2. $V(\bar{x}_{eq}) = \frac{1}{N^2} \sum_h^L \frac{N_h - n_h}{N_h} \frac{N_h^2 S_h^2}{n_h}$ โดยที่ $n_h = n/L$

หรือเมื่อแทนที่ n_h ด้วย n/L จะพบสูตรเฉพาะตัวของ $V(\bar{x}_{eq})$ ดังนี้

$$V(\bar{x}_{eq}) = \frac{1}{N^2} \sum_h^L \frac{N_h - n/L}{N_h} \frac{N_h^2 S_h^2}{n_h}$$

$$= \frac{1}{N^2} \sum_h^L \frac{LN_h - n}{LN_h} \frac{N_h^2 S_h^2}{n/L}$$

$$= \frac{1}{N^2} \sum_h^L \frac{LN_h - n}{N_h} \frac{N_h^2 S_h^2}{n}$$

3. $\hat{V}(\bar{x}_{eq}) = \frac{1}{N^2} \sum_h^L \frac{LN_h - n}{N_h} \frac{N_h^2 s_h^2}{n_h}$

4. $\hat{T}_{eq} = \sum_h^L N_h \bar{x}_h$, เมื่อ $\bar{x}_h = \frac{1}{n/L} \sum_i^{n/L} x_{hi}$

5. $V(\hat{T}_{eq}) = \sum_h^L \frac{LN_h - n}{N_h} \frac{N_h^2 S_h^2}{n}$

6. $\hat{V}(\hat{T}_{eq}) = \sum_h^L \frac{LN_h - n}{N_h} \frac{N_h^2 s_h^2}{n}$

ในกรณีที่ขนาดตัวอย่างที่จัดสรรให้แก่ชั้นภูมิใดมีขนาดใหญ่กว่าขนาดของชั้นภูมิ เช่น ในชั้นภูมิที่ s พบว่า $n_s > N_s$ กรณีเช่นนี้ให้แก้ปัญหาการสุ่มตัวอย่างด้วยการสำรวจทุกหน่วยในชั้นภูมิที่ s แล้วแบ่งขนาดตัวอย่างที่ยังเกินอยู่ $n_s - N_s$ หน่วยให้แก่ชั้นภูมิอื่น ๆ ที่เหลืออย่างเท่าเทียมกัน คือ $\frac{n_s - N_s}{L-1}$ หน่วย ด้วยเหตุนี้ชั้นภูมิอื่น ๆ เว้นชั้นภูมิที่ s จะได้รับการจัดสรรจำนวนตัวอย่างได้เท่า ๆ กันเท่ากับ $n/L + (n_s - N_s)/(L-1)$ หน่วย ส่วนในกรณีที่เกิดเหตุการณ์ทำนองนี้แก่หลายชั้นภูมิ ก็ให้แก้ปัญหาได้ในทำนองเดียวกัน ขอให้สังเกตว่าถ้ามีปัญหาเช่นนี้เกิดขึ้น ขนาดตัวอย่างจากแต่ละชั้นภูมิจะไม่เท่ากันตลอดคือ n_h จะไม่เท่ากับ n/L เสมอไป อาจสูงกว่าหรือต่ำกว่า n/L ซึ่งมีผลให้เราไม่สามารถใช้สูตร 1-6 ข้างต้นวิเคราะห์ได้ วิธีปฏิบัติคือให้ใช้สูตรแม่บทตามทฤษฎี 3.1 และบทแทรก 3.1

และขอย้ำไว้ในที่นี้ว่าโดยปกติเราจะใช้หรือจดจำเฉพาะสูตรแม่บทเท่านั้นเพราะสามารถใช้ได้ตลอดเวลาไม่ว่าจะจัดสรรขนาดตัวอย่างด้วยวิธีใด เพราะการจัดสรรตัวอย่างจะทำให้ทราบค่าของ n_h แตกต่างกันไปตามวิธีการที่แตกต่างกัน เมื่อนำค่า n_h ที่ได้โดยวิธีใดไปแทนค่า n_h ในสูตรแม่บท ผลลัพธ์ก็จะเป็นค่าประมาณของแผนการจัดสรรนั้น

ตัวอย่างเช่น กลุ่มประชากรกลุ่มหนึ่งสามารถจำแนกเป็นชั้นภูมิต่าง ๆ ได้ 3 ชั้น $(L=3)$ ซึ่งมีขนาด $N_1=50, N_2=170,$ และ $N_3=300$ ตามลำดับ กำหนดขนาดตัวอย่าง $n=180$ เมื่อใช้วิธีจัดสรรอย่างเท่าเทียมกันจะพบว่า $n_h = \frac{n}{3}$ หรือ $n_1=n_2=n_3=180/3=60$ ในกรณีนี้จะเห็นได้ว่า $n_1 > N_1$ วิธีปฏิบัติสำหรับแก้ปัญหานี้ก็คือให้สำรวจทุกหน่วยในชั้นภูมิที่ 1 ซึ่งเมื่อกระทำเช่นนี้แล้วจะยังเหลือจำนวนตัวอย่างอีก $n_1-N_1 = 60-50 = 10$ หน่วย ให้แบ่งจำนวนตัวอย่างที่เหลือ 10 หน่วยนี้แก่ชั้นภูมิที่ 2 และ 3 เท่า ๆ กันเท่ากับ $\frac{10}{2} = 5$ หน่วย ดังนั้นขนาดตัวอย่างในขั้นสุดท้ายก็คือ $n_1 = 50$, $n_2 = 65$ และ $n_3 = 65$

อย่างไรก็ตามปัญหาเช่นนี้จะไม่เกิดขึ้นบ่อยนักเว้นแต่นักวิจัยมุ่งความสะดวกรวดเร็วเป็นประมาณ เพราะถ้าขนาดชั้นภูมิแตกต่างกันมากเราไม่นิยมใช้การจัดสรรตามวิธีนี้

ข. การจัดสรรตัวอย่างตามสัดส่วนของชั้นภูมิ (Proportional Allocation)

ในกรณีที่ชั้นภูมิมีขนาดแตกต่างกันและไม่อาจถือได้ว่าเท่ากันได้โดยประมาณ ซึ่งเป็นกรณีที่ปรากฏขึ้นเป็นปกติ ในกรณีเช่นนี้ให้ยึดถือเอาขนาดของชั้นภูมิเป็นเกณฑ์กำหนดน้ำหนัก หมายความว่าชั้นภูมิใดมีขนาดใหญ่กว่าก็ควรจัดสรรจำนวนตัวอย่างให้มากกว่า ชั้นภูมิใดมีขนาดเล็กกว่าก็จัดสรรตัวอย่างให้น้อยกว่าโดยคิดเทียบตามหลักบัญญัติไตรยางค์ หรืออัตราส่วนดังนี้

สมมุติกำหนดขนาดตัวอย่างไว้เท่ากับ n และประชากรมีขนาด N โดยที่ประชากรขนาด N จำแนกเป็นชั้นภูมิต่าง ๆ L ชั้นภูมิมีขนาดเท่ากับ N_h ; $h = 1,2,\ldots,L$

เมื่อพิจารณาเฉพาะชั้นภูมิที่ h ใด ๆ จะพบว่าขนาดตัวอย่างที่ควรจัดสรรให้แก่ชั้นภูมิที่ h จะมีค่าเท่ากับ $\frac{n}{N} \cdot N_h ; h=1,2,\ldots,L$ ซึ่งมีวิธีคำนวณได้ง่าย ๆ ดังนี้คือ

