บทที่ 2

ทฤษฎีที่ใช้ในการวิจัยและผลงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่ใช้ในการวิจัย

ทฤษฎีที่ใช้ในการปรับความเอนเอียงของตัวประมาณค่า อันเนื่องมาจากการไม่ตอบข้อถามของตัวแปรที่ผู้วิจัยสนใจหรือไม่ได้รับแบบสอบถามคืนมาทั้งฉบับ เมื่อเก็บข้อมูลโดยส่งแบบสอบ-ถามทางไปรษณีย์นั้น วิทยานิพนธ์นี้จะศึกษาวิธีการประมาณค่าดังกล่าว 2 วิธีคือ

(1) วิธีแฮนเซ็น-เฮอร์วิทซ์

(2) วิธีเอล-บาดรี้

2.1.1 วิธีแฮนเซ็น-เฮอร์วิทซ์

วิธีแฮนเซ็น-เฮอร์วิทซ์ เป็นวิธีการที่ผสมผสานเอาเทคนิคการประมาณค่าพารามิเตอร์เมื่อเก็บข้อมูลโดยส่งแบบสอบถามทางไปรษณีย์ กับเมื่อเก็บข้อมูลโดยการสัมภาษณ์หน่วยตัวอย่าง ซึ่งมีขั้นตอนของวิธีการประมาณค่าพารามิเตอร์ดังนี้

ขั้นตอนที่ 1 สมมติว่าประชากรที่สนใจมีขนาดเท่ากับ N หน่วย ซึ่งทราบที่อยู่อาศัย และจะต้องส่งแบบสอบถามให้ แล้วเลือกตัวอย่างมาขนาด n หน่วยจากประชากรขนาด N หน่วย โดยเลือกตัวอย่างแบบสุ่มอย่างง่าย (Simple Random Sampling) แบบไม่ใส่คืน (Without Replacement) และส่งแบบสอบถามทางไปรษณีย์ให้กับหน่วยตัวอย่างที่ได้ขนาด n หน่วย สมมติว่าได้จำนวนหน่วยตัวอย่างที่ตอบข้อถามของตัวแปรที่ผู้วิจัยสนใจเท่ากับ n_1 หน่วย และได้จำนวนหน่วยตัวอย่างที่ไม่ตอบข้อถามขอตัวแปรที่ผู้วิจัยสนใจเท่ากับ n_2 หน่วย โดยที่ $n = n_1 + n_2$

ขั้นตอนที่ 2 จากจำนวนหน่วยตัวอย่างที่ไม่ตอบขนาด n_2 หน่วยนั้นเลือกตัวอย่างย่อย(Subsample)ขนาด n_2' หน่วย จากหน่วยตัวอย่างขนาด n_2 หน่วยอีกครั้งหนึ่งโดยเลือกแบบสุ่มอย่างง่ายไม่ใส่คืน (Simple Random Sampling Without Replacement)

โดยที่ $$n_2' = \frac{n_2}{k_{opt}}$$

k_{opt} คือ ค่าคงที่ค่าหนึ่งซึ่งเป็นส่วนกลับของ Subsampling fraction $\left(\frac{1}{n'_2/n_2}\right)$ ที่ต้องการให้ค่าใช้จ่ายต่ำสุด เมื่อกำหนด ความแปรปรวนของค่าประมาณที่ไม่เกิดปัญหาการไม่ตอบ (ε^2) ให้ และโดยวิธีนี้ ค่า k_{opt} มากกว่า 1

<u>ขั้นตอนที่ 3</u> ส่งพนักงานไปสัมภาษณ์หน่วยตัวอย่างย่อยขนาด n'_2 ทุกหน่วยที่ปรากฏอยู่ในตัวอย่างย่อย และสมมติว่า หน่วยตัวอย่างย่อยขนาด n'_2 หน่วยเมื่อถูกสัมภาษณ์แล้วจะต้องตอบข้อถามที่ผู้วิจัยสนใจโดยตอบทุกหน่วย

วิธีแฮนเซ็น-เฮอร์วิทซ์ นี้สามารถนำเทคนิคของ การสุ่มตัวอย่างสองครั้ง (Two-phases Sampling หรือ Double Sampling) สำหรับการแบ่งชั้นภูมิ (Stratification) มาประยุกต์ใช้ได้ โดยที่ประชากรแบ่งออกเป็น 2 ชั้นภูมิ ให้ชั้นภูมิที่ 1 เป็นหน่วยที่ตอบคำถามที่สนใจที่มีขนาด n_1 หน่วย และให้ชั้นภูมิที่ 2 เป็นหน่วยที่ไม่ตอบข้อถามที่สนใจซึ่งมีขนาด n_2 หน่วย จากหน่วยตัวอย่างขนาด n_2 หน่วยจะเลือกตัวอย่างย่อยมาขนาด n'_2 หน่วย โดยที่ $n'_2 = \frac{n_2}{k_{opt}}$

ให้ชั้นภูมิที่ 1 ประกอบด้วย หน่วยตัวอย่างที่ตอบสำหรับ การส่งคำถามทางไปรษณีย์ครั้งแรกด้วยขนาดตัวอย่าง n_1 หน่วย

ให้ชั้นภูมิที่ 2 ประกอบด้วย หน่วยตัวอย่างที่ไม่ตอบข้อถามในการส่งแบบสอบถามทางไปรษณีย์ครั้งแรกและตอบข้อถามหลังจากถูกสัมภาษณ์ซึ่งมีขนาด n_2หน่วย

ให้ $$\bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_{1i} \quad \text{(ชั้นภูมิที่ 1)} \quad \ldots\ldots\ldots\ldots(1)$$

$$\bar{x}'_2 = \frac{1}{n'_2}\sum_{i=1}^{n'_2} x_{2i} \quad \text{(ชั้นภูมิที่ 2)} \quad \ldots\ldots\ldots\ldots(2)$$

ใช้สมการ (1) และ (2) เป็นตัวประมาณค่าเฉลี่ยของประชากร ($\bar{X}$) ได้ โดยจะประมาณ $\bar{X}$ ด้วย $\bar{x}$ โดยที่

$$\bar{x} = w_1 \cdot \bar{x}_1 + w_2 \cdot \bar{x}_2'$$

$$= \frac{n_1}{n} \cdot \bar{x}_1 + \frac{n_2}{n} \cdot \bar{x}_2' \qquad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

สมการ (3) จะเป็นตัวประมาณที่ไม่เอนเอียงสำหรับ $\bar{X}$ (Zarkovich, S.S, 1966 : 152) เพราะว่า

$$E\left[\frac{n_1}{n} \cdot \bar{x}_1\right] = E\left[E(w_1 \cdot \bar{x}_1 \mid w_1)\right]$$

$$= E\left[w_1 \cdot \bar{X}_1\right] = E\left[\frac{n_1}{n} \cdot \bar{X}_1\right]$$

$$= \frac{N_1}{N} \cdot \bar{X}_1$$

และ
$$E\left[\frac{n_2}{n} \cdot \bar{x}_2'\right] = E\left[E(w_2\ \bar{x}_2' \mid w_2)\right]$$

$$= E\left[w_2 \cdot \bar{x}_2\right]$$

$$= E\left[E(w_2\ \bar{x}_2) \mid w_2\right]$$

$$= E\left[w_2\ \bar{X}_2\right] = E\left[\frac{n_2}{n} \cdot \bar{X}_2\right]$$

$$= \frac{N_2}{N} \cdot \bar{X}_2$$

ดังนั้น
$$E(\bar{x}) = E\left[\frac{n_1}{n} \cdot \bar{x}_1 + \frac{n_2}{n} \cdot \bar{x}_2'\right]$$

$$= E\left[\frac{n_1}{n} \cdot \bar{x}_1\right] + E\left[\frac{n_2}{n}\ \bar{x}_2'\right]$$

$$= \frac{N_1}{N} \cdot \bar{X}_1 + \frac{N_2}{N} \cdot \bar{X}_2$$

$$= \bar{X}$$

2.1.1.1 การหาความแปรปรวนของค่าเฉลี่ยของตัวอย่าง ($V(\bar{x})$)

จาก $$\bar{x} = \frac{n_1}{n} \cdot \bar{x}_1 + \frac{n_2}{n} \cdot \bar{x}_2' \quad \text{................(4)}$$

เนื่องจากมีการสุ่มตัวอย่างสองครั้ง คือ ตัวอย่างขนาด n_1 หน่วยที่ตอบ และตัวอย่างย่อยขนาด n_2' หน่วยซึ่งเลือกมาจากตัวอย่างขนาด n_2 หน่วยที่ไม่ตอบ ดังนั้นจึงหาค่าความแปรปรวนของตัวประมาณค่าเฉลี่ยประชากรในสมการ (4) ภายใต้การสุ่มตัวอย่างสองครั้ง (Over the Two Phase of Randomization) (Murthy 1967 : 464-465) คือ

$$V_{12}(\bar{x}) = V_1E_2(\bar{x}) + E_1V_2(\bar{x}) \quad \text{.............. (5)}$$

พิจารณา $E_2(\bar{x})$

$$E_2(\bar{x}) = E_2\left[\frac{n_1}{n} \cdot \bar{x}_1 + \frac{n_2}{n} \cdot \bar{x}_2'\right]$$

$$= \frac{n_1}{n} \cdot \bar{x}_1 + \frac{n_2}{n} \cdot \bar{x}_2 = \frac{1}{n}\left\{\sum_{i=1}^{n_1} x_{1i} + \sum_{i=1}^{n_2} x_{2i}\right\}$$

= ค่าเฉลี่ยของตัวอย่างสุ่มขนาด n ที่เลือกมาจากประชากรทั้งหมดนั่นเอง

ดังนั้น $$V_1E_2(\bar{x}) = \left(\frac{N-n}{N}\right) \cdot \frac{S^2}{n} \quad \text{...............(6)}$$

พิจารณา $V_2(\bar{x})$

จาก $$V_2(\bar{x}) = V_2\left[\frac{n_1}{n} \cdot \bar{x}_1 + \frac{n_2}{n} \cdot \bar{x}_2'\right]$$

$$= \left(\frac{n_2}{n}\right)^2 V_2(\bar{x}_2')$$

$$= \left(\frac{n_2}{n}\right)^2 \left[\frac{(n_2-n_2')}{n_2} \cdot \frac{s_2^2}{n_2'}\right] \quad ; \quad s_2^2 = \frac{1}{(n_2-1)} \sum_{i=1}^{n_2} (x_{2i}-\bar{x}_2)^2$$

$$\therefore \quad V_2(\bar{x}) = \frac{1}{n}\left[\left(\frac{n_2}{n}\right)\left(\frac{n_2}{n_2'} - 1\right) \cdot s_2^2\right] \quad \text{.......... (7)}$$

นำค่า $V_1E_2(\bar{x})$ ในสมการ (6) และค่า $V_2(\bar{x})$ ในสมการ (7) ไปแทนในสมการ (5) จะได้

$$V_{12}(\bar{x}) = \left(\frac{N-n}{N}\right)\frac{S^2}{n} + E_1\left\{\left(\frac{1}{n}\right)\cdot\left(\frac{n_2}{n}\right)(k_{opt} - 1)\ .\ s_2^2\right\} \ldots\ldots (8)$$

เมื่อ $k_{opt} = \dfrac{n_2}{n_2'}$

เนื่องจากค่าคาดหวังแบบมีเงื่อนไขของ s_2^2 เมื่อกำหนด n_2 (the Conditional Expected Value of s_2^2 given n_2) คือ $(s_2^2|\ n_2) = S_2^2$ และค่าคาดหวัง (Expected Value) ของ $\dfrac{n_2}{n}$ คือ $\dfrac{N_2}{N}$ ดังนั้น

