น $E(\frac{m-1}{m} s_b^2) = E_i\{\frac{1}{m}\sum_i^m(\frac{N_i-n_i}{N_i}\frac{N_i^2S_i^2}{n_i}+T^2_i)\} -$

$$\frac{1}{M^2}(M^2\frac{M-m}{M}\frac{S_b^2}{m}+\frac{M}{m}\sum_i^M\frac{N_i-n_i}{N_i}\frac{N_i^2S_i^2}{n_i}+T^2)$$

ให้ $\frac{N_i-n_i}{N_i}\frac{N_i^2S_i^2}{n_i}+T_i^2 = V_i$

$$=> E_i\{\frac{1}{m}\sum_i^m(\frac{N_i-n_i}{N_i}\frac{N_i^2S_i^2}{n_i}+T_i^2)\} = E_i(\frac{1}{m}\sum_i^m V_i)$$

$$= \frac{1}{m}\sum_i^m E_i(V_i) = \frac{1}{m}\sum_i^m(\sum_i^M V_i\frac{1}{M})$$

$$= \frac{1}{m}\cdot m\cdot\frac{1}{M}\sum_i^M V_i = \frac{1}{M}\sum_i^M V_i$$

$$= \frac{1}{M}\sum_i^M(\frac{N_i-n_i}{N_i}\frac{N_i^2S_i^2}{n_i}+T_i^2)$$

ให้ $A = \sum_i^M\frac{N_i-n_i}{N_i}\frac{N_i^2S_i^2}{n_i}$

ดังนั้น $E(\frac{m-1}{m}s_b^2) = \frac{1}{M}(A+\sum_i^M T_i^2)-\frac{1}{M^2}(M^2\frac{M-m}{M}\frac{S_b^2}{m}+\frac{M}{m}\cdot A+T^2)$

$$= \frac{A}{M}+\frac{1}{M}\sum_i^M T_i^2-\frac{M-m}{M}\frac{S_b^2}{M}-\frac{A}{Mm}-(\frac{T}{M})^2$$

$$= \frac{A}{M}-\frac{A}{Mm}+\{\frac{1}{M}\sum_i^m T_i^2-(\frac{T}{M})^2\}-\frac{M-m}{M}\frac{S_b^2}{m}$$

$$= \frac{1}{M}(\frac{m-1}{m})A+\frac{1}{M}(\sum_i^M T_i^2-M\bar{T}^2)-\frac{M-m}{M}\frac{S_b^2}{m}$$

$$= \frac{1}{M}(\frac{m-1}{m})A+\frac{1}{M}\{(M-1)S_b^2\}-\frac{M-m}{M}\frac{S_b^2}{m}$$

$$= \frac{1}{M}(\frac{m-1}{m})A+\frac{S_b^2}{M}\{(M-1)-\frac{M-m}{m}\}$$

$$= \frac{1}{M}(\frac{m-1}{m})A+\frac{S_b^2}{M}(\frac{mM-m-M+m}{m})$$

$$=> E\left(\frac{m-1}{m} s_b^2\right) = \frac{1}{M}\left(\frac{m-1}{m}\right) A + \left(\frac{m-1}{m}\right) S_b^2$$

$$=> E(s_b^2) = S_b^2 + A/M$$

$$= S_b^2 + \frac{1}{M} \sum_{i}^{M} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

$=> E(s_b^2) \neq S_b^2$ หรือ s_b^2 มิใช่ตัวประมาณค่าที่ปราศจากอคติของ S_b^2

3. $E(\hat{V}(\hat{T})) = V(\hat{T})$ หรือ $\hat{V}(\hat{T})$ เป็นตัวประมาณค่าที่ปราศจากอคติของ $V(\hat{T})$

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

พิสูจน์ $E(\hat{V}(\hat{T})) = M^2 \frac{M-m}{Mm} E(s_b^2) + \frac{M}{m} E_i \sum_{i}^{m} \frac{N_i - n_i}{N_i} \frac{N_i^2}{n_i} E_j(s_i^2)$

$$= A + B$$

$$A = M^2 \frac{M-m}{Mm} E(s_b^2) = M^2 \frac{M-m}{Mm} \left(S_b^2 + \frac{1}{M} \sum_{i}^{M} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}\right)$$

$$= M^2 \frac{M-m}{M} \frac{S_b^2}{m} + \frac{M-m}{m} \sum_{i}^{M} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

$$B = \frac{M}{m} E_i \sum_{i}^{m} \frac{N_i - n_i}{N_i} \frac{N_i^2}{n_i} E_j(s_i^2)$$

$$= \frac{M}{m} E_i \sum_{i}^{m} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

ให้ $V_i = \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$

$$=> B = \frac{M}{m} \sum_{i}^{m} E_i(V_i) = \frac{M}{m} \sum_{i}^{m} \left(\sum_{i}^{M} V_i \frac{1}{M} \right)$$

$$= \frac{M}{m} \cdot m \sum_{i}^{M} \frac{V_i}{M} = \sum_{i}^{M} V_i$$

$$= \sum_{i}^{M} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

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

$$+ \sum_i^M \frac{N_i-n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

$$= M^2 \frac{M-m}{M} \frac{S_b^2}{m} + \left(\frac{M-m}{m} + 1\right) \sum_i^M \frac{N_i-n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$$

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

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

ตัวอย่าง 5.1 หน่วยงานของรัฐแห่งหนึ่งซึ่งประกอบด้วยแผนกต่าง ๆ $M = 50$ แผนกต้องการกะประมาณยอดรวมของจำนวนประชาชนที่เข้ามาติดต่อขอรับบริการต่าง ๆ ในระหว่างสัปดาห์แรกของเดือน

ดำเนินการสุ่มตัวอย่างแผนก (psu) มาเป็นตัวอย่าง $m = 5$ แผนกแต่ละแผนกตัวอย่างประกอบด้วยพนักงาน $N_1 = 10, N_2 = 15, N_3 = 10, N_4 = 20, N_5 = 15$ คนตามลำดับ และสุ่มตัวอย่างพนักงาน (ssu) จากแผนกตัวอย่างมาแผนกละ 4,5,3,4 และ 4 คนตามลำดับ และจากการสัมภาษณ์พนักงานเหล่านี้ถึงจำนวนประชาชนที่มาติดต่อและเขาให้การบริการในระหว่างสัปดาห์แรกของเดือนปรากฏข้อมูลดังนี้

แผนกที่	N_i	n_i	x_{ij}
1	10	4	5,3,0,8
2	15	5	11,6,5,9,4
3	10	3	4,375
4	20	4	6,8,2,4
5	15	4	2,0,9,9

ก. จงกะประมาณยอดรวมของประชากรที่หน่วยราชการแห่งนี้ให้การบริการในระหว่างสัปดาห์แรกของเดือน

ข. จงกะประมาณยอดรวมด้วยช่วงเชื่อมั่น 95%

วิธีทำ ก. เราสามารถเตรียมตารางวิเคราะห์ได้ดังนี้

แผนกที่	N_i	n_i	x_{ij}	Σx_{ij}	$\bar{\bar{x}}_i = \frac{1}{n_i}\sum_j^{n_i} x_{ij}$	$\hat{T}_i = N_i\bar{\bar{x}}_i$
1	10	4	5,3,0,8	16	4	40
2	15	5	11,6,5,9,4	35	7	105
3	10	3	4,3,5	12	4	40
4	20	4	6,8,2,4	20	5	100
5	20	4	2,0,9,9	20	5	75
รวม						360