หน่วยสำรวจจากกลุ่มประชากรทั้งสิ้น N หน่วยได้รับเลือกมาเป็นตัวอย่างทั้งสิ้น

$$= n \text{ หน่วย}$$

หน่วยสำรวจจากกลุ่มประชากรทั้งสิ้น 1 หน่วยได้รับเลือกมาเป็นตัวอย่างทั้งสิ้น

$$= \frac{n}{N} \text{ หน่วย}$$

หน่วยสำรวจจากกลุ่มประชากรทั้งสิ้น N_h หน่วยได้รับเลือกมาเป็นตัวอย่างทั้งสิ้น

$$= \frac{n}{N} \cdot N_h \text{ หน่วย}$$

นั่นก็คือขนาดตัวอย่างที่พึงจัดสรรให้แก่ชั้นภูมิที่ h จะมีค่าเท่ากับ $\frac{n}{N} \cdot N_h$ หรือ $n_h = \frac{n}{N} \cdot N_h$; $h = 1,2,\ldots,L$ เราเรียก $\frac{n}{N}$ ว่าอัตราการสุ่มตัวอย่าง (Sampling Fraction) ซึ่งอัตรานี้จะคงที่เสมอ เมื่อนำอัตราดังกล่าวไปคูณกับขนาดของชั้นภูมิ ผลลัพธ์ก็คือขนาดตัวอย่างสำหรับชั้นภูมินั้น ๆ ผลที่น่าสังเกตอีกประการหนึ่งก็คือ การกระทำเช่นนี้จะมีผลให้อัตราการสุ่มตัวอย่างจากแต่ละชั้นภูมิมีค่าคงที่เท่ากันเท่ากับ $\frac{n}{N}$ เสมอ[1]

ตามตัวอย่างในตอน ก. ที่ผ่านมา กำหนดให้ $N_1 = 50$, $N_2 = 170$, $N_3 = 300$, $n = 180$ ซึ่งจะเห็นได้ว่า $N = N_1 + N_2 + N_3 = 50 + 170 + 300 = 520$

ดังนั้น อัตราการสุ่มตัวอย่าง $\frac{180}{520} = .346 = 34.6\%$ หรือประมาณ 35%

นั่นก็คือเราควรสุ่มตัวอย่างจากแต่ละชั้นภูมิประมาณ 35% ของขนาดชั้นภูมินั้น ๆ[2] ดังนั้น

$$n_1 = (.346)(50) = 17.3 \simeq 17 \text{ หน่วย}$$

$$n_2 = (.346)(170) = 58.8 \simeq 59 \text{ หน่วย}$$

$$n_3 = (.346)(300) = 103.8 \simeq 104 \text{ หน่วย}$$

---

1 จาก $n_h = \frac{n}{N} \cdot N_h$; $h = 1,2,\ldots,L$ ดังนั้น $\frac{n_h}{N_h} = \frac{n}{N}$; $h = 1,2,\ldots,L$

1 ความจริงข้อนี้ชี้ให้เห็นอย่างชัดเจนว่าชั้นภูมิใดมีขนาดใหญ่กว่าก็จะได้รับการจัดสรรตัวอย่างให้มากกว่า ชั้นภูมิใดมีขนาดเล็กกว่าจะได้รับการจัดสรรตัวอย่างให้น้อยกว่า

นักศึกษาจะสังเกตได้ว่า $\frac{n_h}{N_h} = \frac{n}{N}$ เสมอ[1] ตามตัวอย่าง $\frac{n_1}{N_1} = \frac{17}{50} = .34$ $\frac{n_2}{N_2} = \frac{59}{170} = .347$ และ $\frac{n_3}{N_3} = \frac{104}{300} = .346$

การจัดสรรตามสัดส่วนของขนาดชั้นภูมิเป็นวิธีที่นิยมใช้ในทางปฏิบัติเพราะนอกจากจะสะดวกรวดเร็วและง่ายต่อการชี้แจงหรือทำความเข้าใจแล้วยังสมเหตุสมผลและยุติธรรม นอกจากนี้ยังให้ค่าประมาณที่แม่นยำแม้จะไม่สูงเท่าอีก 2 วิธีซึ่งจะกล่าวถึงต่อไปในตอน ค และ ง ซึ่งรอบคอบรัดกุมกว่าเพราะพิจารณาถึงค่าใช้จ่ายในการสำรวจด้วยก็ตาม[2] อย่างไรก็ตามแผนสำรวจแบบแบ่งชั้นภูมิที่ยึดถือการจัดสรรตัวอย่างตามวิธีนี้จะมีความแม่นยำสูงมากถ้าใช้กลุ่มตัวอย่าง n ที่ใหญ่พอ

การวิเคราะห์และประมาณค่าพารามิเตอร์ต่าง ๆ สำหรับแผนสำรวจแบบแบ่งชั้นภูมิที่จัดสรรตัวอย่างตามวิธีนี้สามารถกระทำได้โดยง่ายเพียงแต่แทนค่า n_h ในสูตรแม่บท (ทฤษฎี 3.1 และบทแทรก 3.1) ด้วย $n_h = \frac{n}{N} \cdot N_h ; h=1,2,\ldots,L$ หรือจะใช้สูตรเฉพาะแผนก็ได้ซึ่งสามารถพัฒนาได้โดยง่ายดังนี้คือ

1. $\hat{\bar{X}} = \bar{x}_{prop} = \frac{1}{N} \sum_{h}^{L} N_h \bar{x}_h$ เมื่อ $\bar{x}_h = \frac{1}{n_h} \sum_{i}^{n_h} x_{hi}$ โดยที่ $n_h = \frac{n}{N} \cdot N_h$

$h = 1,2,\ldots,L$

หรือ $\bar{x}_{prop} = \frac{1}{N} \sum_{h}^{L} N_h \bar{x}_h$ เมื่อ $\bar{x}_h = \frac{N}{nN_h} \sum_{i}^{nN_h/N} x_{hi}$ $; h=1,2,\ldots,L$

---

1 $\frac{n_h}{N_h} \simeq \frac{n}{N}$ ในกรณีปัดเศษ การปัดเศษอาจมีผลให้ $\sum_{h}^{L} n_h \neq n$ ดังนั้นอาจจำเป็นต้องปรับค่า n_h เพื่อให้ $\sum_{h}^{L} n_h = n$

2 แม้จะมีความแม่นยำไม่สูงเท่ากับวิธี Optimum Allocation และ Neyman Allocation แต่ Proportional Allocation มีจุดเด่นที่เหนือกว่าตรงที่ง่ายกว่าและสะดวกรวดเร็วกว่าทั้งในด้านการวางแผนและวิเคราะห์ข้อมูล

$$=> \bar{x}_{prop} = \frac{1}{n} \sum_{h}^{L} \sum_{i}^{nN_h/N} x_{hi}$$

2. $V(\bar{x}_{prop}) = \frac{1}{N^2} \sum_{h}^{L} \frac{N_h - n_h}{N_h} \frac{N_h^2 S_h^2}{n_h}$ โดยที่ $n_h = \frac{n}{N} \cdot N_h;\ h = 1,2,\ldots,L$