จากสมการ (8) พิจารณาเทอม $E_1\left\{\left(\frac{1}{n}\right)\left(\frac{n_2}{n}\right)(k_{opt} - 1)\ .s_2^2\right\}$

$$E_1\left\{\left(\frac{1}{n}\right)\left(\frac{n_2}{n}\right)(k_{opt} - 1)\ .\ s_2^2\right\} = \frac{1}{n}(k_{opt} - 1)\ E_1\left\{\frac{n_2}{n}\ .\ s_2^2\right\}$$

$$= \frac{1}{n}(k_{opt} - 1)\ E_1\left\{E_1\left(\frac{n_2}{n}.s_2^2\ |\ n_2\right)\right\}$$

$$= \frac{1}{n}(k_{opt} - 1)\ E_1\left\{\left(\frac{n_2}{n}\right)\ E_1\ (s_2^2|\ n_2)\right\} \ldots (9)$$

จากสมการ (9) พิจารณา $E_1(s_2^2|\ n_2)$

$$s_2^2 = \frac{1}{(n_2 - 1)}\sum_{i=1}^{n_2}(x_{2i} - \bar{x}_2)^2$$

$$= \frac{1}{(n_2 - 1)}\sum_{i=1}^{n_2}\left\{(x_{2i} - \bar{X}_2)^2 - n_2(\bar{x}_2 - \bar{X}_2)^2\right\}$$

ดังนั้น $E_1\ (s_2^2\ |n_2) = E_1\left[\left(\frac{1}{(n_2-1)}\left\{\sum_{i=1}^{n_2}(x_{2i} - \bar{X}_2)^2 - n_2\ (\bar{x}_2 - \bar{X}_2)^2\right\}\right]$

$$= \frac{1}{(n_2-1)}\left[E_1\left\{\sum_{i=1}^{n_2}(x_{2i}-\bar{X}_2)^2\right\} - E_1\left\{n_2(\bar{x}_2-\bar{X}_2)^2\right\}\right]$$

$$= \frac{1}{(n_2-1)}\left[\left\{\sum_{i=1}^{n_2} E_1\,(x_{2i}-\bar{X}_2)^2\right\} - E_1\left\{n_2\,(\bar{x}_2-\bar{X}_2)^2\right\}\right]$$

$$= \frac{1}{(n_2-1)}\left[\left\{\sum_{i=1}^{n_2} \cdot \sum_{i=1}^{N_2}\frac{1}{N_2}\,(x_{2i}-\bar{X}_2)^2\right\} - n_2\,E_1(\bar{x}_2-\bar{X}_2)^2\right]$$

$$= \frac{1}{(n_2-1)}\left[\left\{\frac{1}{N_2}\cdot n_2 \cdot \sum_{i=1}^{N_2}(x_{2i}-\bar{X}_2)^2\right\} - \left\{n_2\,.V(\bar{x}_2)\right\}\right]$$

$$= \frac{1}{(n_2-1)}\left[\left\{\frac{1}{N_2}\cdot n_2(N_2-1).S_2^2\right\} - \left\{n_2\cdot\frac{(N_2-n_2)}{N_2}\cdot\frac{S_2^2}{n_2}\right\}\right]$$

$$= \frac{1}{(n_2-1)}\left[n_2S_2^2 - \frac{n_2}{N_2}\cdot S_2^2 - S_2^2 + \frac{n_2}{N_2}S_2^2\right]$$

$$= \frac{1}{(n_2-1)}\left[(n_2-1)\cdot S_2^2\right]$$

$= S_2^2$ (without replacement){ Taro Yamane 1967 : 78-79}

ดังนั้น

$$E_1\,(s_2^2\,|n_2) = S_2^2 \qquad \text{..............(10)}$$

เพราะฉะนั้น นำค่า $E_1(s_2^2\,|n_2)$ ในสมการ (10) แทนในสมการ (9) จะได้

$$E_1\left\{\frac{1}{n}\cdot\left(\frac{n_2}{n}\right)\cdot(k_{opt}-1)\,s_2^2\right\} = \frac{1}{n}\,(k_{opt}-1)\;E_1\left\{\left(\frac{n_2}{n}\right)\cdot s_2^2\right\}$$

$$= \frac{1}{n}\,(k_{opt}-1).\;S_2^2\,E_1\left(\frac{n_2}{n}\right)$$

$$= \frac{1}{n}\,(k_{opt}-1).\;S_2^2\cdot\frac{N_2}{N} \qquad \text{.....(11)}$$

ดังนั้น จากสมการ (8) และ (11) จะได้

$$V_{12}(\bar{x}) = \frac{(N-n)}{N} \cdot \frac{S^2}{n} + \frac{1}{n}(k_{opt} - 1) \cdot \frac{N_2}{N} \cdot S_2^2$$

$$\therefore \; V(\bar{x}) = \frac{(N-n)}{N} \cdot \frac{S^2}{n} + \frac{(k_{opt}-1)}{n} \cdot P_2 S_2^2 \qquad \text{...........(12)}$$

เมื่อ $S^2 = \sum_{i=1}^{N} \frac{(x_i - \bar{X})^2}{N-1}$

= ความแปรปรวนของประชากร

$S_2^2 = \sum_{i=1}^{N_2} \frac{(x_{2i} - \bar{X}_2)^2}{N_2-1}$

= ความแปรปรวนของประชากรที่ไม่ตอบ

และ $P_2 = \frac{N_2}{N}$

= อัตราส่วนของการไม่ตอบในประชากร

2.1.1.2 การหาค่า k_{opt}

จาก $k_{opt} = n_2/n_2'$

เมื่อ n_2 คือ ขนาดตัวอย่างของกลุ่มที่ไม่ตอบหลังจากส่งแบบสอบถามทางไปรษณีย์

n_2' คือ ขนาดตัวอย่างที่สุ่มมาจาก n_2 และจะต้องถูกสัมภาษณ์ โดยสมมติว่า เมื่อถูกสัมภาษณ์แล้วทุกหน่วยจะต้องตอบ

การหาขนาดตัวอย่าง n ที่เหมาะสมและ k_{opt} นั้น โดยการทำให้ค่าคาดหวังของค่าใช้จ่ายมีค่าต่ำสุด ภายใต้ ความแปรปรวนของ $\bar{x}$ กรณีที่ตอบ 100% (ε^2) ที่กำหนดให้

$$\text{จาก} \quad C = c_o n + c_1 n_1 + c_2 n_2'$$

$$\therefore E(C) = E\left[c_o n + c_1 n_1 + c_2 n_2'\right]$$

$$= E(c_o n) + E(c_1 n_1) + E(c_2 n_2')$$

$$= n\left[c_o + E(c_1 \cdot \frac{n_1}{n}) + E(c_2 \cdot \frac{n_2'}{n} \cdot \frac{n_2}{n_2})\right]$$

$$= n\left[c_o + c_1 \cdot \frac{N_{11}}{N} + c_2 E\left(\frac{n_2}{n} \cdot \frac{1}{k}\right)\right] \; ; \quad k = n_2/n_2'$$

$$= n\left[c_o + c_1 \cdot \frac{N_{11}}{N} + c_2 \frac{N_{12}}{kN}\right]$$

ให้ $F = E(C) + \lambda(S_{\bar{x}}^2 - \varepsilon^2)$ ภายใต้เงื่อนไข (Subject to) $S_{\bar{x}}^2 - \varepsilon^2 = 0$

เมื่อ λ คือ Lagrance Multiplier

$$\therefore \frac{\partial F}{\partial n} = (c_o + c_1 \cdot \frac{N_{11}}{N} + c_2 \cdot \frac{N_{12}}{kN}) + \lambda\left[\frac{S^2}{n^2} + \frac{(k-1)}{n^2} \cdot P_{12} S_{12}^2\right] = 0$$

$$\lambda = \frac{\left[c_o + c_1 \frac{N_{11}}{N} + c_2 \cdot \frac{N_{12}}{kN}\right] n^2}{\left[S^2 + (k-1) P_{12} S_{12}^2\right]} \qquad \ldots\ldots\ldots\ldots (11)$$

$$\frac{\partial F}{\partial k} = - c_2 n. \frac{N_{12}}{k^2 N} + \left(\frac{P_{12} S_{12}^2}{n}\right) \lambda = 0$$

$$\therefore \quad c_2 \frac{n.N_{12}}{k^2 N} = \left(\frac{P_{12} S_{12}^2}{n}\right) \lambda \quad \ldots\ldots\ldots\ldots (12)$$

แทนค่า λ ในสมการ (12) จะได้

$$c_2 \frac{n.N_{12}}{k^2 N} = \left(\frac{P_{12} . S_{12}^2}{n}\right) \frac{\left(c_o + c_1 \frac{N_{11}}{N} + c_2 \frac{N_{12}}{kN}\right) . n^2}{(S^2 + (k-1)P_{12}S_{12}^2).}$$

$$\frac{c_2 N_{12}}{k^2 N} \left(S^2 + (k-1) \frac{N_{12}}{N} . S_{12}^2\right) = \left(c_o + c_1 \frac{N_{11}}{N} + c_2 \frac{N_{12}}{kN}\right)\left(\frac{N_{12}}{N} . S_{12}^2\right)$$

$$\frac{c_2}{k^2} . S^2 + \frac{c_2}{k} \frac{N_{12}}{N} S_{12}^2 - \frac{c_2}{k^2} \frac{N_2}{N} . S_{12}^2 = c_o S_{12}^2 + c_1 \frac{N_{11}}{N} S_{12}^2 + \frac{c_2}{k} . \frac{N_{12}}{N} . S_{12}^2$$

$$\frac{1}{k^2} \left(c_2 S^2 - c_2 \frac{N_{12}}{N} . S_{12}^2\right) = \left(c_o S_{12}^2 + c_1 \frac{N_{11}}{N} . S_{12}^2\right)$$

$$k^2 = \frac{c_2 (S^2 - P_{12} S_{12}^2)}{S_{12}^2 (c_o + c_1 P_{11})}$$

เมื่อ $P_{12} = \frac{N_{12}}{N}$

$P_{11} = \frac{N_{11}}{N}$

$$\therefore \; k_{opt} = \sqrt{\frac{c_2(S^2 - P_{12}S_{12}^2)}{S_{12}^2(c_o + c_1P_{11})}} \qquad \ldots\ldots\ldots\ldots (13)$$

2.1.1.3 การหา ขนาดตัวอย่างที่เหมาะสมที่สุด (Optimum Sample Size ; n_{opt})