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

$$= 50 \times \frac{360}{5} = 50 \times 72$$

$$= 3{,}600$$

คาดว่าในสัปดาห์แรกของเดือนจะมีประชาชนมาติดต่อขอรับบริการจากส่วนราชการดังกล่าวประมาณ 3,600 คน

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

จาก $s_b^2 = \frac{1}{m-1}\sum_i^m (\hat{T}_i - \hat{\bar{T}})^2$ และ $s_i^2 = \frac{1}{n_i-1}\sum_j^{n_i}(x_{ij}-\bar{x}_i)^2;\ i = 1,2,\ldots,m$

$$s_b^2 = \frac{1}{m-1}\sum_i^m (\hat{T}_i - \hat{\bar{T}})^2 = \frac{1}{m-1}\left\{\sum_i^m \hat{T}_i^2 - \frac{(\Sigma\hat{T}_i)^2}{m}\right\}$$

$$= \frac{1}{5-1}\left\{(40^2+105^2+40^2+100^2+75^2) - \frac{(360)^2}{5}\right\}$$

$$= \frac{1}{4}(29{,}850 - 25{,}920)$$

$$= 982.5$$

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

$; \; i = 1,2,\ldots,m$

ดังนั้น $s_1^2 = \frac{1}{3}\left\{(25+9+0+64) - \frac{16^2}{4}\right\} = 11.33$

$$s_2^2 = \frac{1}{3}\left\{(121+36+25+81+16) - \frac{35^2}{5}\right\} = 8.50$$

$$s_3^2 = \frac{1}{3}\left\{(16+9+25) - \frac{12^2}{3}\right\} = 1.00$$

$$s_4^2 = \frac{1}{3}\left\{(36+64+4+16) - \frac{20^2}{4}\right\} = 6.67$$

$$s_5^2 = \frac{1}{3}\left\{(4+0+81+81) - \frac{20^2}{4}\right\} = 22.00$$

ดังนั้นเราสามารถเตรียมตารางสำหรับคำนวณหา $\hat{V}(\hat{T})$ ได้ดังนี้

แผนที่	N_i	n_i	$(N_i - n_i)/N_i$	N_i^2	s_i^2	$N_i^2 s_i^2$	$\frac{N_i - n_i}{N_i} \frac{N_i^2 s_i^2}{n_i}$
1	10	4	.60	100	11.33	1133	169.95
2	15	5	.67	225	8.50	1912.28	256.28
3	10	3	.70	100	1.00	100	23.33
4	20	4	.80	400	6.67	2,668	533.60
5	15	4	.73	225	22.00	4,950	903.38
รวม							1,886.54

$$\text{ดังนั้น } \hat{V}(\hat{T}) = M^2 \frac{M-m}{M} \frac{s_b^2}{m} + \frac{M}{m} \sum_{i}^{m} \frac{N_i - n_i}{N_i} \frac{N_i^2 s_i^2}{n_i}$$

$$= 50(50-5) \cdot \frac{982.5}{5} + \frac{50}{5} \cdot 1886.54$$

$$= 460,990.4$$

$$\Rightarrow \quad s(\hat{T}) = 678.96$$

ดังนั้นช่วงเชื่อมั่น 95% ที่คาดว่าค่าจริงของยอดรวมของจำนวนราษฎรที่มาติดต่อราชการ ณ หน่วยราชการแห่งนี้คือ $(T + Z_{\alpha/2} s(\hat{T}), T + Z_{1-\alpha/2} s(\hat{T}))$

$= (3600 - (1.96)(678.96), 3600 + (1,96)(678.96)) = (2269.2, 4930.7)$ หรือประมาณ 2,269 ถึง 4,931 คน

ตัวอย่าง 5.2 ในการวิจัยครั้งหนึ่งผู้วิจัยต้องการกะประมาณจำนวนครัวเรือนในพื้นที่แห่งหนึ่งเพื่อใช้เป็นกรอบตัวอย่าง สมมุติว่าในเขตพื้นที่แห่งนั้นสามารถจัดได้เป็น

8,000 บล๊อค สุ่มตัวอย่างบล๊อคมา m = 80 บล๊อค และในแต่ละบล๊อคเราสามารถแบ่งพื้นที่ออกเป็น 4 ส่วน (segment) เท่า ๆ กันแล้วสุ่มตัวอย่าง segment จากแต่ละบล๊อคตัวอย่างมาบล๊อคละ 2 segment เท่า ๆ กันแล้วนับจำนวนครัวเรือนในแต่ละ segment

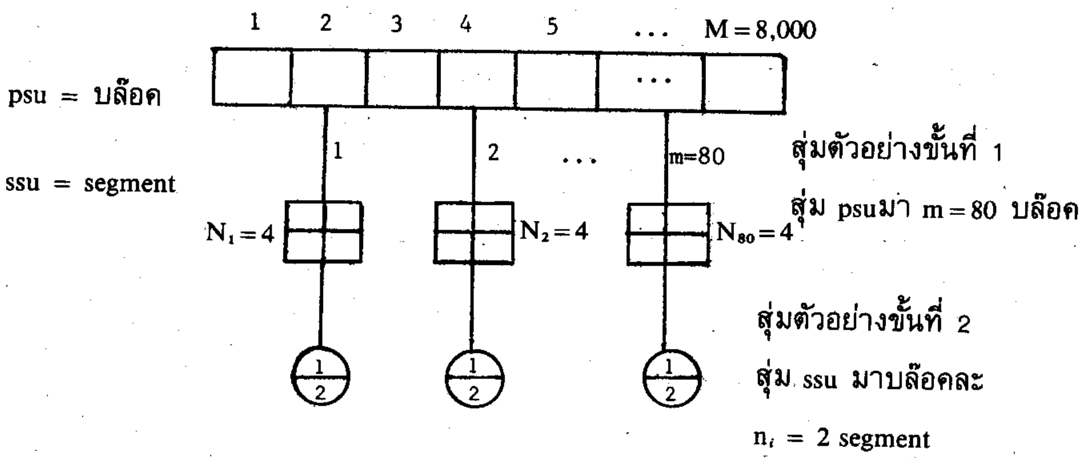
ก. จงวาดไดอะแกรมแสดงแผนสำรวจครั้งนี้

ข. จงแสดงให้เห็นว่า $\hat{V}(\hat{T}) \cong (\frac{M}{m})^2(80s_b^2 + 0.08 \sum_i^m \frac{s_i^2}{n_i})$

ค. ในสถานการณ์ใดจึงจะถือได้ว่า $\hat{V}(\hat{T}) \cong (\frac{M}{m})^2(80s_b^2)$

วิธีทำ

ก. แผนการสำรวจปรากฏดังไดอะแกรมต่อไปนี้



ข. เนื่องจาก $M = 8,000$ และ $m = 80$ จะเห็นว่า $M >> m$ ดังนั้น $\frac{M-m}{M} \approx 1$