หรือถ้าแทนที่ n_h ด้วย $\frac{n}{N} \cdot N_h$ จะพบสูตรเฉพาะของกรณีการจัดสรรตามสัดส่วนของชั้นภูมิดังนี้

$$V(\bar{x}_{prop}) = \frac{1}{N^2} \sum_{h}^{L} \frac{N_h - nN_h/N}{N_h} \frac{N_h^2 S_h^2}{nN_h/N}$$

$$= \frac{1}{N} \sum_{h}^{L} \frac{N-n}{N} \frac{N_h S_h^2}{n}$$

3. $\hat{V}(\bar{x}_{prop}) = \frac{1}{N^2} \sum_{h}^{L} \frac{N_h - n_h}{N_h} \frac{N_h^2 s_h^2}{n_h}$ โดยที่ $n_h = \frac{n}{N} \cdot N_h;\ h = 1,2,\ldots,L$

หรือ $\hat{V}(\bar{x}_{prop}) = \frac{1}{N} \sum_{h}^{L} \frac{N-n}{N} \frac{N_h s_h^2}{n}$

4. $\hat{T}_{prop} = \sum_{h}^{L} N_h \bar{x}_h$ เมื่อ $\bar{x}_h = \frac{1}{nN_h/N} \sum_{i}^{nN/N_h} x_{hi}$

5. $V(\hat{T}_{prop}) = \sum_{h}^{L} \frac{N_h - n_h}{N_h} \frac{N_h^2 S_h^2}{n_h}$ เมื่อ $n_h = \frac{n}{N} \cdot N_h;\ h = 1,2,\ldots,L$

หรือเมื่อแทนค่า n_h ด้วย $\frac{n}{N} \cdot N_h$ จะพบว่า

$$V(\hat{T}_{prop}) = \sum_{h}^{L} \frac{N_h - nN_h/N}{N_h} \frac{N_h^2 S_h^2}{nN_h/N}$$

$$= N \sum_{n}^{L} \frac{N-n}{N} \frac{N_h S_h^2}{n}$$

6. $$\hat{V}(\hat{T}_{prop}) = \sum_{h}^{L} \frac{N_h - n_h}{N_h} \frac{N_h^2 s_h^2}{n_h} \quad \text{เมื่อ } n_h = \frac{n}{N} \cdot N_h;\ h = 1,2,\ldots,L$$

หรือ $$\hat{V}(\hat{T}_{prop}) = N \sum_{h}^{L} \frac{N-n}{N} \frac{N_h s_h^2}{n}$$

ทั้งนี้ $$S_h^2 = \frac{1}{N_h - 1} \sum_{i}^{N_h} (x_{hi} - \bar{X}_h)^2;\ h = 1,2,\ldots,L$$

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i}^{n_h} (x_{hi} - \bar{x}_h)^2 \quad \text{เมื่อ } n_h = \frac{n}{N} \cdot N_h;\ h = 1,2,\ldots,L$$

หรือ $$s_h^2 = \frac{1}{nN_h/N - 1} \sum_{i}^{nN_h/N} (x_{hi} - \bar{x}_h)^2$$

ข้อแนะนำ ในทางปฏิบัติไม่ควรใช้สูตรเฉพาะกรณีเพราะอาจทำให้สับสนได้ ควรจะใช้สูตรแม่บทตามทฤษฎี 3.1 และบทแทรก 3.1 เพียงแต่แทนที่ n_h ด้วย $\frac{n}{N} \cdot N_h$

ตัวอย่าง 3.2 ในการกะประมาณยอดขายและยอดขายถัวเฉลี่ยของห้างสรรพสินค้าต่อวันของเมืองหนึ่ง ซึ่งมีห้างสรรพสินค้าทั้งสิ้น 500 แห่ง ซึ่งเราสามารถจำแนกออกได้เป็น 3 ชั้นภูมิตามขนาด คือ ห้างขนาดใหญ่ 50 แห่ง ขนาดกลาง 150 แห่งและขนาดเล็ก 300 แห่งตามลำดับ และจากการสำรวจเบื้องต้นพบว่า $s_1^2 = 16, s_2^2 = 4, s_3^2 = 1$ สมมุติว่า $\bar{x}_1 = 24, \bar{x}_2 = 12, \bar{x}_3 = 4$ และกำหนดขนาดตัวอย่างให้ $n = 40$

ก. จงจัดสรรตัวอย่างโดยวิธีการจัดสรรอย่างเท่าเทียมกันแล้วกะประมาณยอดขายรวม และยอดขายเฉลี่ยต่อวัน พร้อมทั้งประมาณค่าดังกล่าวด้วยช่วงเชื่อมั่น 95%

ข. จงจัดสรรตัวอย่างโดยวิธีจัดสรรตามสัดส่วนของชั้นภูมิแล้วกะประมาณยอดขายรวมและยอดขายเฉลี่ยต่อวัน พร้อมทั้งประมาณค่าดังกล่าวด้วยช่วงเชื่อมั่น 95%

หมายเหตุ โดยปกติเราจะทราบ $\bar{x}_h$ ภายหลังจากการจัดสรรตัวอย่างและใช้ขนาดตัวอย่างนั้น ๆ ไปสุ่มตัวอย่างมาจากแต่ละชั้นภูมิแล้ว ในที่นี้กำหนดค่า $\bar{x}_h$ มาให้เพื่อความสะดวกในการแสดงให้เห็นการใช้สูตรต่าง ๆ เท่านั้น

วิธีทำ ก. เนื่องจาก $n = 40$ และ $L = 3$ ดังนั้น $n_h = n/L = 40/3 \cong 13; h=1,2,3$ แต่จะเห็นว่า $n_1+n_2n_3 \neq 40$ เราจึงควรจัดสรรตัวอย่างให้แก่ชั้นภูมิใดชั้นภูมิหนึ่ง 14 หน่วย ชั้นภูมิอื่น ๆ จัดสรรให้ 13 หน่วย สมมุติให้ $n_3=14$ ดังนั้น $n_1=13$, $n_2=13$, $n_3=14$

ดังนั้น เราสามารถเตรียมตารางการวิเคราะห์ข้อมูลได้ดังนี้

ชั้นภูมิ	N_h	n_h	s_h^2	$\bar{x}_h$	$N_h\bar{x}_h$	$(N_h-n_h)/N_h$	$N_h^2s_h^2/n_h$	$(N_h-n_h)/N_h.(N_h^2s_h^2/n_h)$
(ห้างสรรพสินค้า)								
ขนาดใหญ่	50	13	16	24	1200	.740	3076.923	2276.92
ขนาดกลาง	150	13	4	12	1800	.913	6923.076	6320.77
ขนาดเล็ก	300	14	1	4	1200	.953	6428.571	6126.43
รวม	500	40			4200			14,724.12

ดังนั้น
$$\bar{x}_{eq} = \frac{1}{N}\sum_{h}^{3} N_h\bar{x}_h = \frac{4200}{500} = 8.4$$

$$\hat{V}(\bar{x}_{eq}) = \frac{1}{N^2}\sum_{h}^{L}\frac{N_h-n_h}{N_h}\frac{N_h^2s_h^2}{n_h} = \frac{1}{(500)^2}\,14,724.12 = 0.059$$