หาได้จากสมการ $S_{\bar{x}}^2 - \varepsilon^2 = 0$

เพราะฉะนั้น $S_{\bar{x}}^2 = \varepsilon^2$

จากสมการ (10) จะได้ $\left(\frac{N - n}{N}\right) . \frac{S^2}{n} + \frac{(k_{opt} - 1)}{n} . P_{12} . S_{12}^2 = \varepsilon^2$

$$\frac{S^2}{n} - \frac{S^2}{N} + \frac{1}{n}(k_{opt} - 1) P_{12} S_{12}^2 = \varepsilon^2$$

$$\frac{1}{n}\left[S^2 + (k_{opt} - 1) P_{12} S_{12}^2\right] = \varepsilon^2 + \frac{S^2}{N}$$

$$\therefore \; n_{opt} = \frac{S^2 + (k_{opt} - 1) P_{12} S_{12}^2}{\varepsilon^2 + S^2/N}$$

$$= \frac{NS^2 + N(k_{opt} - 1) P_{12} S_{12}^2}{N\varepsilon^2 + S^2}$$

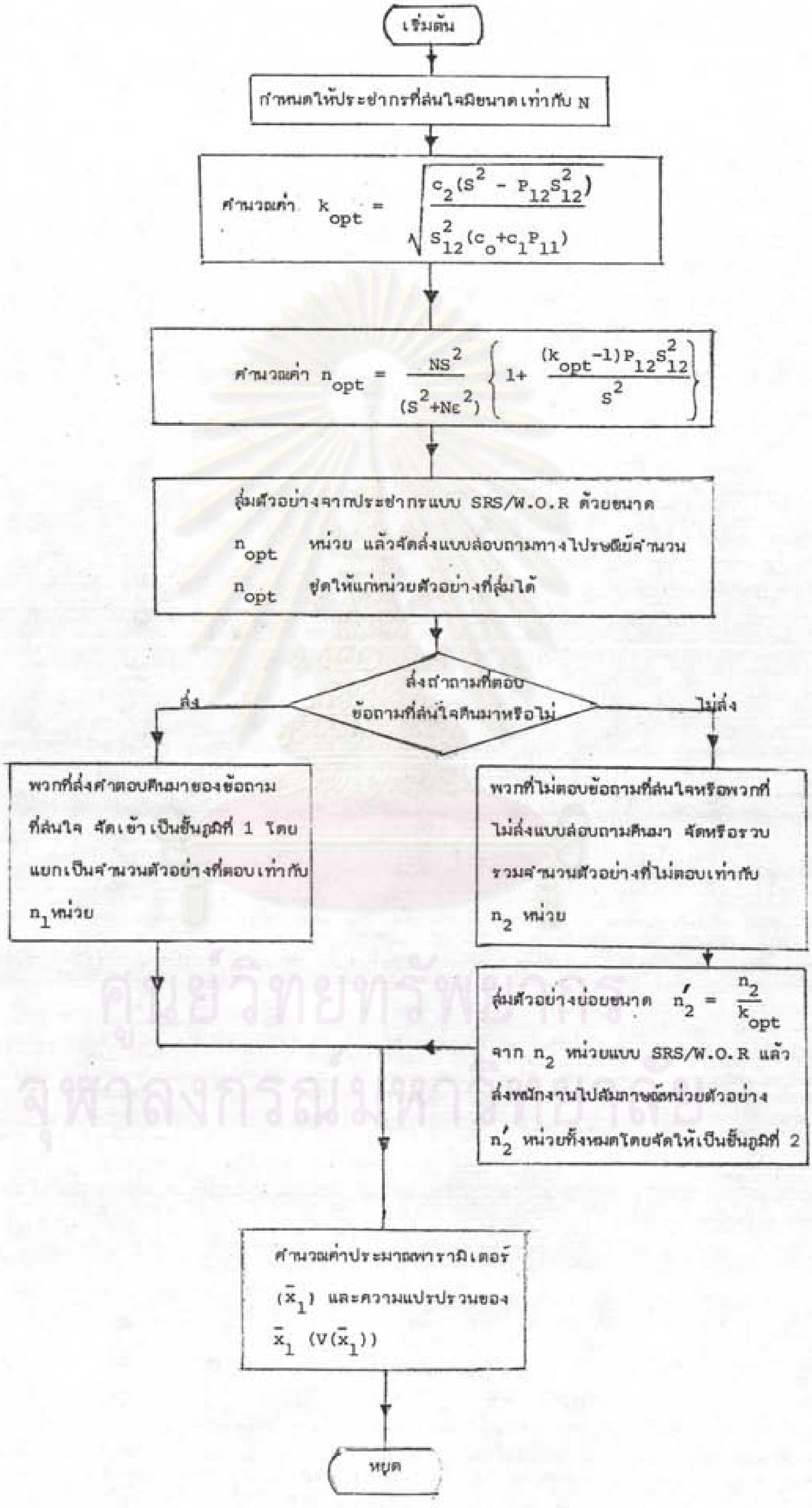
$$n_{opt} = \frac{NS^2}{(S^2+N\varepsilon^2)}\left[1 + \frac{(k_{opt} - 1) P_{12} S_{12}^2}{S^2}\right] \qquad \ldots\ldots (14)$$

เมื่อ ε^2 คือ $V(\bar{x})$ ที่เป็น 100% Response Rate

สำหรับบางค่าของ n' โดยที่ $\varepsilon^2 = \left(\frac{N - n'}{N}\right) . \frac{S^2}{n'}$

[Cochran 1963:23] กล่าวว่า ขนาดตัวอย่าง 500 หน่วย ที่เลือกมาจากประชากรขนาด 200,000 หน่วย จะให้ค่าประมาณของค่าเฉลี่ยประชากร มีค่าใกล้เคียงกับการประมาณค่าเฉลี่ยประชากรเมื่อใช้ขนาดตัวอย่าง 500 หน่วย ซึ่งเลือกมาจากประชากรขนาด 10,000 หน่วย [A sample of 500 from a population of 200,000 gives almost as precise an estimate of the population mean as a sample of 500 from a population of 10,000] ดังนั้นวิทยานิพนธ์นี้จึงกำหนดให้ $n' = 500$ หน่วย

แผนผังแสดงขั้นตอนของวิธีแฮนเซ็น-เฮอร์วิทซ์



แผนผังที่ 1

2.1.2 <u>วิธีเอล-บาดรี้</u>

(M.A. EL-BADRY 1956 : 209) ได้ขยายวิธีแฮนเซ็น-เฮอร์วิทซ์ โดยการไปเก็บข้อมูลเมื่อส่งแบบสอบถามทางไปรษณีย์จากพวกที่ไม่ตอบ แทนที่จะไปเก็บครั้งเดียวเหมือนวิธีแฮนเซ็น-เฮอร์วิทซ์ แต่จะทำการเก็บข้อมูลโดย Mail Survey Subsampling จากพวกที่ไม่ตอบโดยทำหลาย ๆ ครั้ง และครั้งสุดท้ายจะสุ่มตัวอย่างย่อย (Subsampling) จากพวกที่ไม่ตอบในการส่งแบบสอบถามทางไปรษณีย์ครั้งสุดท้าย (ครั้งที่ L) แล้วส่งพนักงานไปสัมภาษณ์หน่วยตัวอย่างย่อยที่ได้ทั้งหมดจากการสุ่มตัวอย่างย่อยครั้งสุดท้ายดังกล่าว ซึ่งขั้นตอนสุดท้ายนี้จะเหมือนกับวิธีแฮนเซ็น-เฮอร์วิทซ์ที่กล่าวมาเบื้องต้น

ขั้นตอนการประมาณค่าพารามิเตอร์ โดยวิธีเอล-บาดรี้ มีดังต่อไปนี้

<u>ขั้นตอนที่</u> 1 สุ่มตัวอย่างมาขนาด n หน่วย จากประชากรที่สนใจขนาด N หน่วย โดยสุ่มตัวอย่างอย่างง่ายแบบไม่ใส่คืน [SRS/W.O.R]

<u>ขั้นตอนที่</u> 2 ขนาดตัวอย่าง n หน่วยที่ได้นี้ จัดส่งแบบสอบถามทางไปรษณีย์ครั้งที่ 1 (L = 1) (the First Mail Attempt) ปรากฏว่า ได้รับคำตอบของตัวแปรที่สนใจคืนมาขนาด n_{11} หน่วย และไม่ตอบขนาด n_{12} หน่วย [Subscript 1 หมายถึงการส่งแบบสอบถามทางไปรษณีย์ครั้งที่ 1 และ 2 หมายถึง กลุ่มที่ไม่ตอบ]

สมมติทราบค่า P_{11}(สัดส่วนการตอบ) ค่าเฉลี่ย μ_{11} และความแปรปรวน S^2_{11}

<u>ขั้นตอนที่</u> 3 สุ่มตัวอย่างย่อย (Subsampling) ขนาด $n'_{12} = k_2 n_{12}$ หน่วย เมื่อ k_2 คือ Subsampling fraction โดยที่ $0 < k_2 < 1$ และสุ่มตัวอย่างแบบ SRS/W.O.R จากตัวอย่างขนาด n_{12}หน่วย ตัวอย่างย่อยที่ได้ขนาด n'_{12} หน่วยนี้ จัดแบบสอบถามทางไปรษณีย์เป็นครั้งที่ 2 (แบบสอบถามชุดเดิม) (L = 2) (the Second Mail Attempt) ปรากฏว่าได้รับคำตอบคืนมาเท่ากับ n_{21} หน่วย และที่ไม่ตอบเป็นครั้งที่สองเท่ากับ n_{22} หน่วย [Subscript-2 ตัวแรก หมายถึงการส่งแบบสอบถามทางไปรษณีย์ ครั้งที่ 2 และ Subscript-2 ตัวหลัง หมายถึงกลุ่มที่ไม่ตอบ]

สมมติทราบค่า P_{21} ,μ_{21}และ S^2_{21}

<u>ขั้นตอนที่ 4</u> สุ่มตัวอย่างย่อยขนาด $n'_{22} = k_3 n_{22}$ หน่วย โดยสุ่มตัวอย่างแบบ SRS/W.O.R จากขนาดตัวอย่างขนาด n_{22} หน่วยซึ่งไม่ตอบจากการส่งแบบสอบถามทางไปรษณีย์ครั้งที่ 2 และตัวอย่างย่อยที่ได้ขนาด n'_{22} หน่วยนี้ จัดส่งแบบสอบถามทางไปรษณีย์เป็นครั้งที่ 3 (L = 3) (the Third Mail Attempt) ปรากฏว่าได้รับคำตอบคืนมาขนาด n_{31}หน่วย และไม่ตอบขนาด n_{32} หน่วย

สมมติทราบค่า P_{31} , μ_{31} และ S^2_{31}

<u>ขั้นตอนที่ 5</u> สุ่มตัวอย่างย่อยจากพวกที่ไม่ตอบในการส่งแบบสอบถามทางไปรษณีย์ครั้งต่อไปเรื่อย ๆ จนกว่าจะถึงการส่งแบบสอบถามทางไปรษณีย์ครั้งที่ L โดยได้รับคำตอบคืนมาขนาด n_{L1} หน่วย และไม่ตอบขนาด n_{L2} หน่วย

สมมติทราบค่า P_{L1} , μ_{L1} และ S^2_{L1}

<u>ขั้นตอนที่ 6</u> สุ่มตัวอย่างย่อยจากตัวอย่างขนาด n_{L2}หน่วย แบบ SRS/W.O.R ด้วยขนาดตัวอย่างเท่ากับ $n'_{(L+1)} = w\, n_{L2}$

โดยที่ w คือ the Interviewed Subsampling Fraction, $0 < w < 1$ ซึ่งขั้นตอนสุดท้ายนี้จะส่งพนักงานไปสัมภาษณ์กลุ่มตัวอย่างย่อยขนาด $n'_{(L+1)}$ ทุก ๆ หน่วย โดยสมมติว่า ตัวอย่างย่อยขนาด $n'_{(L+1)}$ หน่วยนี้ จะต้องตอบคำถามของตัวแปรที่สนใจ โดยตอบมาทั้งหมดไม่มียาดหรือเกิน