ดังนั้น $\hat{V}(\hat{T}) \simeq M^2 \frac{s_b^2}{m} + \frac{M}{m} \sum_i^m \frac{N_i - n_i}{N_i} \frac{N_i^2 s_i^2}{n_i}$

$$\simeq (\frac{M}{m})^2(ms_b^2 + \frac{m}{M} \sum_i^m \frac{N_i - n_i}{N_i} \frac{N_i^2 s_i^2}{n_i})$$

$$\simeq \left(\frac{M}{m}\right)^2 \left(80s_b^2 + \frac{80}{8000} \sum_i^{80} \frac{4-2}{4}\, 4^2 \frac{s_i^2}{n_i}\right.$$

$$\simeq \left(\frac{M}{m}\right)^2 \left(80s_b^2 + \frac{80 \times 8}{8000} \sum_i^{80} \frac{s_i^2}{n_i}\right)$$

$$\simeq \left(\frac{M}{m}\right)^2 \left(80s_b^2 + .08 \sum_i^{80} \frac{s_i^2}{n_i}\right)$$

ค. จาก $\hat{V}(\hat{T}) \simeq \left(\frac{M}{m}\right)^2 \left(ms_b^2 + \frac{m}{M} \sum_i^m \frac{N_i - n_i}{N_i} \frac{N_i^2 s_i^2}{n_i}\right.$

ถ้า $n_i = N_i$ หรือในทุกบล๊อคตัวอย่างเรานับจำนวนครัวเรือนในทุก segment

ดังนั้น $N_i - n_i = 0$

$$=> V(T) \simeq \left(\frac{M}{m}\right)^2 (ms_b^2 + 0)$$

$$\simeq \left(\frac{M}{m}\right)^2 (80s_b^2)$$

5.2.3 การประมาณค่าเฉลี่ยประชากร $\bar{\bar{X}}$ [1]

ดังที่ได้กล่าวไว้แล้วในตอนต้นแล้วว่าในแผนสำรวจแบบ CS นั้น ถ้าเราสามารถประมาณค่ายอดรวมของประชากร T ได้ การประมาณค่าเฉลี่ย $\bar{\bar{X}}$ และสัดส่วน P ย่อมกระทำได้โดยง่าย เพียงแต่นำขนาดประชากร N ไปหาร $\hat{T}$ ก็จะได้ค่าประมาณของ $\bar{\bar{X}}$ และ P ตามต้องการ เฉพาะในส่วนที่เกี่ยวกับ P นั้นค่าของ x_{ij} จะมีได้เพียง 2 ค่า คือ $x_{ij} = 0$ เมื่อ $x_{ij} \in$ C และ $x_{ij} = 1$ เมื่อ $x_{ij} \in C$ ซึ่งจะได้กล่าวถึงเป็นการเฉพาะในตอน 5.2.4

ด้วยเหตุผลดังที่กล่าว เราจึงสามารถประมาณค่าเฉลี่ย $\bar{\bar{X}}$ พร้อมทั้งช่วงเชื่อมั่น $(1-\alpha)100\%$ ได้ดังนี้

---

1. $\bar{X}$ = ค่าจริงของค่าเฉลี่ยต่อ ssu

บทแทรก 5.1[1] เมื่อดำเนินการสุ่มตัวอย่างโดยใช้แผนสำรวจแบบ 2CS แล้วเราสามารถประมาณค่าเฉลี่ยประชากร $\bar{\bar{X}}$ ได้ดังนี้

1. $\hat{\bar{\bar{X}}} = \frac{\hat{T}}{N} = \frac{1}{N}\frac{M}{m}\sum_{i}^{m}\frac{N_i}{n_i}\sum_{j}^{n_i}x_{ij}$ เป็นตัวประมาณค่าที่ปราศจากอคติของ $\bar{\bar{X}}$[2]

2. $$V(\hat{\bar{\bar{X}}}) = V(\hat{T}/N) = \frac{1}{N^2}V(\hat{T})$$
$$= \frac{1}{N^2}M^2\frac{M-m}{M}\frac{S_b^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i}^{M}\frac{N_i-n_i}{N_i}\frac{N_i^2S_i^2}{n_i}$$

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

หรือ $$\hat{V}(\hat{\bar{\bar{X}}}) = \frac{1}{N^2}M^2\frac{M-m}{M}\frac{s_b^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i}^{m}\frac{N_i-n_i}{N_i}\frac{N_i^2s_i^2}{n_i}$$

ง. ช่วงเชื่อมั่นที่ $(1-\alpha)100\%$ ของ $\bar{\bar{X}}$ คือช่วง $(\hat{\bar{\bar{X}}} + Z_{\alpha/2}s(\hat{\bar{\bar{X}}}), \hat{\bar{\bar{X}}} + Z_{1-\alpha/2}s(\hat{\bar{\bar{X}}}))$

สำหรับการพิสูจน์บทแทรกนี้ผู้เขียนจะขอเว้นไว้เป็นแบบฝึกหัด นักศึกษาสามารถพิสูจน์ได้เองโดยง่ายเพราะบทแทรกเหล่านี้เป็นกรณีเฉพาะของทฤษฎี 5.1

---

[1] บทแทรกของทฤษฎี 5.1

[2] $\hat{\bar{\bar{X}}} = \frac{\hat{T}}{N}$ มิใช่ sample mean เพราะ sample mean คือ $\bar{\bar{x}} = \frac{1}{n}\sum_{i}^{m}\sum_{j}^{n_i}x_{ij} = \frac{1}{n_1+n_2+\ldots+n_m}(\sum_{j}^{n_1}x_{1j} + \sum_{j}^{n_2}x_{2j}+\ldots+\sum_{j}^{n_m}x_{mj})$ แต่ $\bar{\bar{x}}$ เป็นตัวประมาณค่าที่มีอคติของ $\bar{\bar{X}}$ เราจึงไม่นิยมใช้ สำหรับ $\bar{\bar{x}}$ นักศึกษาต้องระมัดระวังเป็นพิเศษอย่านำมาใช้สำหรับประมาณ $\bar{\bar{X}}$ เกี่ยวกับเรื่องนี้นักวิจัยที่ไม่เข้าใจเทคนิคการสำรวจดีพอมักจะสับสนกันและนำมาใช้ปะปนกันกับ $\hat{\bar{\bar{X}}}$ $\bar{\bar{x}}$ หมายถึงค่าเฉลี่ยอันเกิดจากการนำข้อมูลจากหน่วยสำรวจจากทุก psu ตัวอย่างมารวมกันโดยตรงแล้วเฉลี่ยด้วยขนาดตัวอย่าง n ส่วน $\hat{\bar{\bar{X}}} = \frac{\hat{T}}{N}$ เกิดจากการประมาณค่าเป็นลำดับชั้นตามแผน 2CS เริ่มด้วยการประมาณค่าเฉลี่ยต่อ ssu ใน psu ที่ i แล้วจึงหาค่าเฉลี่ยต่อ psu นำค่าเฉลี่ยนี้ไปประมาณยอดรวมแล้วจึงเฉลี่ยต่อ ssu ด้วยการหารตลอดด้วย N

สำหรับการพิสูจน์ว่า $\bar{\bar{x}}$ เป็นตัวประมาณค่าที่มีอคติของ $\bar{\bar{X}}$ หรือ $E(\bar{\bar{x}}) \neq \bar{\bar{X}}$ จะขอเว้นไว้เป็นแบบฝึกหัด