ช่วงเชื่อมั่น 95% ของ $\bar{X}$ คือ

$$\{8.4 - (1.96)(\sqrt{.059}),\ 8.4 + (1.96)(\sqrt{.059})\} = (7.923, 8.876)$$

นั่นคือ ในวันหนึ่ง ๆ ห้างสรรพสินค้าในเมืองดังกล่าวจะมียอดขายเฉลี่ยแห่งละประมาณ 8.4 หน่วย หรือสามารถเชื่อถือได้ถึง 95% ว่าในวันหนึ่ง ๆ ห้างสรรพสินค้าแต่ละแห่งจะมียอดขายเฉลี่ยอยู่ในช่วงระหว่าง 7.923 ถึง 8.876 หน่วย

สำหรับ การประมาณยอดรวมจะขอเว้นไว้เป็นแบบฝึกหัด

ข. เนื่องจาก $n = 40$ และ $\frac{n}{N} = \frac{40}{500} = .08 = 8\%$ หรือ $n_h = (.08)\,N_h$

ดังนั้น $n_1 = (.08)(50) = 4$, $n_2 = (.08)(150) = 12$, $n_3 = (.08)(300) = 24$

เตรียมตารางวิเคราะห์ข้อมูลได้ดังนี้

ขั้นภูมิ (ห้างสรรพสินค้า)	N_h	n_h	s_h^2	$\bar{x}_h$	$N_h\bar{x}_h$	$(N_h - n_h)/N_h$	$N_h^2 s_h^2/n_h$	$(N_h - n_h)/N_h \cdot (N_h^2 s_h^2/n_h)$
ขนาดใหญ่	50	4	16	24	1200	.92	10000	9200
ขนาดกลาง	150	12	4	12	1800	.92	7500	6900
ขนาดเล็ก	300	24	1	4	1200	.92	3750	3450
รวม	500				4200			19,550

ดังนั้น $\bar{x}_{prop} = 4200/500 = 8.4$

$$\hat{V}(\bar{x}_{prop}) = \frac{1}{N^2} \sum_{h}^{L} \frac{N_h - n_h}{N_h} \frac{N_h^2 s_h^2}{n_h} = .078$$

ช่วงเชื่อมั่น 95% ของ $\bar{X}$ คือ

$$\{8.4 - (1.96)(\sqrt{.078}),\ 8.4 + (1.96)(\sqrt{.078})\} = (7.85,\ 8.95)$$

สำหรับกรณีการประมาณยอดขายรวมจะขอเว้นไว้เป็นแบบฝึกหัด

ข้อสังเกต การวิเคราะห์ข้อมูลในตัวอย่างนี้วิเคราะห์โดยอาศัยสูตรแม่บท เพียงแต่แทนค่า n_h ด้วย $\frac{n}{L}$ หรือ $\frac{n}{N} \cdot N_h$ แล้วแต่กรณีเท่านั้น มิได้ใช้สูตรเฉพาะกรณีของการจัดสรร

ค. การจัดสรรแบบอุตมะ (Optimum Allocation)

การจัดสรรแบบอุตมะ คือ การจัดสรรขนาดตัวอย่างให้แก่แต่ละชั้นภูมิในลักษณะที่การจัดสรรดังกล่าวต้องคำนึงถึงค่าใช้จ่ายในการสำรวจต่อหน่วยในแต่ละชั้นภูมิด้วย เพราะค่าใช้จ่ายในการสำรวจนั้นโดยปกติจะแตกต่างกันไปตามชั้นภูมิ เช่น ความยากง่ายในการเข้าถึงหน่วยสำรวจ ระยะทางจากสำนักงานกลางถึงชั้นภูมิหรือตำบลที่ตั้งของหน่วยสำรวจ ระดับค่าครองชีพในตำบลที่ตั้งของหน่วยสำรวจ และอื่น ๆ สิ่งเหล่านี้ย่อมส่งผลสะท้อนต่อค่าจ้างพนักงานสำรวจ เพราะชั้นภูมิใดเป็นท้องที่ที่ห่างไกลจากสำนักงานกลางมากย่อมเสียค่าใช้จ่ายในการเดินทางสูง การคมนาคมที่ไม่สะดวกทางกันดารต้องใช้เวลาเดินทางนานนับชั่วโมงจึงจะเข้าถึงหน่วยสำรวจได้ มีผลให้พนักงานสำรวจท้อถอยและไม่ปรารถนาจะเข้าสำรวจในชั้นภูมินั้นหรือบางท้องที่มีระดับค่าครองชีพสูงมาก พนักงานสำรวจได้รับค่าตอบแทนไม่คุ้มกับค่าใช้จ่าย

ปัจจัยในหลาย ๆ ประการดังกล่าวข้างต้นจึงเป็นเครื่องชี้ให้เห็นว่าการไม่คำนึงถึงค่าใช้จ่ายในการสำรวจย่อมส่งผลสะท้อนต่องานและคุณภาพของงานได้ไม่ทางตรงก็ทางอ้อม ผลสะท้อนทางตรงก็คือ งานอาจไม่ลุล่วงตามเป้าหมายไปตามแผน ทั้งนี้เพราะโดยปกติงานสำรวจหรืองานวิจัยโครงการหนึ่ง ๆ จะได้รับงบประมาณสนับสนุนเป็นจำนวนที่จำกัดพอดี นักวิจัยต้องวางแผนสำรวจและกำหนดขนาดตัวอย่างให้สอดคล้องกับงบประมาณที่มีอยู่ ถ้ามีงบประมาณน้อยการกำหนดขนาดตัวอย่างให้มีขนาดใหญ่เกินไปย่อมเป็นสิ่งที่เป็นไปไม่ได้ และอาจทำให้งานล้มเหลวกลางคัน นอกจากนี้งบประมาณที่มีอยู่มิใช่งบประมาณที่จะนำมาใช้เพื่อการสำรวจได้ทั้งหมด นักวิจัยต้องกันงบประมาณส่วนหนึ่งไว้ใช้สำหรับเป็นค่าใช้จ่ายประจำบางหมวด (fixed cost) เช่น ค่าจ้างเจ้าหน้าที่ธุรการ ค่าเช่าสถานที่หรือค่าใช้จ่าย

สำนักงานและบริการสาธารณูปโภค ค่าใช้จ่ายเพื่อการประมวลผลและเสนอผลและอื่น ๆ ซึ่งมีผลให้งบประมาณเพื่อการสำรวจมีจำนวนไม่มากนัก ถ้านักวิจัยไม่คำนึงถึงประเด็นนี้ อีกทั้งยังไม่คำนึงถึงค่าใช้จ่ายที่แตกต่างกันในการสำรวจแต่ละชั้นภูมิด้วยแล้วผลเสียย่อมกระทบถึงงานโดยตรง เช่น ไม่อาจดำเนินงานต่อไปได้ หรือต้องเปลี่ยนแผนกลางคันทำให้ระดับความแม่นยำและน่าเชื่อถือผิดไปจากเดิม เป็นต้น ส่วนผลในทางอ้อมก็คือผลอันเนื่องมาจากพนักงานสำรวจ ความยากง่ายในการเข้าถึงหน่วยสำรวจและความสิ้นเปลืองค่าใช้จ่ายในการเดินทาง และปัจจัยการยังชีพซึ่งควรมีผลให้ได้รับค่าตอบแทนที่แตกต่างกัน ถ้ากำหนดค่าตอบแทนไม่เหมาะสม ย่อมมีผลในด้านกำลังใจของพนักงานสำรวจซึ่งสะท้อนไปสู่คุณภาพของข้อมูล ซึ่งจะย้อนไปมีผลกระทบต่อความแม่นยำของงานในที่สุด