สมมติทราบค่า P_{L2} , μ_{L2} และ S^2_{L2}

โดยวิธีเอล-บาตรีนี้สัดส่วนการตอบที่ได้ทั้งหมดคือ $\sum_{i=1}^{L} P_{i1} + P_{L2} = 1$ และความสัมพันธ์ระหว่างสัดส่วนของการตอบหรือไม่ตอบในการส่งแบบสอบถามทางไปรษณีย์แต่ละครั้งเป็นดังนี้คือ $P_{i2} = P_{(i+1)1} + P_{(i+1)2}$, $i = 1,2,3,\ldots, L$

ตัวประมาณค่าของค่าเฉลี่ยประชากร (Population Mean) โดยวิธีเอล-บาตรี จะมีคุณสมบัติดังต่อไปนี้

(1) ความเป็นเส้นตรง (Linearity)

(2) ไม่เอนเอียง (Unbiased) และ

(3) มีความแปรปรวนต่ำสุด (Minimum Variance) ภายใต้เงื่อนไข (1) และ (2)

2.1.2.1 ตัวประมาณที่ไม่เอนเอียงของ $\bar{X}$

เมื่อ $\bar{x} = \frac{1}{n}\left[n_{11}\bar{x}_1 + \frac{n_{21}}{k_2} \cdot \bar{x}_2 + \frac{n_{31}}{k_2 k_3} \bar{x}_3 + \ldots + \frac{n_{L1}}{\prod_{i=2}^{L} k_i} \cdot \bar{x}_{L1} + \frac{n_{L2}}{\prod_{i=2}^{L} k_i} \cdot \bar{x}_{(L+1}\right.$

พิสูจน์

ให้ $\bar{x} = c \sum_{h=1}^{n_{11}} x_{11}^{h} + d \sum_{h=1}^{n_{21}} x_{21}^{h} + \ldots + g \sum_{h=1}^{wn_{L2}} x_{(L+1).1}^{h}$ ......(1)

เมื่อ c, d, ..., g เป็นค่าคงที่ และ

เมื่อ x_{i1}^{h} $(i = 1, 2, \ldots,(L+1))$ แทนค่าสังเกตที่ h ในการไปเก็บข้อมูลในการส่งแบบสอบถามทางไปรษณีย์ครั้งที่ i

$\therefore \; E(\bar{x}) = c\mu_{11}\, En_{11} + d\mu_{21}\, En_{21} + \ldots + g\mu_{L2}\, wEn_{L2}$

$= ncP_{11}\mu_{11} + ndk_2 P_{21}\mu_{21} + \ldots + ngw(\prod_{i=2}^{L} k_i)\, P_{L2}\mu_{L2}$ ....(2)

เนื่องจากเราต้องการให้ $E(\bar{x}) = \mu$ (เป็น Unbiased Estimator ของ μ) .....(3)

เมื่อ $\mu = P_{11}\mu_{11} + P_{21}\mu_{21} + P_{31}\mu_{31} + \ldots + P_{L1}\mu_{L1} + P_{L2}\mu_{L2}$ ....(4)

จากสมการ (2) (3) และ (4) จะได้ว่าค่าคงที่ในสมการที่ (2) จะต้องมีค่าดังต่อไปนี้

$$c = \frac{1}{n}\ ,\ d = \frac{1}{nk_2}\ ,\ e = \frac{1}{nk_2k_3}\ ,\ \ldots,\ f = \frac{1}{n\left(\prod\limits_{i=2}^{L} k_i\right)} \quad \text{และ } g = \frac{1}{nw}\ \frac{1}{\left(\prod\limits_{i=2}^{L} k_i\right)}$$

แทนค่าคงที่ c , d, ..., g ในสมการ (1) จะได้

$$\bar{x} = \frac{1}{n}\sum_{h=1}^{n_{11}} x_{11}^{h} + \frac{1}{nk_2}\sum_{h=1}^{n_{21}} x_{21}^{h} + \frac{1}{nk_2k_3}\sum_{h=1}^{n_{31}} x_{31}^{h} + \ldots +$$

$$\frac{1}{nw\left(\prod\limits_{i=2}^{L} k_i\right)} \sum^{wn_{L2}} x_{(L+1).1}$$

$$\therefore\ \bar{x} = \frac{1}{n}\left[n_{11}\bar{x}_{11} + \frac{n_{21}}{k_2}\bar{x}_{21} + \frac{n_{31}}{k_2k_3}\cdot\bar{x}_{31} + \ldots + \frac{n_{L1}}{\left(\prod\limits_{i=2}^{L} k_i\right)}\cdot\bar{x}_{L1} + \right.$$

$$\left.\frac{n_{L2}}{\left(\prod\limits_{i=2}^{L} k_i\right)}\cdot\bar{x}_{(L+1)1}\right]$$

2.1.2.2 ความแปรปรวนของ $\bar{x}$

$$\bar{x} = \frac{1}{n}\left[n_{11}\bar{x}_{11} + \frac{n_{21}}{k_2}\bar{x}_{21} + \frac{n_{31}}{k_2k_3}\cdot\bar{x}_{31} + \ldots + \frac{n_{L1}}{\left(\prod\limits_{i=2}^{L} k_i\right)}\bar{x}_{L1} + \right.$$

$$\left.\frac{n_{L2}}{\left(\prod\limits_{i=2}^{L} k_i\right)}\cdot\bar{x}_{(L+1).1}\right]$$

$$
\begin{aligned}
&= \frac{1}{n}\Bigg[(n_{11}\bar{x}_{11} + n_{12}\bar{x}_{12}) + \left(\frac{n_{21}}{k_2}\cdot\bar{x}_{21} + \frac{n_{22}}{k_2}\cdot\bar{x}_{22} - n_{12}\bar{x}_{12}\right) + \left(\frac{n_{31}}{k_2k_3}\cdot\bar{x}_{31} + \right. \\
&\qquad \left.\frac{n_{32}}{k_2k_3}\,\bar{x}_{32} - \frac{n_{22}}{k_2}\bar{x}_{22}\right) + \left(\frac{n_{41}}{k_2k_3k_4}\cdot\bar{x}_{41} + \frac{n_{42}}{k_2k_3k_4}\cdot\bar{x}_{41} - \frac{n_{32}}{k_2k_3}\cdot\bar{x}_{32}\right) + \\
&\qquad \ldots + \frac{n_{L2}}{\prod_{i=2}^{L} k_i}\,(\bar{x}_{(L+1).1} - \bar{x}_{L2})\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n}\Bigg[n\bar{x}_{1.} + \left\{\frac{1}{k_2}\,(n_{21}\bar{x}_{21} + n_{22}\bar{x}_{22}) - n_{12}\bar{x}_{12}\right\} + \left\{\frac{1}{k_2}\left(\frac{n_{31}\bar{x}_{31} + n_{32}\bar{x}_{32}}{k_3} - n_{22}\bar{x}_{22}\right)\right\} \\
&\qquad + \left\{\frac{1}{k_2k_3}\left(\frac{n_{41}\bar{x}_{41} + n_{42}\bar{x}_{42}}{k_4} - n_{32}\bar{x}_{32}\right)\right\} + \ldots + \frac{n_{L2}}{\left(\prod_{i=2}^{L} k_i\right)}\,(\bar{x}_{(L+1).1} - \bar{x}_{L2})\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n}\Bigg[n\bar{x}_{1.} + \left\{n_{12}\left(\frac{n_{21}\bar{x}_{21} + n_{22}\bar{x}_{22}}{n_{21} + n_{22}}\right) - n_{12}\bar{x}_{12}\right\} + \left\{\frac{1}{k_2}\,n_{22}\left(\frac{n_{31}\bar{x}_{31} + n_{32}\bar{x}_{32}}{n_{31} + n_{32}}\right)\right. \\
&\qquad \left. - n_{22}\,\bar{x}_{22}\right\} + \ldots + \frac{n_{L2}}{\left(\prod_{i=2}^{L} k_i\right)}\,(\bar{x}_{(L+1).1} - \bar{x}_{L2})\Bigg]
\end{aligned}
$$

$$
= \frac{1}{n}\left[n\bar{x}_{1.} + n_{12}\,(\bar{x}_{2.} - \bar{x}_{12}) + \frac{n_{22}}{k_2}(\bar{x}_{3.} - \bar{x}_{22}) + \ldots + \frac{n_{L2}}{\left(\prod_{i=2}^{L} k_i\right)}\,(\bar{x}_{(L+1)1} - \bar{x}_{L2})\right]
$$

เมื่อ $\bar{x}_{i.}$ $(i = 1,2,\ldots, L)$ คือค่าเฉลี่ยในตัวอย่างที่ i ของกลุ่มที่ตอบและไม่ตอบ

และ $\bar{x}_{i2}$ $(i = 1,2,\ldots, L)$ คือค่าเฉลี่ยในตัวอย่างที่ i ของกลุ่มที่ไม่ตอบ

จาก $V(\bar{x}) = E(\bar{x} - \mu)^2$

$$= E\left[\frac{1}{n^2}(n\bar{x} - n\mu)^2\right]$$

$$= E\left[\frac{1}{n^2}\left\{n(\bar{x}_{1.} - \mu) + n_{12}(\bar{x}_{2.} - \bar{x}_{12}) + \frac{n_{22}}{k_2}(\bar{x}_{3.} - \bar{x}_{22}) + \frac{n_{32}}{k_2 k_3}(\bar{x}_{4.} - \bar{x}_{32}) + \ldots + \frac{n_{L2}}{(\prod_{i=2}^{L} k_i)}(\bar{x}_{(L+1).1} - \bar{x}_{L2})\right\}^2\right]$$

$$= \frac{1}{n^2}\left\{n^2 E(\bar{x}_{1.} - \mu)^2 + En_{12}^2(\bar{x}_{2.} - \bar{x}_{12})^2 + \frac{1}{k_2^2}En_{22}^2 .(\bar{x}_{3.} - \bar{x}_{22})^2 + \frac{1}{k_2^2 k_3^2}E(\bar{x}_{4.} - \bar{x}_{32})^2 + \ldots + \frac{1}{(\prod_{i=2}^{L} k_i)^2} . En_{L2}^2(\bar{x}_{(L+1).1} - \bar{x}_{L2})^2\right\} \qquad \ldots\ldots\ldots\ldots (5)$$

ไม่มีเทอมผลคูณ (Product term) เพราะว่ากลุ่มที่ไม่ตอบในการส่งแบบสอบถามทางไปรษณีย์ครั้งที่ i จะถูกสุ่มตัวอย่างย่อยต่อไปอีกในการส่งแบบสอบถามทางไปรษณีย์ครั้งที่ i + 1จะเห็นว่า อยู่คนละชั้นภูมิกันจึงไม่มีเทอมผลคูณของโควาเรียนซ์ (Covariance)

จากสมการ (5)

พิจารณาเทอม $En_{i2}^2(\bar{x}_{(i+1).} - \bar{x}_{i2})^2$

$$= En_{i2}^2\left[(\bar{x}_{(i+1).} - \mu_{i2}) - (\bar{x}_{i2} - \mu_{i2})\right]^2$$

$$= En_{i2}^2(\bar{x}_{(i+1).} - \mu_{i2})^2 + (-2)\,En_{i2}^2(\bar{x}_{(1+1).} - \mu_{i2}).(\bar{x}_{i2} - \mu_{i2}) + En_{i2}^2(\bar{x}_{i2} - \mu_{i2})^2 \quad \ldots\ldots\ldots (6)$$

จากสมการ (6)