ตัวอย่าง 5.3 จากตัวอย่างที่ 5.1 จงกะประมาณดูว่าโดยถัวเฉลี่ยแล้วในสัปดาห์ที่ 1 ของเดือน เจ้าหน้าที่ของรัฐคนหนึ่ง ๆ ให้การบริการแก่ประชาชนที่มาติดต่อราชการกี่คน สมมุติว่าหน่วยราชการดังกล่าวมีเจ้าหน้าที่ทั้งสิ้น 1200 คน

วิธีทำ จาก $\hat{T} = 3600$ และ $\hat{V}(\hat{T}) = 460{,}990.4$
ดังนั้น

$$\hat{\bar{\bar{X}}} = \hat{T}/N = 3600/1200 = 3$$

นั่นคือ ในระหว่างสัปดาห์แรกของเดือนเจ้าหน้าที่แต่ละคนได้ให้บริการแก่ประชาชนเฉลี่ยรายละ 3 คน

$$\hat{V}(\hat{\bar{\bar{X}}}) = \frac{1}{N^2} V(\hat{T}) = \frac{1}{(1200)^2} \cdot 460{,}990.6 = .3201322;\ s(\hat{\bar{\bar{X}}}) = .566$$

ดังนั้นช่วงเชื่อมั่น 99% ที่คาดว่าค่าจริงของ $\bar{\bar{X}}$ ปรากฏอยู่คือ

$$\begin{aligned}(\hat{\bar{\bar{X}}} + Z_{.005}s(\hat{\bar{\bar{X}}}),\ \hat{\bar{\bar{X}}} + Z_{.995}s(\hat{\bar{\bar{X}}})) &= \{3-(2.58)(.566),\ 3+(2.58)(.566)\} \\ &= (1.54,\ 4.46)\end{aligned}$$

หรือ $1.54 \leqslant \bar{\bar{X}} \leqslant 4.46$

5.2.4 การประมาณค่าสัดส่วนประชากร P

ด้วยเหตุที่ P เป็นกรณีเฉพาะของ $\bar{\bar{X}}$ ดังนั้นเราสามารถประมาณค่า P ได้โดยอาศัยบทแทรก 5.1 ได้ดังนี้

บทแทรก 5.2 เมื่อดำเนินการสุ่มตัวอย่างโดยใช้แผนสำรวจแบบ 2CS แล้วเราสามารถประมาณค่าสัดส่วนประชากร P ได้ดังนี้

ก. $\hat{P} = \frac{1}{N} \frac{M}{m} \sum_{i}^{m} N_i p_i$ เป็นตัวประมาณค่าที่ปราศจากอคติของ

$P = \frac{1}{N} \sum_{i}^{M} \sum_{j}^{N_i} x_{ij}$ โดยที่ $p_i = \bar{x}_i = \frac{1}{n_i} \sum_{j}^{n_i} x_{ij}$ เมื่อ $x_{ij} = 0,1$

ข. $V(\hat{P}) = \frac{1}{N^2} M^2 \frac{M-m}{M} \frac{S_b^2}{m} + \frac{1}{N^2} \frac{M}{m} \sum_{i}^{M} \frac{N_i - n_i}{N_i} \frac{N_i^2 S_i^2}{n_i}$

โดยที่ $S_b^2 = \frac{1}{M-1} \sum_{i}^{M} (N_i P_i - \bar{N}P)^2 ; \bar{N} = N/M$

$S_i^2 = \frac{N_i P_i Q_i}{N_i - 1} , P_i = \frac{1}{N_i} \sum_{j}^{N_i} x_{ij} ; i = 1,2,\ldots,M$

ค. $\hat{V}(\hat{P}) = \frac{1}{N^2} M^2 \frac{M-m}{M} \frac{s_b^2}{m} + \frac{1}{N^2} \frac{M}{m} \sum_{i}^{m} \frac{N_i - n_i}{N_i} \frac{N_i^2 s_i^2}{n_i}$

โดยที่ $s_b^2 = \frac{1}{m-1} \sum_{i}^{m} (N_i p_i - \frac{1}{m} \sum_{i}^{m} N_i p_i)^2$

$s_i^2 = \frac{n_i p_i q_i}{n_i - 1} ; i = 1,2,\ldots,m$

พิสูจน์ สิ่งที่เราต้องพิสูจน์สำหรับการใช้บทแทรกนี้ก็คือพิสูจน์ให้เห็นว่า

$S_b^2 = \frac{1}{M-1} \sum_{i}^{M} (N_i P_i - \bar{N}P_i)^2$ และ $s_b^2 = \frac{1}{m-1} \sum_{i}^{m} (N_i p_i - \frac{1}{m} \sum_{i}^{m} N_i p_i)^2$
เท่านั้น ส่วน S_i^2 และ s_i^2 เป็นสิ่งที่เคยพิสูจน์มาแล้วในบทที่ 2

1. $\because S_b^2 = \frac{1}{M-1} \sum_{i}^{m} (T_i - \bar{T})^2$

พิจารณา T_i จะพบว่า $T_i = \sum_{j}^{N_i} x_{ij} = N_i \cdot \frac{1}{N_i} \sum_{j}^{N_i} x_{ij} = N_i P_i ; i = 1,2,\ldots,M$

และ $\bar{T} = \frac{T}{M} = \frac{1}{M} \sum_{i}^{M} T_i = \frac{1}{M} \cdot N \cdot \frac{1}{N} \sum_{i}^{M} T_i = \frac{N}{M} \cdot \frac{1}{N} \sum_{i}^{M} \sum_{j}^{N_i} x_{ij} = \frac{N}{M} \cdot P = \bar{N}P$ [1.]

ดังนั้น $S_b^2 = \frac{1}{M-1} \sum_{i}^{M} (T_i - \bar{T})^2 = \frac{1}{M-1} \sum_{i}^{M} (N_i P_i - \bar{N}P)^2$

2. $\because s_b^2 = \frac{1}{m-1} \sum_{i}^{m} (\hat{T}_i - \hat{\bar{T}})^2$

พิจารณา $\hat{T}_i$ จะพบว่า $\hat{T}_i = N_i \bar{x}_i = N_i \frac{1}{n_i} \sum_{j}^{n_i} x_{ij} = N_i p_i ; i = 1,2,\ldots, m$

---

1. $\frac{N}{M} = \frac{1}{M} (N_1 + N_2 + \ldots + N_M) = \bar{N}$

และ $\hat{\bar{T}} = \frac{1}{m}\sum_{i}^{m}\hat{T}_i = \frac{1}{m}\sum_{i}^{m}N_ip_i$

ดังนั้น $s_b^2 = \frac{1}{m-1}\sum_{i}^{m}(\hat{T}_i-\hat{\bar{T}})^2 = \frac{1}{m-1}\sum_{i}^{m}(N_ip_i - \frac{1}{m}\sum_{i}^{m}N_ip_i)^2$