ดังนั้น แผนการจัดสรรตัวอย่างที่ดีจึงควรเป็นแผนการณ์ที่นอกจากจะสอดคล้องกับงบประมาณที่มีอยู่แล้วยังจะต้องเป็นแผนการณ์ที่มีผลให้งานมีความแม่นยำสูงสุด หรือ $V(\hat{\theta})$ มีค่าต่ำสุดด้วย หรือนัยหนึ่ง ต้องเป็นแผนการณ์ที่ทำให้ $V(\hat{\theta})$ มีค่าต่ำสุดภายใต้วงงบประมาณที่มีอยู่อย่างจำกัด ขนาดตัวอย่างที่จัดสรรให้แก่แต่ละชั้นภูมิตามปรัชญานี้เรียกว่า การจัดสรรแบบอุดมะ (Optimum Allocation)

ฟังก์ชันค่าใช้จ่ายเพื่อการวิจัย (Cost Function) นิยามดังนี้

$$
\begin{aligned}
C &= c_0 + n_1c_1 + n_2c_2 + \ldots + n_Lc_L \\
&= c_0 + \sum_{h}^{L} n_hc_h
\end{aligned}
$$

โดยที่ c_0 คือ ค่าใช้จ่ายประจำหรือค่าใช้จ่ายคงที่ (fixed cost) เช่น ค่าใช้จ่ายในการดำเนินการโครงการ ค่าวัสดุอุปกรณ์ เครื่องมือ ค่าใช้จ่ายสำนักงาน ค่าใช้จ่ายในการวางแผนสำรวจ ค่าใช้จ่ายในการสำรวจกรอบตัวอย่าง ปรับปรุงกรอบตัวอย่าง หรือสร้างกรอบตัวอย่าง ค่าใช้จ่ายในการฝึกอบรมพนักงานสำรวจ ค่าใช้จ่ายในการวิเคราะห์ประมวลผล เช่น เวลาเช่าคอมพิวเตอร์หรืออื่น ๆ ที่ไม่ผันแปรไปตามขนาดของตัวอย่าง

c_h คือ ค่าใช้จ่ายต่อหน่วยในการสำรวจชั้นภูมิที่ h หรือค่าใช้จ่ายผันแปร (variable cost) ค่าใช้จ่ายส่วนนี้จะแตกต่างกันไปตามธรรมชาติของชั้นภูมิ เช่น ความยากง่ายในการ

เข้าถึง ความห่างไกลจากสำนักงานกลาง (สำนักวิจัยหรือสถานที่วิจัย) ความทุรกันดาร ระดับการครองชีพ และอื่น ๆ เช่น กำหนดให้ $c_1 = 10$ บาท $c_2 = 18$ บาท $c_3 = 40$ บาท เป็นต้น $c_1 = 10$ บาท หมายความว่าค่าใช้จ่ายในการสำรวจหน่วยสำรวจในชั้นภูมิที่ 1 เท่ากับ 10 บาท ต่อ 1 หน่วย $c_2 = 18$ บาท หมายความว่าค่าใช้จ่ายในการสำรวจหน่วยสำรวจในชั้นภูมิที่ 2 เท่ากับ 18 บาท ต่อ 1 หน่วย $c_3 = 40$ บาท หมายความว่าค่าใช้จ่ายในการสำรวจในชั้นภูมิที่ 3 เท่ากับ 40 บาท ต่อ 1 หน่วย เป็นต้น ค่าใช้จ่ายต่อหน่วยอาจหมายถึงเฉพาะค่าจ้างพนักงานสำรวจ หรือผนวกค่าใช้จ่ายอื่น ๆ เข้าไปด้วย เช่น ค่าแบบสอบถามต่อ 1 ชุด ค่าใช้จ่ายในการประมวลผลซึ่งอาจรวมค่าบัตร ค่าจ้างเจาะบัตร และการตรวจสอบความถูกต้อง ค่าเช่าที่พักและอาหาร ค่าขนส่ง ตลอดจนค่าจ้างผู้ควบคุมงานสนาม งานวิจัยเอาไว้ด้วย แต่ถ้าจะให้การวิเคราะห์ค่าใช้จ่ายเป็นไปโดยง่าย ควรจำแนกค่าใช้จ่ายบางรายการที่ผนวกกันอยู่นี้ออกเป็นค่าใช้จ่ายคงที่ และถ้าเป็นไปได้ควรจัดให้ c_h หมายถึง เฉพาะค่าจ้างพนักงานสำรวจและค่าใช้จ่ายผันแปรในรายการที่สำคัญจริง ๆ โดยผลักค่าใช้จ่ายอื่น ๆ ไปเป็นค่าใช้จ่ายคงที่

n_hc_h คือ ค่าใช้จ่ายรวมจากการสำรวจในชั้นภูมิที่ h หรือค่าใช้จ่ายรวมในการสำรวจชั้นภูมิที่ $h = \overbrace{c_h + c_h + \ldots + c_h}^{n_h \text{ ครั้ง}}$ ขอให้สังเกตว่าค่าใช้จ่ายผันแปรจะเพิ่มพูนขึ้นตามขนาดตัวอย่างที่จัดสรรให้แก่แต่ละชั้นภูมิ ถ้าจัดสรร n_h ไว้ไม่เหมาะสม เช่น จัดสรรตัวอย่างให้แก่ชั้นภูมิที่เสียค่าใช้จ่ายต่อหน่วยสูงในจำนวนที่มากเกินไป หรือจัดให้แก่ชั้นภูมิที่เสียค่าใช้จ่ายต่อหน่วยต่ำในจำนวนที่น้อยเกินไป งบประมาณที่จะใช้จ่ายย่อมได้รับความกระทบกระเทือนและส่งผลสะท้อนไปถึง $V(\hat{\theta})$ ด้วย ปัญหาก็คือ เราพึงจัดสรร n_h อย่างไรจึงจะสอดคล้องกับงบประมาณเพื่อการสำรวจ (ไม่นับค่าใช้จ่ายคงที่) และค่าใช้จ่ายต่อหน่วยในชั้นภูมิหนึ่ง ๆ ทั้งยังมีผลให้งานมีความแม่นยำสูงสุด ($V(\hat{\theta})$ ต่ำสุด) เท่าที่จะพึงเป็นไปได้ภายใต้ข้อจำกัดนี้

พิจารณาสมการค่าใช้จ่าย $c = c_0 + \sum_{h}^{L} n_hc_h$ จะพบว่าเราสามารถกันค่าใช้จ่ายคงที่ไว้ต่างหากได้ทำให้ค่าใช้จ่ายที่เหลืออยู่เป็นค่าใช้จ่ายผันแปรเท่านั้น นั่นคือ สมการค่าใช้จ่ายจะลดรูปเป็น

$$C = \sum_{h}^{L} n_h c_h \qquad \text{โดยที่ } C = c - c_0$$

การพิจารณาลดรูปเช่นนี้จะช่วยให้งานพัฒนาสูตรโครงสร้างของ n_h และสูตรอื่น ๆ ที่เกี่ยวข้องมีลักษณะที่ง่ายต่อการจดจำ

1. การพัฒนาสูตรโครงสร้างของ n_h ตามวิธีการจัดสรรแบบอุตมะ

ให้ $V(\bar{x}_{st})$ หรือ $\hat{V}(\bar{x}_{st})$ คือ ฟังก์ชันที่เราต้องการ minimize ภายใต้ข้อจำกัดของงบประมาณ $C = \sum_{h}^{L} n_h c_h$ หรือ $\sum_{h}^{L} n_h c_h - C = 0$