พิจารณาเทอม $-2En_{i2}^2(\bar{x}_{(i+1).}-\mu_{i2})(\bar{x}_{i2}-\mu_{i2})$

$$= -2En_{i2}^2\left[\bar{x}_{(i+1).}\bar{x}_{i2}-\bar{x}_{i2}\,.\,\mu_{i2}-\bar{x}_{(i+1).}\mu_{i2}+\mu_{i2}^2\right]$$

$$= -2En_{i2}^2\,\bar{x}_{(i+1).}\bar{x}_{i2}+2En_{i2}^2\,\bar{x}_{i2}\,\mu_{i2}+2En_{i2}^2\,\bar{x}_{(i+1).}\mu_{i2}-2En_{i2}^2\,\mu_{i2}^2$$

$$= 2En_{i2}^2\,\bar{x}_{i2}\,\mu_{i2}-2En_{i2}^2\,\mu_{i2}^2$$

$$= -2En_{i2}^2(-\bar{x}_{i2}\,\mu_{i2}+\mu_{i2}^2)$$

$$= -2En_{i2}^2\left[(\bar{x}_{i2}^2-2\bar{x}_{i2}\mu_{i2}+\mu_{i2}^2)+(\bar{x}_{i2}\mu_{i2}-\bar{x}_{i2}^2)\right]$$

$$= -2En_{i2}^2(\bar{x}_{i2}-\mu_{i2})^2-2En_{i2}^2(\bar{x}_{i2}\mu_{i2}-\bar{x}_{i2}^2)$$

$$= -2En_{i2}^2(\bar{x}_{i2}-\mu_{i2})^2 \qquad \ldots\ldots\ldots\ldots (7)$$

ดังนั้น $En_{i2}^2(\bar{x}_{(i+1).}-\bar{x}_{i2})^2 = En_{i2}^2(\bar{x}_{(i+1).}-\mu_{i2})^2-2En_{i2}^2(\bar{x}_{i2}-\mu_{i2})^2$

$$+\,En_{i2}^2(\bar{x}_{i2}-\mu_{i2})^2$$

$$= En_{i2}^2(\bar{x}_{(i+1).}-\mu_{i2})^2-En_{i2}^2(\bar{x}_{i2}-\mu_{i2})^2 \ldots (8)$$

จากสมการ (8)

พิจารณาเทอม $En_{i2}^2\,(\bar{x}_{(i+1).} - \mu_{i2})^2 = E\,\{n_{i2}^2\ E_i\ (\bar{x}_{(i+1).} - \mu_{i2})^2\}$

$$= E\,\{n_{i2}^2\ V(\bar{x}_{(i+1).})\}$$

$$= E\left\{n_{i2}^2\ \frac{(N_{i2} - k_{(i+1)}\ n_{i2})}{N_{i2}} \cdot \frac{s_{i2}^2}{(k_{(i+1)}\ n_{i2})}\right\}$$

และ $En_{i2}^2\,(\bar{x}_{i2} - \mu_{i2})^2 = E\{\,n_{i2}^2\ E_{n_{i2}}\ (\bar{x}_{i2} - \mu_{i2})^2\}$

เมื่อ $E_{n_{i2}}$ คือ Expectation เมื่อทราบค่า n_{i2}

$$= E\,\{n_{i2}^2\ V(\bar{x}_{i2})\}$$

$$= E\,\{n_{i2}^2\ \frac{(N_{i2} - n_{i2})}{N_{i2}}\ \frac{s_{i2}^2}{n_{i2}}\}$$

เพราะฉะนั้น เขียนสมการ (8) ใหม่ได้เป็น

$$En_{i2}^2\,(\bar{x}_{(i+1).} - \bar{x}_{i2})^2 = E\left\{n_{i2}^2\ \frac{(N_{i2} - k_{(i+1)}\ n_{i2})}{N_{i2}} \cdot \frac{s_{i2}^2}{(k_{(i+1)}\,n_{i2})}\right\}$$

$$- \quad E\left\{n_{i2}^2\ \frac{(N_{i2} - n_{i2})}{N_{i2}}\ \frac{s_{i2}^2}{n_{i2}}\right\}$$

$$= E\ \frac{n_{i2}^2\ s_{i2}^2}{N_{i2}\ n_{i2}}\left[\frac{(N_{i2} - k_{(i+1)}\ n_{i2})}{k_{(i+1)}} - (N_{i2} - n_{i2})\right]$$

$$= E\,\frac{n_{i2}\, s_{i2}^2}{N_{i2}}\left[\frac{N_{i2} - k_{(i+1)}\, n_{i2} - k_{(i+1)}\, N_{i2} + k_{(i+1)}\, n_{i2}}{k_{(i+1)}}\right]$$

$$= E\,\frac{n_{i2}}{N_{i2}} \cdot s_{i2}^2 \cdot N_{i2}\left[\frac{1 - k_{(i+1)}}{k_{(i+1)}}\right]$$

$$= \left[\frac{1}{k_{(i+1)}} - 1\right] s_{i2}^2\, En_{i2} \quad ; \quad \text{สมมติว่าทราบค่า } s_{i2}^2$$

$$= \left[\frac{1}{k_{(i+1)}} - 1\right] \cdot s_{i2}^2 \cdot (n \prod_{i=2}^{i} k_u)\, P_{i2}\ , \quad i = 1,2,\ldots, L \quad \ldots(9)$$

$$u = 2,3,\ldots, L$$

นำสมการ (9) ไปแทนค่าในสมการ (5) จะได้

$$V(\bar{x}) = \frac{1}{n^2}\left\{ n^2\, V(\bar{x}_{1.}) + \left(\frac{1}{k_2} - 1\right) s_{i2}^2\ nP_{i2} + \frac{1}{k_2^2}\left(\frac{1}{k_3} - 1\right) s_{22}^2\, nk_2\, P_{22} + \frac{1}{k_2^2 k_3^2}\left(\frac{1}{k_4} - 1\right) s_{32}^2\, nk_2k_3P_{32} + \ldots + \frac{1}{(\prod_{i=2}^{L} k_i)^2}\left(\frac{1}{w} - 1\right) s_{L2}^2 \cdot n(\prod_{2}^{L} k_i)\, P_{L2} \right\}$$

$$\therefore\ V(\bar{x}) = \frac{1}{n}\left\{\left(\frac{N - n}{N}\right) . s^2 + \left(\frac{1}{k_2} - 1\right) P_{12}\, s_{12}^2 + \frac{1}{k_2}\left(\frac{1}{k_3} - 1\right) P_{22}\, s_{22}^2 + \frac{1}{k_2k_3}\left(\frac{1}{k_4} - 1\right) P_{32}\, s_{32}^2 + \ldots + \frac{1}{(\prod_{i=2}^{L} k_i)}\left(\frac{1}{w} - 1\right) P_{L2}\, s_{L2}^2\right\} \quad \ldots.(10)$$

จากสมการ (10) กำหนดให้ $R = \frac{V(\bar{x})}{S^2}$ = Relative Variance ของ $\bar{x}$ เทียบกับความแปรปรวนของประชากร (Population Variance) และให้

$$s_i = \frac{P_{i2}\, S_{i2}^2}{S^2} \quad , \quad i = 1,2,\ldots, L$$

เพราะฉะนั้น

$$R = \frac{1}{n}\left[(1-\frac{n}{N}) + \left(\frac{1}{k_2} - 1\right) \frac{P_{12}\, S_{12}^2}{S^2} + \frac{1}{k_2}\left(\frac{1}{k_3} - 1\right) \frac{P_{22}\, S_{22}^2}{S^2} + \ldots + \frac{1}{\left(\prod_{i=2}^{L} k_i\right)}\left(\frac{1}{w} - 1\right) \frac{P_{L2}\, S_{L2}^2}{S^2} \right]$$

$$= \frac{1}{n}\left[(1 - \frac{n}{N}) + (\frac{1}{k_2} - 1)\, s_1 + \frac{1}{k_2}(\frac{1}{k_3} - 1)\, s_2 + \ldots + \frac{1}{\left(\prod_{i=2}^{L} k_i\right)} \cdot (\frac{1}{w} - 1)\, s_L \right]$$

$$= \frac{1}{n}\left[1 - \frac{n}{N} + \frac{s_1}{k_2} - s_1 + \frac{1}{k_2 k_3} \cdot s_2 - \frac{s_2}{k_2} + \ldots + \frac{1}{\left(\prod_{i=2}^{L} k_i\right).w} \cdot s_L - \frac{s_L}{\left(\prod_{i=2}^{L} k_i\right)} \right]$$

$$\therefore R + \frac{1}{N} = \frac{1}{n}\left[(1-s_1) + \left\{ \left(\frac{s_1}{k_2} - \frac{s_2}{k_2}\right) + \left(\frac{s_2}{k_2 k_3} - \frac{s_3}{k_2 k_3}\right) + \left(\frac{s_3}{k_2 k_3 k_4} - \frac{s_4}{k_2 k_3 k_4}\right) + \ldots \right\} + \frac{s_L}{w\left(\prod_{i=2}^{L} k_i\right)} \right]$$

$$= \frac{1}{n}\left[(1 - s_1) + \sum_{j=1}^{(L-1)} \frac{\left(s_j - s_{j+1}\right)}{\left(\prod_{i=2}^{(i+1)} k_i\right)} + \frac{s_L}{w\left(\prod_{i=2}^{L} k_i\right)} \right] \ldots (11)$$

2.1.2.3 การคำนวณหา Subsampling fraction $k_2, k_3, \ldots, k_L$, w ซึ่งจะทำให้ Relative Variance ;R มีค่าต่ำสุด สำหรับค่าใช้จ่ายที่กำหนดให้

จาก Cost $= c_o(n_1 + k_2 n_{12} + k_3 n_{32} + \ldots + k_2 n_{(L-1)2}) + c_1 (n_{11} + n_{21} + n_{31}$

$+ \ldots + n_{(L-1).1}) + c_2 w n_{L2}$

ให้ Expected Cost คือ C = E(cost) สำเหตุที่ใช้ค่าคาดหวังของค่าใช้จ่าย (Expected Cost) นั้น เพราะว่า จำนวนผู้ตอบหรือไม่ตอบในการส่งแบบสอบถามทางไปรษณีย์ในแต่ละครั้งนั้นไม่ทราบล่วงหน้ามาก่อน

$$\therefore \quad C = n_1 c_o \{1 + k_2 P_{12} + k_2 k_3 P_{22} + \ldots + (\prod_{i=2}^{L} k_i) P_{(L-1).2}\} +$$

$$n_1 c_1 \{P_{11} + k_2 P_{21} + k_2 k_3 P_{31} + \ldots + (\prod_{i=2}^{L} k_i) P_{L1}\} + n_1 c_2 (\prod_{i=2}^{L} k_i) w P_{L2}$$

$$= (n_1 c_o + n_1 c_1 P_{11}) + n_1\{ k_2 P_{12} c_o + k_2 k_3 c_o P_{22} + \ldots + c_o(\prod_{i=2}^{L} k_i) P_{(L-1)2}\}$$

$$+ n_1\{ k_2 P_{21} c_1 + k_2 k_3 P_{31} c_1 + \ldots + c_1 (\prod_{i=2}^{L} k_i) P_{L1}\} + n_1 c_2 . (\prod_{i=2}^{L} k_i) w P_{L2}$$

$$= n_1 \{(c_o + c_1 P_{11}) + \sum_{j=1}^{(L-1)} (\prod_{i=2}^{(j+1)} k_i) (c_o P_{j2} + c_1 P_{(j+1).1}) + w (\prod_{i=2}^{L} k_i) c_2 P_{L2} \} \qquad \ldots\ldots\ldots\ldots(12)$$