สำหรับกรณีเฉพาะต่าง ๆ ของบทแทรก 5.1 และ 5.2 [1] ต่อไปนี้จะขอเว้นไว้เป็นแบบฝึกหัด

1. $m = M$ 2. $n_i = N_i$ 3. $M >> m$

4. $N_i >> n_i$ 5. $M >> m$, $N_i >> n_i$

ตัวอย่าง 5.4 ต้องการกะประมาณสัดส่วนของประชากรในท้องที่แห่งหนึ่งว่ามีความพึงพอใจต่อการบริหารงานของรัฐหรือไม่ โดยถือว่าเอาเฉพาะผู้ที่มีสิทธิเลือกตั้งเป็นหน่วยสำรวจ ในท้องที่ดังกล่าวประกอบไปด้วยผู้ที่มีสิทธิเลือกตั้งทั้งสิ้น 80,000 คน จำแนกออกเป็น 40 หน่วยเลือกตั้ง สุ่มตัวอย่างหน่วยเลือกตั้งมาเป็นตัวอย่างชั้นที่ 1 รวม 8 หน่วย แต่ละหน่วยสุ่มตัวอย่างผู้มีสิทธิออกเสียงเลือกตั้งมาเป็นตัวอย่างหน่วยละ 50 คน แล้วสอบถามความรู้สึกที่มีต่อการบริหารงานของรัฐว่าเป็นอย่างไร สมมุติข้อถามหนึ่งในแบบสอบถามชุดนี้ถามว่า “ท่านมีความรู้สึกอย่างไรเกี่ยวกับพระราชกำหนดภาษีน้ำมันของรัฐ” โดยมีคำตอบคือ (ก) เห็นด้วย (1) (ข) ไม่เห็นด้วย (0)

จากการทอดแบบสำรวจปรากฏผลการสำรวจดังนี้คือ

---

[1] ดูข้อแนะนำท้ายทฤษฎี 5.1

หน่วยเลือกตั้งที่	N_i	n_i	Σx_{ij}
2	1600	50	15
8	2500	50	20
11	2100	50	19
17	1900	50	17
22	2400	50	21
27	1800	50	18
32	1800	50	16
35	2100	50	17

จงกะประมาณสัดส่วนของประชากรที่เห็นด้วยและไม่เห็นด้วยกับการกำหนดพระราชกำหนดภาษีน้ำมันด้วยช่วงเชื่อมั่น 95%

วิธีทำ จากข้อมูลเราสามารถเตรียมตารางวิเคราะห์ได้ดังนี้

$$N = 80{,}000,\ M = 40,\ m = 8,\ n_i = 50\ ;\ i = 1,2,\ldots,8$$

หน่วยเลือกตั้ง	N_i	n_i	Σx_{ij}	p_i	N_ip_i	$\frac{N_i-n_i}{N_i}$	$n_ip_iq_i$	$N_i^2 \frac{N_i-n_i}{N_i} \frac{n_ip_iq_i}{n_i(n_i-1)}$
2	1600	50	15	.30	480	.969	10.50	10631.314
8	2500	50	20	.40	1000	.980	12.00	30000.000
11	2100	50	19	.38	798	.976	11.78	20695.102
17	1900	50	17	.34	646	.974	11.22	16102.485
22	2400	50	21	.42	1008	.979	12.18	28034.083
27	1800	50	18	.36	648	.972	11.52	14808.042
32	1800	50	16	.32	576	.972	10.88	13985.373
35	2100	50	17	.34	714	.976	11.22	19711.295
รวม					5870			153967.680

$$\hat{P} = \frac{1}{N} \cdot \frac{M}{m} \sum_i^m N_ip_i = \frac{40 \times 5870}{80000 \times 8} = .3669$$

$$s_b^2 = \frac{1}{m-1} \sum_i^m (Np_i - \frac{1}{m} \sum_i^m N_ip_i)^2 = \frac{1}{m-1} \left\{ \sum_i^m (N_ip_i)^2 - \frac{(\Sigma N_ip_i)^2}{m} \right\}$$

$$= \frac{1}{8-1} \left\{ 480^2 + 1000^2 + 798^2 + \ldots 714^2 - \frac{(5870)^2}{8} \right\}$$

$$= \frac{1}{7} (4562060 - \frac{34456900}{8})$$

$$= 36421.07$$

ดังนั้น $$\hat{V}(\hat{P}) = \frac{1}{N^2} M^2 \frac{M-m}{M} \frac{s_b^2}{m} + \frac{1}{N^2} \frac{M}{m} \sum_i^m \frac{N_i-n_i}{N_i} N_i^2 \frac{n_ip_iq_i}{n_i(n_i-1)}$$

$$= \frac{1}{80000^2} \cdot 40^2 \cdot \frac{40-8}{40} \cdot \frac{36421.07}{8} + \frac{1}{80000^2} \cdot \frac{40}{8} \cdot 153967.68$$

$= \quad 0.00091 + 0.0001202 = .0010302$

$=> s(\hat{P}) = .0321$

ดังนั้น ช่วงเชื่อมั่น 95% ของ P คือ $\{\hat{P}+Z_{\alpha/2}s(\hat{P}),\ \hat{P}+Z_{1-\alpha/2}s(\hat{P})\}$

$= \{.3669-(1.96)(.0321),\ .3669+(1.96)(.0321)\}$

$= (.3039,\ .4298)$

นั่นคือ

ก. คาดว่าจะมีประชาชนในท้องถิ่นดังกล่าวเห็นด้วยกับพระราชกำหนดภาษีน้ำมันประมาณ 37% หรือระหว่าง 30-43%

ข. คาดว่าจะมีประชาชนในท้องที่ดังกล่าวไม่เห็นด้วยกับพระราชกำหนดภาษีน้ำมันประมาณ 63% หรือระหว่าง 57-70%

5.2.5 การประมาณค่าอัตราส่วนประชากร R

จากตอนที่ 3.4.2 เราจะพบว่า $\hat{T}_x = \hat{R}_c T_y$ เมื่อ $T_y = N\overline{Y}$

ขณะเดียวกัน $\hat{R}_c = \hat{T}_x / T_y$

ดังนั้น $V(\hat{T}_x) = T_y^2 V(\hat{R}_c) = (N\overline{Y})^2 V(\hat{R}_c)$ และ $V(\hat{R}_c) = \frac{1}{T_y^2} V(\hat{T}_x)$

และจากตอน 5.2.4 เราจะพบว่า

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

โดยที่ $S_b^2 = \frac{1}{M-1} \sum_i^M (T_i - \overline{T})^2$ และ $S_i^2 = \frac{1}{N_i - 1} \sum_j^{N_i} (x_{ij} - \overline{X}_i)^2$

โดยอาศัยความจริงทั้งสองประการข้างต้นเราสามารถคำนวณหา $V(\hat{R}_c)$ สำหรับกรณี 2CS ได้ดังนี้

$$V(\hat{R}_{cl}) = \frac{1}{T_y^2} V(\hat{T}) = \frac{1}{T_y^2} \left(M^2 \frac{M-m}{M} \frac{S_{1b}^2}{m} + \frac{M}{m} \sum_i^M \frac{N_i - n_i}{N_i} \frac{N_i^2 S_{2i}^2}{n_i}\right)$$