ดังนั้น โดยอาศัย Lagrange Multiplier Method จะพบว่า

$$F = V(\bar{x}_{st}) + \lambda\left(\sum_{h}^{L} n_h c_h - C\right) \text{ เมื่อ } \lambda \text{ คือ Lagrange Multiplier}$$

$$= \frac{1}{N^2} \sum_{h}^{L} \frac{N_h - n_h}{N_h} \frac{N_h^2 S_h^2}{n_h} + \lambda \sum_{h}^{L} n_h c_h - \lambda C$$

เพื่อจะหา n_h ที่ minimize $V(\bar{x}_{st})$ ภายใต้ข้อจำกัดของงบประมาณ (Constraint) $\sum_{h}^{L} n_h c_h - C$ เราจะต้องดิฟเฟอร์เรนชิเอท F เทียบต่อ n_h แล้วให้เท่ากับ 0 คือ $\frac{\partial F}{\partial n_h} = 0$

ดังนี้

$$F = \frac{1}{N^2} \sum_{h}^{L} \left(1 - \frac{n_h}{N_h}\right) \frac{N_h^2 S_h^2}{n_h} + \lambda \sum_{h}^{L} n_h c_h - \lambda C$$

$$= \frac{1}{N^2} \sum_{h}^{L} \frac{N_h^2 S_h^2}{n_h} - \frac{1}{N^2} \sum_{h}^{L} N_h S_h^2 + \lambda \sum_{h}^{L} n_h c_h - \lambda C$$

$$\frac{\partial F}{\partial n_h} = -\frac{1}{N^2} \frac{N_h^2 s_h^2}{n_h^2} - 0 + \lambda c_n - 0 = 0; h = 1,2,\ldots,L'$$

$$n_h^2 = \frac{N_h^2 S_h^2}{N^2 \lambda c_h}$$

$$n_h = \frac{1}{\sqrt{\lambda}} \cdot \frac{1}{\sqrt{c_h}} \cdot \frac{N_h S_h}{N} \text{ ; } h = 1,2,\ldots,L \qquad \ldots\ldots\ldots.(1)$$

แต่ $n_1 + n_2 + \ldots + n_L = n$ หรือ $\sum_h^L n_h = n$ ดังนั้นเรา Sum สมการที่ 1 ทั้งสองด้านของเครื่องหมายเท่ากับจะพบว่า

$$\sum_h^L n_h = \frac{1}{\sqrt{\lambda}} \sum_h^L \frac{N_h S_h}{N\sqrt{c_h}}$$

$$=> \quad n = \frac{1}{\sqrt{\lambda}} \sum_h^L \frac{N_h S_h}{N\sqrt{c_h}}$$

$$=> \quad \sqrt{\lambda} = \frac{1}{n} \sum_n^L \frac{N_h S_h}{N\sqrt{c_h}} \qquad \ldots\ldots\ldots.(2)$$

แทนค่า $\sqrt{\lambda}$ จากสมการที่ 2 ลงใน $\sqrt{\lambda}$ ของสมการที่ (1)

$$=> \quad n_h = \frac{n}{\sum_h^L \frac{N_h S_h}{N\sqrt{c_h}}} \cdot \frac{1}{\sqrt{c_h}} \cdot \frac{N_h S_h}{N}$$

$$= n(N_h S_h/\sqrt{c_h})/(\sum_h^L N_h S_h/\sqrt{c_h}); \; h = 1,2,\ldots,L$$

นั่นคือ ขนาดตัวอย่างที่ควรจัดสรรให้แก่แต่ละชั้นภูมิที่สอดคล้องกับงบประมาณเพื่อการวิจัยและเป็นขนาดตัวอย่างที่ทำให้ $V(\bar{x}_{st})$ ต่ำที่สุดก็คือ

$$n_h = n(N_h S_h/\sqrt{c_h})/(\sum_h^L N_h S_h/\sqrt{c_h}); \; h = 1,2,\ldots,L$$

และในกรณีที่ไม่อาจทราบค่าความแปรปรวนคือ S_h^2 ให้ใช้ s_h^2 แทน ดังนั้น

$$n(N_h s_h/\sqrt{c_h})/(\sum_h^L N_h s_h/\sqrt{c_h});\ h = 1,2,\ldots, L$$

ข้อสังเกต

ก. ผลรวมคือ $\Sigma N_h S_h/\sqrt{c_h}$ จะเป็นค่าคงที่เสมอ ซึ่งแสดงให้เห็นว่าตัวหารในสูตร n_h นี้ เป็นค่าคงที่ K ดังนั้นเราอาจเสนอสูตร n_h ได้ใหม่เป็น

$$n_h = \frac{n(N_h S_h/\sqrt{c_h})}{K} \qquad \text{เมื่อ } K = \sum_h^L N_h S_h/\sqrt{c_h}$$

ข. เมื่อพิจารณาสูตร $n_h = \frac{n(N_h S_h/\sqrt{c_h})}{K}$ จะพบว่าค่าของ n_h จะแตกต่างกันไปดังนี้คือ

(1) ถ้าชั้นภูมิใดมีขนาด (N_h) ใหญ่กว่า ชั้นภูมินั้นจะได้รับการจัดสรรตัวอย่างให้มากกว่า

(2) ถ้าชั้นภูมิใดมีความแปรปรวนภายใน (Within Stratum Variance S_h^2) สูงกว่าจะได้รับการจัดสรรตัวอย่างมากกว่า

(3) ถ้าชั้นภูมิใดเสียค่าใช้จ่ายต่อหน่วย (c_h) สูงกว่าจะได้รับการจัดสรรตัวอย่างให้น้อยกว่า

(4) จากสูตร $n_h = n(N_h S_h/\sqrt{c_h})/(\sum_h^L N_h S_h/\sqrt{c_h})$ ถ้า $S_1 = S_2 = \ldots = S_L = S$ และ $c_1 = c_2 = \ldots c_L = c$ แล้วสูตรของ n_h จะเปลี่ยนแปลงได้ดังนี้คือ

$$n(N_h S_h/\sqrt{c_h})/(\sum_h^L N_h S_h/\sqrt{c_h})$$

$$= n(N_h S/\sqrt{c})/(\sum_h^L N_h S/\sqrt{c})$$

$$= nN_h / \sum_{h}^{L} N_h$$

แต่ $\sum_{h}^{L} N_h = N$

$$=> n_h = \frac{n}{N} \cdot N_h ; h = 1,2,\ldots,L$$

นั่นก็คือ ถ้าค่าใช้จ่ายในการสำรวจต่อหน่วยในชั้นภูมิต่าง ๆ มีค่าเท่ากันและทุกชั้นภูมิมีความแปรปรวนภายในเท่ากันแล้ว แผนจัดสรรตัวอย่างแบบอุตมะจะกลายเป็นแผนจัดสรรอย่างได้สัดส่วนกับขนาดชั้นภูมิ