ให้ $F = (R + \frac{1}{N}) - \lambda(c - T)$ เมื่อ T = Expected Cost

$$\frac{\partial F}{\partial n_1} = -\frac{1}{n_1^2}\left\{(1 - s_1) + \sum_{j=1}^{(L-1)} \frac{(s_j - s_{j+1})}{(\prod_{i=2}^{j+1} k_i)} + \frac{s_L}{w \prod_{i=2}^{L} k_i}\right\} - \lambda\left\{(c_o + c_1 P_{11}) + \sum_{j=1}^{(L-1)} (\prod_{i=2}^{j+1} k_i)(c_o P_{j2} + c_1 P_{(j+1).1}) + w(\prod_{i=2}^{L} k_i)\, c_2 P_{L2}\right\} = 0$$

ให้ $$M = (c_o + c_1 P_{11}) + \sum_{j=1}^{(L-1)} (\prod_{i=2}^{j+1} k_i)(c_o P_{j.2} + c_1 P_{(j+1).1}) + w(\prod_{i=2}^{L} k_i)\, c_2 P_{L2} \qquad \text{.........(13)}$$

และให้ $$L = (1 - s_1) + \sum_{j=1}^{(L-1)} \frac{(s_j - s_{j+1})}{(\prod_{i=2}^{j+1} k_i)} + \frac{s_L}{w \prod_{i=2}^{L} k_i} \qquad \text{.......(14)}$$

$$\therefore \quad \frac{\partial F}{\partial n_1} = -\frac{1}{n_1^2}\{L\} - \lambda M = 0$$

$$\lambda = \frac{1}{n_1^2} \cdot \frac{L}{M} \qquad \text{.......(15)}$$

$$\frac{\partial F}{\partial (w(\prod_{i=2}^{L} k_i))} = -\frac{1}{n_1} \cdot \frac{s_L}{(w \prod_{i=2}^{L} k_i)^2} - \lambda\, n_1 c_2 P_{L2} = 0$$

$$\frac{s_L}{(w \prod_{i=2}^{L} k_i)^2} = \lambda n_1^2\, c_2 P_{L2}$$

$$\therefore \quad \left(w \prod_{i=2}^{L} k_i\right)^2 = \frac{1}{\lambda n_1^2} \cdot \frac{s_L}{c_2 P_{L2}}$$

แทนค่า λ จากสมการ (15) จะได้

$$\left(w \prod_{i=2}^{L} k_i\right)^2 = \frac{1}{\frac{1}{n_1^2} \cdot \frac{L}{M} \cdot n_1^2} \cdot \frac{s_L}{c_2 P_{L2}} = \frac{M}{L} \cdot \frac{s_L}{c_2 P_{L2}} \quad \ldots(16)$$

และ

$$\frac{\partial F}{\partial\left(\prod_{i=2}^{j+1} k_i\right)} = -\frac{1}{n_1} \sum_{j=1}^{(L-1)} \frac{(s_j - s_{j+1})}{\left(\prod_{i=2}^{j+1} k_i\right)^2} - \lambda n_1 \sum_{j=1}^{(L-1)} (c_o P_{j2} +$$

$$c_1 P_{(j+1).1}) = 0$$

$$\therefore \quad \frac{(s_j - s_{j+1})}{\left(\prod_{i=2}^{j+1} k_i\right)^2} = \lambda n_1^2 \left(c_o P_{j2} + c_1 P_{(j+1).1}\right)$$

แทนค่า λ จากสมการ (15) จะได้

$$\left(\prod_{i=2}^{j+1} k_i\right)^2 = \frac{M}{L} \frac{(s_j - s_{j+1})}{(c_o P_{j2} + c_1 P_{(j+1).1})}, \quad j=1,1,2,\ldots, L-1 \quad \ldots(17)$$

พิจารณา $\frac{M}{L}$

จากสมการ (17) เขียนใหม่ได้เป็น

$$\left(\prod_{i=2}^{j+1} k_i\right) (c_o P_{j2} + c_1 P_{(j+1).1}) = \frac{M}{L} \frac{(s_j - s_{j+1})}{\left(\prod_{i=2}^{j+1} k_i\right)}, \quad j = 1,2,\ldots, L-1 \quad \ldots.(18)$$

จากสมการ (13) และ (14) จัดเป็น

$$(w \prod_{i=2}^{L} k_i)(c_2 P_{22}) \quad = \quad \frac{M}{L} \cdot \frac{s_L}{c_2 P_{22}} \quad \text{.........(19)}$$

take $\sum_{j}^{(L-1)}$ ในสมการ (18) แล้วนำมาบวกกับสมการ (19) จะได้

$$\sum_{j}^{(L-1)} (\prod_{i=2}^{j+1} k_i)(c_o P_{j2} + c_1 P_{(j+1).1}) + (w \prod_{i=2}^{L} k_i)(c_2 P_{L2})$$

$$= \frac{M}{L} \left\{ \sum_{j}^{(L-1)} \frac{(s_j - s_{j+1})}{(\prod_{i=2}^{j+1} k_i)} + \frac{s_L}{c_2 P_{L2}} \right\} \quad \text{.........(20)}$$

จากสมการที่ (20) จะได้

$$\frac{M}{L} = \frac{\sum_{j=1}^{(L-1)} (\prod_{i=2}^{j+1} k_i)(c_o P_{j2} + c_1 P_{(j+1).1}) + (w \prod_{i=2}^{L} k_i)(c_2 P_{L2})}{\sum_{j=1}^{(L-1)} \frac{(s_j - s_{j+1})}{(\prod_{i=2}^{j+1} k_i)} + \frac{s_L}{c_2 P_{L2}}} \quad \text{..(21)}$$

และเอาสมการ (13)หารด้วยสมการ (14) จะได้

$$\frac{M}{L} = \frac{(c_o + c_1 P_{11}) + \sum_{j}^{(L-1)} (\prod_{i=2}^{j+1} k_i)(c_o P_{j2} + c_1 P_{(j+1).1}) + w(\prod_{i=2}^{L} k_i)\, c_2 P_{L2}}{(1 - s_1) + \sum_{j}^{(L-1)} \frac{(s_j - s_{(j+1)})}{(\prod_{i=2}^{j+1} k_i)} + \frac{s_L}{c_2 P_{L2}}} \quad \text{..(22)}$$

จะเห็นว่าสมการ (21) เท่ากับ สมการ (22)

ดังนั้นจะได้ค่า $\frac{M}{L} = \frac{c_o + c_1P_{11}}{(1 - s_1)}$ ด้วย ....(23)

แทนค่า $\frac{M}{L}$ ของสมการ (23) ในสมการ (17) จะได้

$$\left(\prod_{i=2}^{j+1} k_i\right)^2 = \frac{(s_j - s_{j+1})}{(c_oP_{j2}+c_1P_{(j+1).1})} \frac{(c_o + c_1P_{11})}{(1 - s_1)} \quad , j=1,2,\ldots,(L-1) \qquad (24)$$

และแทนค่า $\frac{M}{L}$ ของสมการ (23) ในสมการ (16) จะได้

$$\left(w \prod_{i=2}^{L} k_i\right)^2 = \frac{s_L}{c_2P_{L2}} \frac{(c_o + c_1P_{11})}{(1 - s_1)}$$

ดังนั้น Subsampling Fraction เมื่อ L > 1 คือ

$$k_2^2 = \frac{(s_1 - s_2)}{(1 - s_1)} \frac{(c_o + c_1P_{11})}{(c_oP_{12}+c_1P_{21})}$$

$$k_u^2 = \frac{(s_{(u-1)} - s_u)}{(s_{(u-2)}-s_{(u-1)})} \cdot \frac{(c_oP_{(u-2).2}+ c_1P_{(u-1).1})}{(c_oP_{(u-1).2}+ c_1P_{u1})} \quad , u=3,4,\ldots,L \qquad (25)$$

$$w^2 = \frac{s_L}{(s_{(L-1)}-s_L)} \frac{(c_oP_{(L-1).2}+ c_1P_{L1})}{c_2P_{L2}}$$

2.1.2.4 การหาขนาดตัวอย่างที่เหมาะสมที่สุด (n_{opt})

จากสมการ (11)

$$n = \left(R + \frac{1}{N}\right)^{-1} \left\{ (1 - s_1) + \sum_{j=1}^{(L-1)} \frac{(s_j - s_{(j+1)})}{\left(\prod_{i=2}^{j+1} k_i\right)} + \frac{s_L}{w \prod_{i=2}^{L} k_i} \right\}$$

แทนค่า $\prod_{i=2}^{j+1} k_i$ ของสมการที่ (23) จะได้

$$n = \left(R + \frac{1}{N}\right)^{-1} \left\{ (1-s_1) + \sum_{j=1}^{(L-1)} \frac{(s_j - s_{j+1})(1-s_1)^{1/2} (c_o P_{j2} + c_1 P_{(j+1).1})^{1/2}}{(s_j - s_{j+1})^{1/2} (c_o + c_1 P_{11})^{1/2}} \right.$$

$$\left. + \frac{s_L \cdot (1-s_1)^{1/2} (c_2 P_{L2})^{1/2}}{s_L^{1/2} (c_o + c_1 P_{11})^{1/2}} \right\}$$

ดังนั้น

$$n_{opt} = \left(R + \frac{1}{N}\right)^{-1} \left(\frac{1 - s_1}{c_o + c_1 P_{11}}\right)^{1/2} \left[\left\{ (1 - s_1)(c_o + c_1 P_{11}) \right\}^{1/2} \right.$$

$$+ \sum_{j=1}^{(L-1)} \left\{ (c_o P_{j2} + c_1 P_{(j+1).1})(s_j - s_{(j+1)}) \right\}^{1/2} .$$

$$\left. + \left\{ c_2 P_{L2} \cdot s_L \right\}^{1/2} \right] \qquad \ldots\ldots\ldots\ldots\ldots(26)$$

เมื่อ $R = \dfrac{\varepsilon^2}{S^2}$

2.1.2.5 ข้อสังเกตบางประการเกี่ยวกับค่า Subsampling fraction ($k_2,\ldots,w$) ที่เหมาะสมที่สุด

(1) ค่า Subsampling Fraction $k_2, k_3, \ldots, k_L$, w ไม่ขึ้นอยู่กับค่า n ดังนั้น ผลจากการได้ค่า k, w ที่เหมาะสมที่สุด โดยวิธีดังกล่าวจะเหมือนกับการหาค่า k, w ที่เหมาะสมที่สุด โดยวิธีการทำให้ผลคูณของค่าใช้จ่ายกับความแปรปรวนที่ไม่คิดค่า n มีค่าต่ำสุด คุณสมบัติข้อนี้สำคัญคือ ถ้าข่าวสาร (Information) ที่ต้องการสำหรับการคำนวณค่า n ในสมการที่ (26) ไม่สามารถหาได้ (not available) เราสามารถที่จะเดา (guess) ค่า n ได้ โดยที่การหาค่า k, w ที่เหมาะสมที่สุดยังคงทำให้ผลคูณของค่าใช้จ่ายกับความแปรปรวนมีค่าต่ำสุด อยู่เหมือนเดิม [MA. EL-BADRY 1956 : 220]

(2) The Mail Subsampling Fraction ($k_2,\ldots, k_L$) ไม่ได้ขึ้นอยู่กับจำนวนครั้งของการส่งแบบสอบถามทางไปรษณีย์ (L) ดังนั้นจำนวนครั้งของการไปเก็บข้อมูลโดยส่งแบบสอบถามทางไปรษณีย์ (Mail Attempt) จึงไม่ได้กำหนดไว้ให้คงที่เอาไว้ล่วงหน้า โดยที่ผู้วิจัยจะส่งแบบสอถบามทางไปรษณีย์ต่อไปเรื่อย ๆ จนกว่าจะได้ precision ตามที่ต้องการ และในขั้นตอนสุดท้ายจึงจะสัมภาษณ์กลุ่มตัวอย่างที่เลือกมาก่อนหน้านั้น