โดยที่ $\hat{R}_{ct} = \hat{T}_x/\hat{T}_y$ เมื่อ $\hat{T}_x = \frac{M}{m}\sum_{i}^{m}\frac{N_i}{n_i}\sum_{j}^{n_i} x_{ij}$, $\hat{T}_y = \frac{M}{m}\sum_{i}^{M}\frac{N_j}{n_i}\sum_{j}^{n_i} y_{ij}$

$$T_y = N\bar{Y}$$

$$S_{1b}^2 = S_{xb}^2 + R^2S_{yb}^2 - 2R\varrho_b S_{xb}S_{yb}$$

$$S_{2i}^2 = S_{xi}^2 + R^2S_{yi}^2 - 2R\varrho_i S_{xi}S_{yi}$$

$$S_{xb}^2 = \frac{1}{M-1}\sum_{i}^{M}(T_{xi}-\bar{T}_x)^2,\ S_{yb}^2 = \frac{1}{M-1}\sum_{i}^{M}(T_{yi}-\bar{T}_y)^2$$

$$S_{xi}^2 = \frac{1}{N_i-1}\sum_{j}^{N_i}(x_{ij}-\bar{\bar{X}}_i)^2,\ S_{yi}^2 = \frac{1}{N_i-1}\sum_{j}^{N_i}(y_{ij}-\bar{\bar{Y}}_i)^2$$

$$\varrho_b = \frac{1}{M-1}\sum_{i}^{M}\frac{(T_{xi}-\bar{T}_x)(T_{yi}-\bar{T}_y)}{S_{xb}S_{yb}} = \text{Cov ระหว่าง psu}$$

$$\varrho_i = \frac{1}{N_i-1}\sum_{j}^{N_i}\frac{(x_{ij}-\bar{\bar{X}}_i)(y_{ij}-\bar{\bar{Y}}_i)}{S_{xi}S_{yi}} = \text{Cov ระหว่าง ssu ใน psu ที่ i}$$

การพิสูจน์ว่า $S_{1b}^2 = S_{xb}^2 + S_{yb}^2 + -2R\varrho_b S_{xb}S_{yb}$ และ $S_{2i}^2 = S_{xi}^2 + S_{yi}^2 - 2R\varrho_i S_{xi}S_{yi}$ ขอให้ดูวิธีพิสูจน์ 2 ในตอน 3.4.2 และขอเว้นการพิสูจน์ไว้เป็นแบบฝึกหัด อย่างไรก็ตาม ขอแนะนำให้นักศึกษาเข้าใจไว้ในที่นี้ว่า ใน 2CS นั้นความผันแปรย่อมเกิดขึ้น 2 ระยะคือ เมื่อสุ่ม psu คราวหนึ่งกับเมื่อสุ่ม ssu อีกคราวหนึ่ง เมื่อเป็นเช่นนี้ความแปรปรวนจึงมีทั้งความแปรปรวนในระหว่าง psu และในระหว่าง ssu ใน psu ที่ i ดังนั้น $V(\hat{R}_{ct})$ จึงต้องประกอบไปด้วยความผันแปรทั้งสองแหล่งคือ S_{1b}^2 และ S_{2i}^2 การพิสูจน์ความจริงเหล่านี้โดยอาศัยแนวคิดจากตอน 3.4.2 นักศึกษาจึงควรจะคำนึงถึงจุดที่แนะนำนี้ไว้เสมอ

สำหรับตัวประมาณค่าของ $V(\hat{R}_{ct})$ เราจะใช้สูตรดังนี้คือ

$$\hat{V}(\hat{R}_{ct}) = \frac{1}{\hat{T}_y^2}\hat{V}(\hat{T}) = \frac{1}{\hat{T}_y^2}\left(M^2\frac{M-m}{M}\frac{s_{1b}^2}{m} + \frac{M}{m}\sum_{i}^{m}\frac{N_i-n_i}{N_i}\frac{N_i^2 s_{2i}^2}{n_i}\right)$$

โดยที่ $\hat{T}_y = \frac{M}{m}\sum_{i}^{m}\frac{N_i}{n_i}\sum_{j}^{n_i} y_{ij}$

$$s_{1b}^2 = s_{xb}^2 + \hat{R}_{ct}^2 s_{by}^2 - 2\hat{R}\hat{\varrho}_b s_{xb}s_{yb}$$

เมื่อ $s_{xb}^2 = \frac{1}{m-1} \sum_i^m (\hat{T}_{xi} - \bar{\hat{T}}_x)^2$, $s_{yb}^2 = \frac{1}{m-1} \sum_i^m (\hat{T}_{yi} - \bar{\hat{T}}_y)^2$

และ $s_{2i}^2 = s_{xi}^2 + \hat{R}_{c_t} s_{yi}^2 - 2\hat{R}_{c_t}\hat{\varrho}_i S_{xi} S_{yi}$; $i = 1,2,\ldots,m$

เมื่อ $s_{xi}^2 = \frac{1}{n_i - 1} \sum_j^{n_i} (x_{ij} - \bar{\bar{x}}_i)^2$; $s_{yi}^2 = \frac{1}{n_i - 1} \sum_j^{n_i} (y_{ij} - \bar{\bar{y}}_i)^2$

$$\hat{\varrho}_b = \frac{1}{m-1} \sum_i^m \frac{(\hat{T}_{xi} - \bar{\hat{T}}_x)(\hat{T}_{yi} - \bar{\hat{T}}_y)}{s_{xb} s_{yb}}$$

$$\hat{\varrho}_i = \frac{1}{n_i - 1} \sum_j^{n_i} \frac{(x_{ij} - \bar{\bar{x}}_i)(y_{ij} - \bar{\bar{y}}_i)}{s_{xi} s_{yi}} \; ; i = 1,2,\ldots,m$$

ตัวอย่าง 5.5 ต้องการเปรียบเทียบอัตราส่วนของความต้องการซื้อสินค้าประเภทเสื้อผ้าเครื่องแต่งกายระหว่างเดือนเมษายนและพฤษภาคมของประชาชนในท้องที่ตำบลหนึ่ง ซึ่งประกอบด้วยหมู่บ้านต่าง ๆ 6 หมู่บ้าน โดยสุ่มตัวอย่างหมู่บ้านมาเป็นตัวอย่าง 3 หมู่บ้าน แต่ละหมู่บ้านสุ่มร้านค้าในหมู่บ้านมาเป็นตัวอย่าง 50% ของที่มีอยู่ ปรากฏข้อมูลดังนี้

ตำบลตัวอย่าง	n_i	ยอดขายเสื้อผ้า เดือนพฤษภาคม (โหล) : X	ยอดขายเสื้อผ้า เดือนเมษายน (โหล) : Y
1	10	2,4,1,3,1,5,4,4,5,3	3,4,2,4,2,5,4,5,6,5
2	8	5,6,5,7,8,8,9,11	6,8,6,9,11,8,10,13
3	5	11,13,15,12,18	13,12,18,14,24

วิธีทำ

$$\hat{R}_{ct} = \hat{T}_x/\hat{T}_y, \hat{T}_x = \frac{M}{m} \sum_i^m \frac{N_i}{n_i} \sum_j^{n_i} x_{ij} \; ; \; \hat{T}_y = \frac{M}{m} \sum_i^m \frac{N_i}{n_i} \sum_j^{n_i} y_{ij}$$

$$\hat{V}(\hat{R}_{ct}) = \frac{1}{\hat{T}_y^2} \left(M^2 \frac{M-m}{M} \frac{s_{1b}^2}{m} + \frac{M}{m} \sum_i^m \frac{N_i - n_i}{N_i} \frac{N_i^2 s_{2i}^2}{n_i}\right)$$