แต่อย่างไรก็ตามโดยปกติชั้นภูมิต่าง ๆ จะมีความแปรปรวนภายในไม่เท่ากันและจะเสียค่าใช้จ่ายในการสำรวจไม่เท่ากันทั้งนี้เพราะมีความแตกต่างกันในปัจจัยทางภูมิศาสตร์ เศรษฐกิจและสังคม ซึ่งเป็นเงื่อนไขที่สำคัญและบังคับให้ไม่อาจถือว่าเท่ากันได้ ความจริงข้อนี้ชี้ให้เห็นจุดอ่อนอย่างสำคัญของแผนจัดสรรอย่างเป็นสัดส่วนกับขนาดของชั้นภูมิในตอนก่อน ขณะเดียวกันก็ช่วยชี้ให้เห็นว่าในกรณีที่เป็นไปได้คือ $S_1 = S_2 = \ldots = S_L = S$ และ $c_1 = c_2 = \ldots = c_L = c$ นั้น แผนสำรวจที่จัดสรรตัวอย่างแบบอุตมะมิได้มีประสิทธิภาพสูงไปกว่าแผนจัดสรรอย่างได้สัดส่วนกับขนาดประชากรเลย อีกทั้งยังยุ่งยากซับซ้อนกว่าเสียอีกด้วย

2. การประมาณค่าพารามิเตอร์

การประมาณค่าพารามิเตอร์ เช่น $\bar{X}$ และ T กระทำได้โดยง่ายเพียงแต่แทนค่า n_h ในสูตรแม่บทคือทฤษฎี 3.1 และบทแทรก 3.1 ด้วย $n_h = n(N_h S_h / \sqrt{c_h})/K$

เมื่อ $K = \sum_{h}^{L} N_h S_h / \sqrt{c_h}$ หรือจะใช้สูตรเฉพาะกรณีของการจัดสรรก็ได้ซึ่งสามารถพัฒนาได้โดยง่ายดังนี้

1. $\hat{\bar{X}} = \bar{x}_{opt} = \frac{1}{N} \sum_{h}^{L} N_h \bar{x}_h$ เมื่อ $\bar{x}_h = \frac{1}{n_h} \sum_{i}^{n_h} x_{hi}$

โดยที่ $n_h = n(N_hS_h/\sqrt{c_h})/K$; $h = 1,2,\ldots,L$

2. จาก $V(\bar{x}_{st}) = \frac{1}{N^2}\sum_h^L \frac{N_h - n_h}{N_h}\frac{N_h^2S_h^2}{n_h}$

$$= \frac{1}{N^2}\sum_h^L \frac{N_h^2S_h^2}{n_h} - \frac{1}{N_2}\sum_h^L N_hS_h^2$$

แทนค่า n_h ด้วย $n(N_hS_h/\sqrt{c_h})/K$

$$V(\bar{x}_{opt}) = \frac{1}{N^2}\sum_h^L N_h^2S_h^2/\{n(N_hS_h/\sqrt{c_h})K\} - \frac{1}{N^2}\sum_h^L N_hS_h^2$$

$$= \frac{1}{N^2}\sum_h^L \frac{1}{n}(N_hS_h\sqrt{c_h}).K - \frac{1}{N^2}\sum_h^L N_hS_h^2$$

เมื่อ $K = \sum_h^L N_hS_h/\sqrt{c_h}$

$$V(\bar{x}_{opt}) = \frac{1}{N^2}\cdot\frac{1}{n}(\sum_h^L N_hS_h\sqrt{c_h})(\sum_h^L N_hS_h/\sqrt{c_h}) - \frac{1}{N^2}\sum_h^L N_hS_h^2$$

3. $\hat{V}(\bar{x}_{opt}) = \frac{1}{N^2}\cdot\frac{1}{n}(\sum_h^L N_hs_h\sqrt{c_h})(\sum_h^L N_hs_h/\sqrt{c_h}) - \frac{1}{N^2}\sum_h^L N_hs_h^2$

4. $\hat{T}_{opt} = \sum_h^L N_h\bar{x}_h$

5. $V(\hat{T}_{opt}) = \frac{1}{n}(\sum_h^L N_hS_h\sqrt{c_h})(\sum_h^L N_hS_h/\sqrt{c_h}) - \sum_h^L N_hS_h^2$

6. $\hat{V}(\hat{T}_{opt}) = \frac{1}{n}(\sum_h^L N_hs_h\sqrt{c_h})(\sum_h^L N_hs_h/\sqrt{c_h}) - \sum_h^L N_hs_h^2$

หมายเหตุ

1. การวิเคราะห์รายการต่าง ๆ เหล่านั้นนักศึกษาสามารถเลี่ยงไปใช้สูตรแม่บทได้เพียงแต่แทนค่า n_h ในสูตรแม่บทด้วย $n_h = n(N_hS_h/\sqrt{c_h})/K$ หรือ $n_h = n(N_hs_h/\sqrt{c_h})/K$ แล้วแต่กรณี โดยที่

$$K = \sum_{h}^{L} N_hS_h/\sqrt{c_h} \qquad \text{หรือ} \quad K = \sum_{h}^{L} N_hs_h/\sqrt{c_h}$$

2. สูตรรายการ 1-6 ข้างต้นรายการใดมิได้แสดงการพัฒนามาให้ดูโดยละเอียด นักศึกษาสามารถพัฒนาได้เองโดยง่ายในที่นี้จะขอเว้นไว้เป็นแบบฝึกหัด

ง. การจัดสรรตัวอย่างแบบเนย์แมน (Neyman Allocation)

ในสถานการณ์ทางปฏิบัตินั้น บางครั้งค่าใช้จ่ายในการสำรวจต่อหน่วยในชั้นภูมิต่าง ๆ อาจใกล้เคียงกัน หรือ ไม่ต่างกันอย่างเด่นชัด ซึ่งในกรณีเช่นนี้เราอาจถือว่าค่าใช้จ่ายต่อหน่วยเหล่านั้นเท่ากันได้ กล่าวคือ ถือว่า $c_1 = c_2 = \ldots = c_L = c_f$ เมื่อ c_f เป็นค่าคงที่ใด ๆ ซึ่งย่อมมีผลให้ฟังก์ชันของค่าใช้จ่ายเปลี่ยนแปลงไปจากเดิม คือ $C = \sum_{h}^{L} n_hc_h$ กลายเป็น

$$C = \sum_{h}^{L} n_hc_f = c_f \sum_{h}^{L} n_h = nc_f$$ [1]

---

[1] จาก $C = c_fn$ แสดงว่า $c - c_0 = nc_f$ ดังนั้น $n = \frac{c-c_0}{c_f}$ ความจริงข้อนี้มีประโยชน์ประการสำคัญ คือ ถ้างบประมาณเท่ากับ C บาท และทราบว่า (หรือถือว่า) ค่าใช้จ่ายต่อหน่วยในการสำรวจของแต่ละหน่วยจากแต่ละชั้นภูมิมีค่าเท่ากัน (โดยประมาณ) $= c_f$ บาท เราจะสามารถกำหนดขนาดตัวอย่าง n เพื่อการสำรวจได้จากสมการ $n = \frac{c-c_0}{c_f}$ ขนาดตัวอย่างนี้แม้จะยังไม่รัดกุม แต่ก็พอเป็นแนวทางในการกำหนดขนาดตัวอย่างให้แก่นักวิจัยได้ในขณะที่นักวิจัยไม่ทราบข้อสนเทศใด ๆ ของกลุ่มประชากรเลย

ดังนั้นโดยอาศัย Lagrange Multiplier Method จะพบว่า

$$F = V(\bar{x}_{st}) + \lambda(nc_f - C) = V(\bar{x}_{st}) + \lambda(c_f \sum_{h}^{L} n_h - C)$$