2.1.2.6 การพิจารณาจำนวนครั้งของการส่งแบบสอบถามทางไปรษณีย์ที่เหมาะสมที่สุด

จะศึกษาดูว่าจำนวนครั้งที่จะไปเก็บข้อมูลโดยการส่งแบบสอบถามทางไปรษณีย์จากการที่ไม่ตอบนั้นควรจะทำกี่ครั้งจึงจะทำให้ได้ค่า precision ตามต้องการ

วิธีการหาจำนวนครั้งของการส่งแบบสอบถามทางไปรษณีย์ที่เหมาะสมที่สุดนั้น จะพิจารณาดังต่อไปนี้คือ จากความจริงที่ว่า เมื่อค่าใช้จ่ายในการสำรวจทั้งหมด (Total Cost of the Investigations) ได้กำหนดไว้มีค่าคงที่ค่าหนึ่งหรือจำนวนหนึ่งภายใต้ Fixed Total Cost นั้นค่าประมาณของ $\bar{X}$ ควรจะมีความแปรปรวนน้อยที่สุดเท่าที่จะกระทำได้ หรือถ้ากำหนด Precision ของตัวประมาณค่ากำหนดไว้คงที่แล้ว จำนวนครั้งของการส่งแบบสอบถามทางไปรษณีย์ จะทำให้ค่าใช้จ่ายทั้งหมดมีค่าต่ำสุด

ดังนั้นจากสมการที่ (26) เมื่อแทนค่า n_{opt} ของสมการ (26) ลงในสมการ Expected Cost จะได้ดังนี้คือ

$$\text{Expected Cost} = T = \left(R + \frac{1}{N}\right)^{-1} \left[\left\{(1 - s_1)\,(c_o + c_1 P_{11})\right\}^{1/2} + \sum_{j=1}^{(L-1)} \left\{(c_o P_{j2} + c_1 P_{(j+1).1})\,(s_j - s_{j+1})\right\}^{1/2} + \left\{c_2 P_{L2} s_L\right\}^{1/2} \right]^2 \quad \text{..........(27)}$$

จากสมการ (27) จะได้ Relative Variance ต่ำสุดคือ

$$\left(R + \frac{1}{N}\right) = T^{-1} \left[\left\{(1 - s_1)\,(c_o + c_1 P_{11})\right\}^{1/2} + \sum_{j=1}^{(L-1)} \left\{(c_o P_{j2} + c_1 P_{(j+1).1})\,(s_j - s_{j+1})\right\}^{1/2} + \left\{c_2\, P_{L2} s_L\right\}^{1/2} \right]^2 \quad \text{...........(28)}$$

จากสมการ (27) และ (28) แสดงให้เห็นถึงการที่ลด Expected Cost เมื่อกำหนด Precision (R) ให้ และต้องการเพิ่ม Precision เมื่อกำหนด Cost ให้ ตามลำดับ

ดังนั้นจะทำการส่งแบบสอบถามทางไปรษณีย์ครั้งที่ L+1 ต่อไป หลังจากได้ส่งแบบสอบถามทางไปรษณีย์ครั้งที่ L ไปแล้ว ถ้า

$$\left\{(c_o P_{L2} + c_1 P_{(L+1).1})\,(s_L - s_{L+1})\right\}^{1/2} + \left\{c_2 P_{(L+1).2}\, s_{(L+1)}\right\}^{1/2} < \left\{c_2 P_{L2} s_L\right\}^{1/2}$$

$$\left\{ c_o P_{L2} + c_1 P_{(L+1).1} \right\}^{1/2} . \left\{ \frac{P_{L2}\, S^2_{L2}}{S^2} - \frac{P_{(L+1).2}\, S^2_{(L+1).2}}{S^2} \right\}^{1/2} + \left\{ c_2 P_{(L+1).2} \cdot \frac{P_{(L+1).2}\, S^2_{(L+1).2}}{S^2} \right\}^{1/2} < \left\{ c_2 P_{L2} \cdot \frac{P_{L2} \,.S^2_{L2}}{S^2} \right\}^{1/2}$$

$$\left\{ (c_o P_{L2} + c_1 P_{(L+1).1} \right\}^{1/2} \cdot \frac{(P_{L2} S_{L2}^2 - P_{(L+1).2} S_{(L+1).2}^2)^{1/2}}{S}$$

$$+ \frac{c_2^{1/2} \cdot P_{(L+1).2} S_{(L+1).2}}{S} < \frac{c_2^{1/2} \cdot P_{L2} \cdot S_{L2}}{S}$$

$$\left\{ \frac{c_o P_{L2} + c_1 P_{(L+1).1}}{c_2} \right\}^{1/2} \left\{ \frac{P_{L2} S_{L2}^2 - P_{(L+1).2} S_{(L+1).2}^2}{P_{L2}^2 S_{L2}^2} \right\}^{1/2}$$

$$+ \frac{P_{(L+1).2}}{P_{L2}} \cdot \frac{S_{(L+1).2}}{S_{L2}} < 1$$

$$\left\{ 1 - \frac{P_{(L+1).2}}{P_{L2}} \cdot \frac{S_{(L+1).2}}{S_{L2}} \right\} > \left\{ \frac{c_o P_{L2} + c_1 P_{(L+1).1}}{c_2} \right\}^{1/2} \cdot$$

$$\left\{ \frac{P_{L2} S_{L2}^2 - P_{(L+1).2} S_{(L+1).2}^2}{P_{L2}^2 S_{L2}^2} \right\}^{1/2}$$

จาก $\quad P_{L2} = P_{(L+1).1} + P_{(L+1).2}$

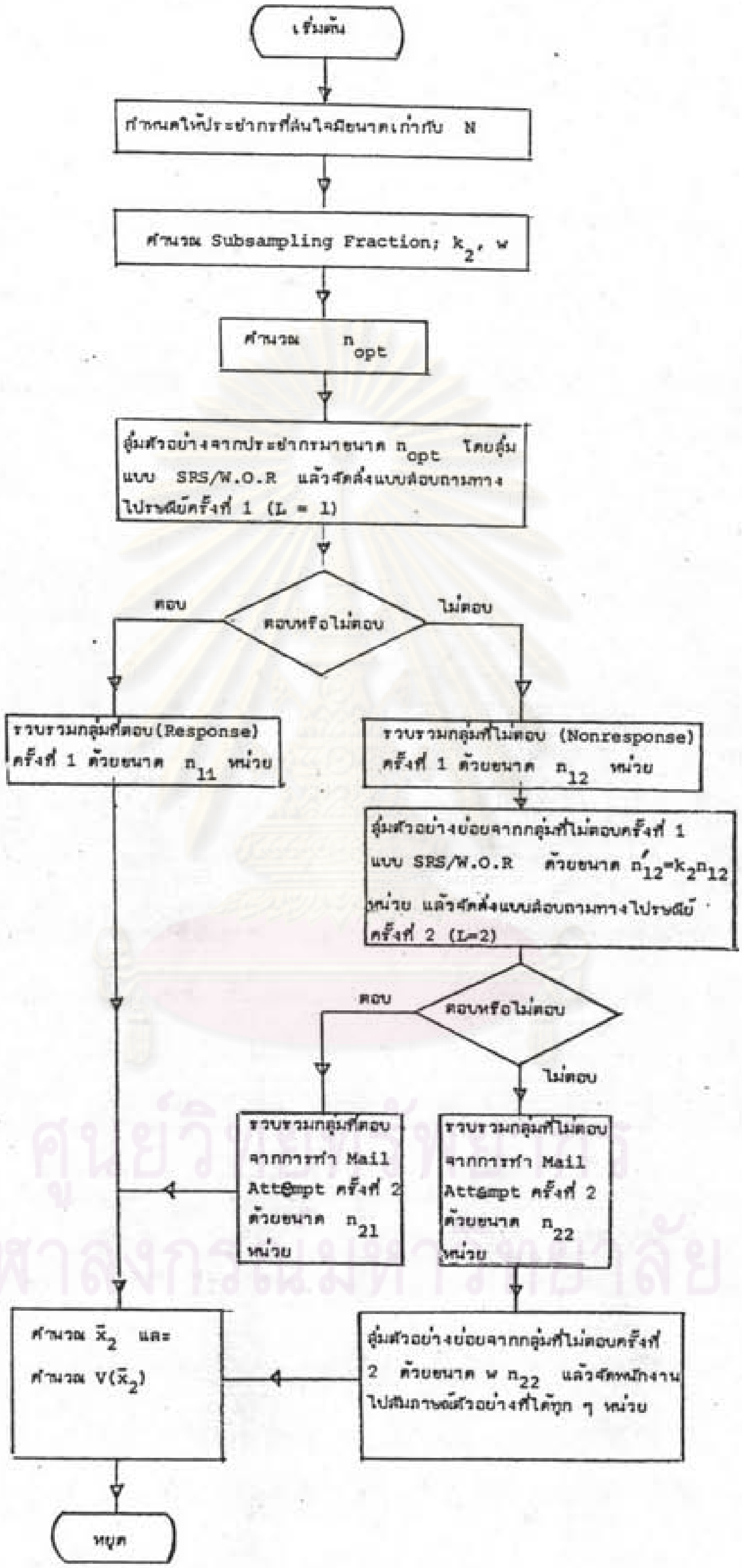
ดังนั้น $\left\{ 1 - \frac{P_{(L+1).2}}{P_{L2}} \cdot \frac{S_{(L+1).2}}{S_{L2}} \right\}^2 > \left\{ 1 - \frac{P_{(L+1).2}}{P_{L2}} \left(\frac{S_{(L+1).2}}{S_{L2}} \right)^2 \right\} \cdot$

$$\left\{ \frac{(c_o + c_1)}{c_2} - \frac{c_1}{c_2} \frac{P_{(L+1).2}}{P_{L2}} \right\} \qquad \text{..........(29)}$$

เพราะฉะนั้น จำนวนครั้งของการส่งแบบสอบถามทางไปรษณีย์ที่เหมาะสมที่สุดคือ ค่าที่ทำให้อสมการ (29) เป็นจริง

แต่ในการวิจัยนี้ กำหนดให้ L = 3 เท่านั้น ซึ่ง Kish (1655 : 539) ได้กล่าวว่า "การที่จะให้ได้คำตอบเพิ่มขึ้นจากการไม่ตอบแบบสอบถามทางไปรษณีย์นั้นควรจะทำการส่งแบบสอบถามทางไปรษณีย์ 3 หรือ 4 ครั้ง จึงจะทำให้ได้อัตราการตอบที่ต้องการมากกว่า 80 เปอร์เซ็นต์หรือมากกว่า 90 เปอร์เซ็นต์ และการทำ Interview Followup จากกลุ่มตัวอย่างที่ไม่ตอบแบบสอบถามทางไปรษณีย์นั้นจะทำให้ได้อัตราการตอบเพิ่มขึ้น" ดังนั้นการวิจัยในครั้งนี้ L = 3 เท่านั้น

แผนผังแสดงขั้นตอนของวิธีเอล-บาดรี้เมื่อ L=2



แผนผังที่ 2

2.2 ผลงานวิจัยที่เกี่ยวข้อง

สำหรับผลงานวิจัยที่เกี่ยวข้องนั้น ได้มีนักวิจัยโดยเฉพาะนักวิจัยทางด้านสังคมศาสตร์เป็นจำนวนมากที่ได้ศึกษาถึงปัญหาการเกิด การไม่ตอบ (Nonresponse) ในการเก็บข้อมูลทางไปรษณีย์ (Mailed Survey)