โดยที่

$$s^2_{1b} = s^2_{xb} + \hat{R}^2_{ct}s^2_{yb} - 2\hat{R}_{ct}\hat{\varrho}_b s_{xb}s_{yb}$$

$$s^2_{xb} = \frac{1}{m-1}\sum_i^m (\hat{T}_{xi} - \hat{\bar{T}}_x)^2 \ , \ s^2_{yb} = \frac{1}{m-1}\sum_i^m (\hat{T}_{yi} - \hat{\bar{T}}_y)^2$$

$$\hat{\varrho}_b = \frac{1}{m-1}\sum_i^m \frac{(\hat{T}_{xi} - \hat{\bar{T}}_x)(\hat{T}_{yi} - \hat{\bar{T}}_y)}{s_{xb}s_{yb}}$$

$$s^2_{2i} = s^2_{xi} + \hat{R}^2_{ct}s^2_{yi} - 2\hat{R}_{ct}\hat{\varrho}_i s_{xi}s_{yi}$$

โดยที่ $s^2_{xi} = \frac{1}{n_i-1}\sum_j^{n_i} (x_{ij} - \bar{\bar{x}}_i)^2 \ , \ s^2_{yi} = \frac{1}{n_i-1}\sum_j^{n_i} (y_{ij} - \bar{\bar{y}}_i)^2$

$$\hat{\varrho}_i = \frac{1}{n_i-1}\sum_j^{n_i} \frac{(x_{ij} - \bar{\bar{x}}_i)(y_{ij} - \bar{\bar{y}}_i)}{s_{xi}s_{yi}} \ ; i = 1,2,\ldots,m$$

เราสามารถเตรียมตารางวิเคราะห์ข้อมูลได้ดังนี้คือ

ตำบล	N_i	n_i	Σx_{ij}	$\bar{\bar{x}}_i$	Σy_{ij}	$\bar{\bar{y}}_i$	$\hat{T}_{xi} = N_i\bar{\bar{x}}_i$	$\hat{T}_{yi} = N_i\bar{\bar{y}}_i$	$\hat{T}_{xi}\hat{T}_{yi}$
1	20	10	32	3.2	40	4	64	80	5120
2	12	8	59	7.38	71	8.875	118	142	16756
3	10	5	69	13.8	81	16.2	138	162	22356
รวม							320	384	44232

$$\hat{\bar{T}}_x = \frac{1}{m}\sum_i^m \hat{T}_{xi} = \frac{1}{3}(320) = 106.67, \ \hat{T}_x = M\hat{\bar{T}}_x = 6(106.67) = 640$$

$$\hat{\bar{T}}_y = \frac{1}{m}\sum_i^m \hat{T}_{yi} = \frac{1}{3}(384) = 128 \ , \ \hat{T}_y = M\hat{\bar{T}}_y = 6(128) = 768$$

$$\hat{R}_{ct} = \frac{\hat{T}_x}{\hat{T}_y} = 640/768 = .833$$

นั่นคือ อัตราส่วนระหว่างความต้องการซื้อสินค้าประเภทเสื้อผ้าเครื่องนุ่งห่มของตำบลดังกล่าวในเดือนพฤษภาคมและเดือนเมษายนเท่ากับ .833 หรือความต้องการซื้อสินค้าประเภทเสื้อผ้าในเดือนพฤษภาคมลดลงจากเดือนเมษายน 16.7%

$$s_{xb}^2 = \frac{1}{m-1} \sum_i^m (\hat{T}_{xi} - \bar{\hat{T}}_x)^2 = \frac{1}{m-1} \left\{ \sum_i^m \hat{T}_{xi}^2 - \frac{(\sum^m \hat{T}_{xi})^2}{m} \right\}$$

$$= \frac{1}{2}\left\{(64^2+118^2+138^2) \frac{320^2}{3} \right\} = 1465.33$$

$$s_{xb} = 38.28$$

$$s_{yb}^2 = \frac{1}{m-1} \sum_i^m (\hat{T}_{yi} - \bar{\hat{T}}_y)^2 = \frac{1}{m-1} \left\{ \sum_i^m \hat{T}_{yi}^2 - \frac{(\sum_i^m \hat{T}_{yi})^2}{m} \right\}$$

$$= \frac{1}{2}\left\{(80^2+142^2+162^2) - \frac{384^2}{3} \right\} = 1828$$

$$s_{yb} = 42.76$$

$$\frac{1}{m-1} \sum_i^m (\hat{T}_{xi} - \bar{\hat{T}}_x)(\hat{T}_{yi} - \bar{\hat{T}}_y) = \frac{1}{m-1} \left(\sum_i^m \hat{T}_{xi} \hat{T}_{yi} - m \bar{\hat{T}}_x \bar{\hat{T}}_y \right)$$

$$= \frac{1}{2}\left\{44232 - 3(106.67)(128)\right\} = 1635.36$$

$$\hat{\varrho}_b = \frac{1635.36}{(38.28)(42.76)} = .999$$

$$s_{1b}^2 = s_{xb}^2 + \hat{R}_{cl}^2 s_{yb}^2 - 2\hat{R}_{cl}\hat{\varrho}_b s_{xb} s_{yb}$$

$$= 1465.33 + (.833)^2(1828) - 2(.833)(.999)(38.28)(42.76)$$

$$= 9.4893$$

$$s_{xi}^2 = \frac{1}{n_i-1} \sum_j^{n_i} (x_{ij} - \bar{\bar{x}}_i)^2 = \frac{1}{n_i-1} \left\{ \sum_j^{n_i} x_{ij}^2 - \frac{(\sum^{n_i} x_{ij}^2)^2}{n_i} \right\}; i = 1,2,\ldots,m$$

$$s_{x1}^2 = \frac{1}{10-1} \left\{(2^2+4^2+\ldots+3^2) - \frac{32^2}{10}\right\} = 2.178,\ s_{x1} = 1.48$$

$$s_{x2}^2 = \frac{1}{8-1} \left\{(5^2+6^2+\ldots+11^2) - \frac{59^2}{8}\right\} = 4.27,\ s_{x2} = 2.07$$

$$s_{x3}^2 = \frac{1}{5-1} \left\{(11^2+13^2+\ldots+18^2) - \frac{69^2}{5}\right\} = 7.7,\ s_{x3} = 2.77$$

$$s_{yi}^2 = \frac{1}{n_i - 1} \sum_j^{n_i} (y_{ij} - \bar{\bar{y}}_i)^2 = \frac{1}{n_i - 1} \left\{ \sum_j^{n_i} y_{ij}^2 - \frac{(\sum^{n_i} y_{ij})^2}{n_i} \right\} ; i = 1,2,\ldots,m$$

ดังนั้น

$$s_{y1}^2 = \frac{1}{9}\left\{(3^2 + 4^2 + \ldots + 5^2) - \frac{40^2}{10}\right\} = 1.78 , s_{y1} = 1.33$$

$$s_{y2}^2 = \frac{1}{7}\left\{(6^2 + 8^2 + \ldots + 13)^2 - \frac{71^2}{8}\right\} = 5.84 , s_{y2} = 2.42$$

$$s_{y3}^2 = \frac{1}{4}\left\{(13^2 + 12^2 + \ldots + 24)^2 - \frac{81^2}{5}\right\} = 24.2, s_{y3} = 4.92$$

$$\varrho_i = \frac{1}{n_i - 1} \sum_j^{n_i} \frac{(x_{ij} - \bar{\bar{x}}_i)(y_{ij} - \bar{\bar{y}}_i)}{s_{xi} s_{yi}}$$

$$= \frac{1}{n_i - 1} \cdot \frac{1}{s_{xi} s_{yi}} \left(\sum_j^{n_i} x_{ij} y_{ij} - n_i \bar{\bar{x}}_i \bar{\bar{y}}_i \right) ; i = 1,2,\ldots,m$$