อาศัยวิธีการในทำนองเดียวกันกับในตอน ค. เราสามารถคำนวณหาขนาดตัวอย่าง n_h ที่สอดคล้องกับค่าใช้จ่ายลักษณะนี้ได้ และเป็นขนาดตัวอย่างที่มีผลให้ $V(\bar{x}_{st})$ มีค่าต่ำสุดได้ดังนี้

$$F = \frac{1}{N^2}\sum_{h}^{L}\frac{N_h^2S_h^2}{n_h} - \frac{1}{N^2}\sum_{h}^{L} N_hS_h^2 + \lambda c_f \sum_{h}^{L} n_h - \lambda C$$

$$\frac{\partial F}{\partial n_h} = -\frac{1}{N^2}\frac{N_h^2S_h^2}{n_h^2} - 0 + \lambda c_f - 0 = 0 \; ; h = 1,2,\ldots,L$$

$$=> \quad n_h^2 = \frac{N_h^2S_h^2}{N^2\lambda c_f}$$

$$n_h = \frac{1}{\sqrt{\lambda}} \cdot \frac{1}{\sqrt{c_f}} \cdot \frac{N_hS_h}{N} \; ; h = 1,2,\ldots,L \qquad \ldots\ldots\ldots.(1)$$

เมื่อรวมกันตลอดในทุกค่าของ h จะพบว่า

$$\sum_{h}^{L} n_h = \frac{1}{\sqrt{\lambda}}\sum_{h}^{L} N_hS_h/N\sqrt{c_f}$$

$$=> \quad n = \frac{1}{\sqrt{\lambda}}\sum_{h}^{L} N_hS_h/N\sqrt{c_f}$$

$$\sqrt{\lambda} = \frac{1}{n}\sum_{h}^{L} N_hS_h/N\sqrt{c_f} \qquad \ldots\ldots\ldots.(2)$$

แทนค่า $\sqrt{\lambda}$ จากสมการที่ 2 ลงใน $\sqrt{\lambda}$ ของสมการที่ 1

$$n_h = \frac{N_hS_h}{N\sqrt{c_f}} / \{ \frac{1}{n} \sum_{h}^{L} N_hS_h/N\sqrt{c_f}\}$$

นั่นคือ $$n_h = n \cdot \frac{N_hS_h}{\sum_{h}^{L} N_hS_h} ; h=1,2,\ldots,L$$

ข้อสังเกต

1. จากสูตร $n_h = \frac{N_hS_h}{\sum N_hS_h} \cdot n$ ส่วนคือ $\sum_{h}^{L} N_hS_h$ จะมีค่าคงที่เสมอ

ถ้าให้ $\sum_{h}^{L} N_hS_h = K$ เราอาจเสนอสูตรในรูปใหม่ได้ดังนี้

$$n_h = \frac{nN_hS_h}{K} ; h=1,2,\ldots,L$$

2. เมื่อพิจารณาสูตร $n_h = n \cdot \frac{N_hS_h}{K}$ จะพบว่าขนาดตัวอย่างที่จัดสรรให้แก่ชั้นภูมิที่ h จะต้องผันแปรไปตามขนาดของชั้นภูมิและความแปรปรวนภายในชั้นภูมิ กล่าวคือ ชั้นภูมิใดมีขนาดใหญ่กว่าเราจะต้องจัดสรรตัวอย่างให้มากกว่า และชั้นภูมิใดมีความแปรปรวนภายในมากกว่าหรือขาดความเป็นเนื้อเดียวกันเราจะต้องจัดสรรตัวอย่างให้มากกว่า

3. สูตร $n_h = \frac{N_hS_h}{\sum N_hS_h} \cdot n$ สามารถพัฒนาขึ้นมาจากสูตร

$$n_h = n(N_hS_h/\sqrt{c_h})/(\sum_{h}^{L} N_hS_h/\sqrt{c_h})$$ ได้ดังนี้คือ

เมื่อถือว่า $c_1 = c_2 = \ldots = c_L = c_f$ ดังนั้น

$$n_h = n(N_hS_h/\sqrt{c_f})/(\sum_{h}^{L} N_hS_h/\sqrt{c_f}) ; h=1,2,\ldots,L$$

$$= nN_hS_h / \sum_{h}^{L} N_hS_h$$

$$= \frac{N_hS_h}{\Sigma N_hS_h} \cdot n \quad ; \; h = 1,2,\ldots,L$$

ด้วยเหตุนี้เราจึงกล่าวได้ว่าการจัดสรรตัวอย่างแบบเนย์แมนเป็นกรณีเฉพาะของการจัดสรรแบบอุตมะ เมื่อถือว่าค่าใช้จ่ายในการสำรวจต่อหน่วยในชั้นภูมิต่าง ๆ ใกล้เคียงกันจนสามารถถือว่าเท่ากันได้

4. ถ้า $S_1 = S_2 = \ldots = S_L = S$ **หรือนัยหนึ่ง** ถ้าความแปรปรวนภายในชั้นภูมิต่าง ๆ มีค่าเท่ากัน จะพบว่า $n_h = \frac{N_hS}{\Sigma N_hS} \cdot n = \frac{N_h}{N} \cdot n = \frac{n}{N} . N_h ; h=1,2,\ldots,L$ นั่นก็คือ ถ้าความแปรปรวนภายในชั้นภูมิต่าง ๆ มีค่าเท่ากันหรือใกล้เคียงกันมาก แผนจัดสรรแบบเนย์แมน จะลดรูปเป็นแผนจัดสรรอย่างได้สัดส่วนกับขนาดของชั้นภูมิ หรือนัยหนึ่ง ถ้าความแปรปรวนภายในชั้นภูมิต่าง ๆ มีค่าเท่ากันหรือใกล้เคียงกันมาก เราควรจะใช้แผนจัดสรรตัวอย่างให้เป็นสัดส่วนต่อขนาดของชั้นภูมิแทนเพราะสะดวกรวดเร็วกว่าแต่มีประสิทธิภาพเท่ากันหรือใกล้เคียงกัน

สำหรับการวิเคราะห์ข้อมูลเพื่อกะประมาณพารามิเตอร์คือ $\bar{X}$ และ T นั้นให้ปฏิบัติเช่นเดียวกันกับแผนอื่น ๆ คือให้อาศัยสูตรแม่บท เพียงแต่แทนค่า n_h ในสูตรแม่บทด้วย $n_h = \frac{N_hS_h}{\Sigma N_hS_h} \cdot n$ และในกรณีที่ไม่ทราบค่า S_h ให้ประมาณค่า S_h ด้วย s_h หรือจะใช้สูตรเฉพาะแผนก็ได้ สูตรดังกล่าวพัฒนาได้ง่าย ๆ ดังนี้คือ

1. $\hat{\bar{X}} = \bar{x}_{ney} = \frac{1}{N} \sum_{h}^{L} N_h\bar{x}_h$ เมื่อ $\bar{x}_h = \frac{1}{n_h} \sum_{i}^{n_h} x_{hi}$

โดยที่ $n_h = \frac{N_hS_h}{\Sigma N_hS_h} \cdot n \quad ; \; h=1,2,\ldots,L$

2. $V(\bar{x}_{ney}) = \frac{1}{N^2} \sum_{h}^{L} \frac{N_h - n_h}{N_h} \frac{N_h^2S_h^2}{n_h}$ เมื่อ $n_h = \frac{N_hS_h}{\Sigma N_hS_h} \cdot n$