ดังนั้น

$$\varrho_1 = \frac{1}{9} \cdot \frac{1}{(1.48)(1.33)} \left\{(2 \times 3 + 4 \times 4 + \ldots + 3 \times 5) - (10)(3.2 \times 4)\right\}$$

$$= 0.903$$

$$\varrho_3 = \frac{1}{7} \cdot \frac{1}{(2.07(2.42)} \left\{(5 \times 6 + 6 \times 8 + \ldots + 11 \times 13 - 8(7.375 \times 8.875)\right\}$$

$$= 0.923$$

$$\varrho_3 = \frac{1}{4} \cdot \frac{1}{(2.77)(4.92)} \left\{(11 \times 13 + 13 \times 12 + \ldots + 18 \times 24) - 5(13.8 \times 16.2)\right\}$$

$$= 0.939$$

$$s_{xi}^2 = s_{xi}^2 + \hat{R}_{cl}^2 s_{yi}^2 - 2\hat{R}_{cl} \varrho_i s_{xi} s_{yi} : i = 1,2,3$$

ดังนั้น

$$s^2_{21} = 2.178 + (8.33)^2(1.78) - 2(8.33)(.903)(1.48)(1.33) = .452$$

$$s^2_{22} = 4.27 + (.833)^2(5.84) - 2(.833)(.923)(2.07)(2.42) = .619$$

$$s^2_{23} = 7.7 + (.833)^2(24.2) - 2(.833)(.939)(2.77)(4.92) = 3.172$$

$$V(\hat{R}_{cl}) = \frac{1}{\hat{T}^2_y}\left(M^2 \frac{M-m}{M} \frac{s^2_{1b}}{m} + \frac{M}{m}\sum_{i}^{m} \frac{N_i - n_i}{N_i} \frac{N_i^2 s^2_{2i}}{n_i}\right)$$

$$= \frac{1}{768^2}\left\{6^2 \frac{6-3}{6} \frac{99.4893}{3} + \frac{6}{3}\left\{\frac{20-10}{20} \cdot 20^2 \frac{(.452)}{10}\right.\right.$$

$$\left.\left. + \frac{16-8}{16} \cdot 16^2 \frac{(.619)}{8} + \frac{10-5}{10} \cdot 10^2 \cdot \frac{3.172}{5}\right\}\right\}$$

$$= \frac{1}{768^2}(56.9358 + 18.08 + 19.808 + 63.44)$$

$$= .0002683$$

$$s_{\hat{R}_{cl}} = .0164$$

ดังนั้นช่วงเชื่อมั่น 95% ที่คาดว่าค่าจริงของอัตราส่วน R_{cl} จะปรากฏอยู่คือ

$$.833 - (1.96)(.0164) < R_{cl} < .833 + (1.96)(.0164)$$

$$.800 < R_{cl} < .865$$

5.2.6 การประมาณค่าในกรณีที่ Cluster มีขนาดเดียวกันหรือใกล้เคียงกันและขนาดตัวอย่างที่สุ่มจากแต่ละ Cluster มีขนาดเดียวกัน

ในกรณีทางปฏิบัติหลายกรณีเราอาจพบปัญหาที่สำคัญคือการไม่ทราบขนาดประชากร N ซึ่งในกรณีเช่นนี้จะมีผลให้ไม่อาจคำนวณค่าประมาณของค่าเฉลี่ยและสัดส่วนประชากรได้ ตัวอย่างเช่นในการประมาณค่าเฉลี่ยของครอบครัวเกษตรกรในเขตจังหวัด

ขอนแก่น ถ้าไม่ทราบจำนวนครัวเรือน เกษตรกรในจังหวัดขอนแก่นการประมาณขนาดครอบครัวโดยเฉลี่ยย่อมกระทำมิได้ ทั้งนี้เพราะ $\hat{\bar{X}} = \hat{T}/N$ เราสามารถประมาณยอดรวม T ได้ด้วย $\hat{T}$ แต่ถ้าไม่ทราบค่า N ก็ไม่สามารถหาค่า $\bar{\bar{X}}$ ได้ ในทำนองเดียวกันการประมาณสัดส่วนประชากร P ก็ไม่อาจกระทำได้เพราะ $\hat{P} = \hat{T}/N$; $x_{ij} = 0,1$

ปัญหาดังกล่าวสามารถแก้ไขได้ 2 วิธีคือ

วิธีที่ 1 ถือว่า Cluster มีขนาดเท่าเทียมกัน (Equal Cluster Size)

วิธีที่ 2 ใช้วิธีสำรวจแบบ Probability Proportional to Size (pps) [1]

การถือว่า Cluster มีขนาดเท่ากันนั้น เป็นวิธีปฏิบัติที่ถือว่าเป็นเพียงกรณีเฉพาะของ 2CS และมิได้ใช้ได้เสมอไปถ้าความจริงมิได้เป็นเช่นนั้น กล่าวคือ ถ้า Cluster มีขนาดที่แตกต่างกันมาก เพราะถ้าขนาด Cluster แตกต่างกันมากแต่เรากลับถือว่ามีขนาดเท่าเทียมกันหรือใกล้เคียงกันย่อมเป็นการฝืนความจริง แต่ในหลายกรณีและหลายสถานการณ์เราจะพบว่าขนาด Cluster จะเท่ากันหรือใกล้เคียงกันเสมอ เช่น การสุ่มตัวอย่างนักเรียนจากแต่ละชั้นเรียนที่เป็นตัวอย่าง โดยที่ชั้นเรียนเป็น psu ซึ่งจะพบว่า โดยปกติชั้นเรียนหนึ่ง ๆ มักจะมีขนาดใกล้เคียงกันประมาณ 45 คน หรือทำการสำรวจโดยอาศัยภาพถ่ายทางอากาศซึ่งแผนที่จะแบ่งพื้นที่ออกเป็นตารางขนาดเท่า ๆ กันอาจเป็น Square Grid หรือ Rectangular Grid ก็ได้ ซึ่งเมื่อขยายมาตราส่วนแล้ว 1 ตารางอาจมีพื้นที่เท่ากับ 20 ไร่เท่า ๆ กัน ดังนี้เป็นต้น

สำหรับการสุ่มตัวอย่างจากแต่ละ Cluster เราสามารถสุ่มมาได้ในจำนวนใด ๆ ตามความเหมาะสมแต่โดยปกติโดยทั่วไปควรใช้ Equal Allocation เพราะเมื่อ Cluster มีขนาดใกล้เคียงกันแล้วการสุ่มตัวอย่าง ssu มาเป็นหน่วยตัวอย่าง Cluster ละเท่า ๆ กันย่อมเหมาะสม ถ้าจะใช้ Proportional Allocation ก็ได้เพราะมิได้ให้ผลแตกต่างไปจาก Equal Allocation มากนัก

---

1. จะกล่าวถึงในตอนต่อไป