เนื่องจากการส่งแบบสอบถามทางไปรษณีย์นั้นจะประหยัดค่าใช้จ่ายในการสำรวจตัวอย่างมากกว่าการสำรวจตัวอย่างโดยการสัมภาษณ์และการส่งแบบสอบถามทางไปรษณีย์นั้นจะกระทำได้สะดวกกว่าและมีความถูกต้องของข้อมูลมากกว่า การที่ไปเก็บข้อมูลโดยการสัมภาษณ์หรือทางโทรศัพท์ เพราะว่าการส่งแบบสอบถามทางไปรษณีย์นั้น ทำให้ผู้ตอบสามารถตรวจสอบคำถามได้อย่างอิสระหรือสามารถปรึกษาข้อถามกับคนอื่น ๆ ได้(Nuckols, Robert C 1964 : 11-16) แต่ว่า

การเก็บข้อมูลโดยส่งแบบสอบถามทางไปรษณีย์นั้น ก็มีข้อเสียที่สำคัญ ๆ คือ ส่วนมากจะได้อัตราการตอบของข้อถามบางข้อ หรือส่งคำถามคืนมาทั้งฉบับน้อยมาก จึงทำให้เกิดปัญหาความเอนเอียง (Bias) ของตัวประมาณค่าได้ (Deming W. Edward 1953 : 743)

เมื่อเกิดปัญหาดังกล่าวจึงได้มีผู้คิดค้นวิจัย พยายามที่จะเพิ่มอัตราการตอบ (Respanse Rate) ในการส่งแบบสอบถามทางไปรษณีย์ โดยเขียนเป็นรายงานการวิจัยไว้ ซึ่งจะพิจารณาได้ 2 กรณีคือ

2.2.1 กรณีที่ต้องอาศัยเวลา (Timing) ในการที่จะได้อัตราการตอบเพิ่มขึ้น เช่น วิธี Follow-up Technique

วิธี Follow-up Technique หมายถึง วิธีการที่จะให้ได้ อัตราการตอบเพิ่มขึ้น ซึ่งใช้มากในการเก็บข้อมูลทางไปรษณีย์ โดยการกลับไปถามผู้ที่ไม่ตอบหลาย ๆ ครั้งจนกว่าจะได้คำตอบหรืออัตราการตอบเป็นที่น่าพอใจซึ่ง Kish(1655 : 539) ได้กล่าวว่า "การที่จะให้ได้อัตราการตอบเพิ่มขึ้นจากการที่ไม่ตอบแบบสอบถามทางไปรษณีย์นั้นควรจะทำ Mail Follow-up 3 หรือ 4 ครั้ง จึงจะทำให้ได้อัตราการตอบที่ต้องการมากกว่า 80 เปอร์เซ็นต์ หรือมากกว่า 90 เปอร์เซ็นต์ และกระทำ Interview Follow-up จากกลุ่มตัวอย่างที่ไม่ตอบแบบสอบถามทางไปรษณีย์นั้น จะทำให้ได้อัตราการตอบเพิ่มขึ้น" และสก๊อต (Scott 1961 : 164) ได้ทำ

Pilot Survey เกี่ยวกับการบริการในการใช้โทรศัพท์ (a Telephone Survice Pilot Servey) โดยใช้ขนาดตัวอย่างเท่ากับ 1,050 หน่วย และทำ Mailed Follow-up 2 ครั้ง โดยครั้งแรกได้รับคำตอบคืนมา 74.80 เปอร์เซ็นต์ หลังจากนั้น 6 สัปดาห์ ส่งแบบสอบถามให้อีก ปรากฏว่าได้รับคำตอบคืนมา 95.60 เปอร์เซ็นต์ สำหรับ คลอสเซ่นและฟอร์ด (Clausen & Ford 1947 : 497-599) พบว่า "สัดส่วนของการได้รับคำตอบกลับคืนมาจากการทำ Follow-up 1 ครั้ง จะให้ผลเหมือนกันกับการส่งคำถามทางไปรษณีย์ครั้งแรก"

2.2.2 กรณีที่ต้องอาศัยเทคนิคบางอย่างที่จะทำให้ได้อัตราการตอบ (Response Rate) เพิ่มขึ้น เช่น

2.2.2.1 เทคนิคการจัดความเหมาะสมของความยาวของแบบสอบถาม

การจัดรูปแบบความยาวของแบบสอบถามให้เหมาะสมนั้น จะมีผลต่อการได้รับคำตอบคืนมามากกว่าเดิมได้ แต่อย่างไรก็ตาม ผลงานวิจัยที่จะมาสอดคล้องในเรื่องนี้มีน้อยมาก เช่น สก๊อต(Scott 1961 : 167) ได้ศึกษาถึงจำนวนผู้ใช้วิทยุและโทรทัศน์ โดยใช้ประชากรขนาด 4536 คน โดยจัดแบบสอบถามออกเป็น 2 แบบ กล่าวคือ แบบที่หนึ่งเป็นแบบสอบถามแบบยาว (รวมเอาแบบสั้น 2 ฉบับ รวมไว้ในฉบับเดียวกันเป็นแบบยาว 1 ฉบับ) แบบที่สอง เป็นแบบสอบถามแบบสั้น หรือเป็นแบบสอบถามที่มีเนื้อหาที่จะถามน้อย เมื่อนำแบบสอบถามแบบสั้น 2 ฉบับที่ใช้คำถามต่างกันไปถามตัวอย่างที่สุ่มได้ ปรากฏว่า ได้อัตราการตอบเท่ากับ 90.50 เปอร์เซ็นต์ และนำแบบสอบถามแบบยาว 1 ฉบับ นำไปถามตัวอย่างชุดเดียวกัน ปรากฏว่าได้รับคำตอบคืนมา 89.60 เปอร์เซ็นต์ ซึ่งผลที่ได้ดังกล่าวนี้พบว่าไม่แตกต่างกัน หรือไม่มีระดับนัยสำคัญ (Not Significant) ดังนั้นจากผลการวิจัยดังกล่าวไม่ได้มาสนับสนุนข้อความที่ว่า "แบบสอบถามแบบสั้น จะได้อัตราการตอบมากกว่าแบบสอบถามแบบยาว" เสมอไป

2.2.2.2 เทคนิคในการหาผู้มาสนับสนุนในด้านค่าใช้จ่ายให้พอเพียง (Sponsorship)

เทคนิคนี้มีความสำคัญมาก ถ้าการสำรวจตัวอย่างในเรื่องที่เราสนใจนั้นจะต้องเสียค่าใช้จ่ายมาก ซึ่งสก๊อต (Scott 1961 : 168) ได้กล่าวว่า "การที่จะให้ได้รับอัตราการตอบมากขึ้นจากการสำรวจตัวอย่างเกี่ยวกับ British Social Survey โดยการส่งแบบสอบถามทางไปรษณีย์นั้น รัฐบาลควรจะเป็นผู้สนับสนุนในด้านค่าใช้จ่าย เพราะว่าถ้ารัฐบาลเข้ามาเกี่ยวข้องในเรื่องการสำรวจในระดับประเทศแล้วละก็จะทำให้ผู้ตอบมองเห็นความ

สำคัญในข้อนี้ซึ่งจะมีผลทำให้ได้รับคำตอบคืนมามากขึ้น"

2.2.2.3 เทคนิคในการบรรจุซองตอบรับพร้อมติดแสตมป์ให้แก่ผู้ตอบ (Return Envelopes)

เป็นเทคนิคที่ผู้วิจัย นำซองจดหมายตอบรับพร้อมกับติดแสตมป์ให้แก่ผู้ตอบเพื่อ ให้ผู้ตอบคืนแบบสอบถามที่ตอบแล้วให้แก่ผู้วิจัย ซึ่งวิธีนี้เป็นวิธีที่ใช้กันมากในทางปฏิบัติโดยทั่ว ๆ ไปสำหรับการทำ Mailed Survey (SCOTT 1961 : 170-172)

2.2.2.4 เทคนิคแอนนอนนิมมิติ (Anonymity)

หมายถึง การที่แบบสอบถามที่ใช้ในการวิจัย จะต้องไม่ถามชื่อของผู้ตอบ หรือไม่ให้ผู้ตอบเขียนลายเซ็นลงในแบบสอบถาม ซึ่งจะทำให้ผู้ตอบมีทัศนคติที่ไม่ดีต่อการตอบแบบสอบถาม เพราะว่าจะทำให้ผู้ตอบไม่กล้าที่จะตอบแบบสอบถาม ซึ่งผู้ตอบกลัวว่า ถ้าตอบไปแล้ว เกรงจะมีผลทางด้านกฎหมาย เช่น กลัวสรรพากร ทราบรายได้ที่แท้จริงของตนเองเป็นต้น จนเป็นเหตุทำให้อัตราการตอบลดลงได้ แต่ในบางกรณีอาจทำให้อัตราการตอบสูงขึ้นก็ได้ เช่น ค็อซ แอนเดอร์สันและฟุลเช่อร์ (Cox, Anderson and Fulcher 1974 : 413) ได้ศึกษาถึงเรื่อง Anonymity โดยใช้ประชากรคือ จำนวนผู้ออกเงินบำรุงเกี่ยวกับกิจการขององค์การโทรศัพท์ ขนาด 4,000 คน โดยส่งแบบสอบถามให้กับตัวอย่างที่สุ่มได้ โดยจัดแบบสอบถามออกเป็น 2 แบบคือ แบบที่หนึ่งจะถามชื่อและที่อยู่ลงในตัวแบบสอบถามด้วย แต่แบบที่สองเป็นแบบสอบถามที่ไม่ได้ถามชื่อและที่อยู่ลงในตัวแบบสอบถามเลย เมื่อนำแบบสอบถามทั้งสองแบบไปถามตัวอย่างชุดเดียวกัน แต่ใช้เวลาต่างกัน ปรากฏว่า คำตอบจากแบบสอบถามแบบที่หนึ่งจะได้รับคำตอบคืนมา 21.5 เปอร์เซ็นต์ ขณะที่แบบสอบถามแบบที่สองได้รับคำตอบคืนมาเพียง 14.5 เปอร์เซ็นต์เท่านั้น ซึ่งผลการทดลองความแตกต่างดังกล่าว ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญ" แสดงให้เห็นว่าผู้ตอบ (Respondent) ต้องการให้คนอื่นรู้ว่าตนเองมีส่วนรวมในด้านความช่วยเหลือกิจการขององค์-การโทรศัพท์ อยากจะให้ตนเองมีชื่อเสียงในสังคม จึงทำให้อัตราการตอบแบบสอบถามแบบที่หนึ่งมีมากกว่าความที่สองได้

2.2.2.5 เทคนิคเกี่ยวกับการกำหนดรูปเล่มและสีของแบบสอบถาม

สก็อต (Scott 1961 : 177) ได้สำรวจเกี่ยวกับ Motor Cyclists เพื่อที่จะศึกษาถึงข้อมูลเกี่ยวกับอายุ และประสบการณ์ของเจ้าของจักรยานยนต์ โดยใช้แบบสอบถาม 2 แบบ กล่าวคือ แบบที่หนึ่งจะเย็บเล่มและพิมพ์แบบสอบถาม ทำเป็นรูปเล่ม

อย่างดี แต่แบบที่สองจะไม่เย็บเป็นรูปเล่ม แต่จะเป็นลักษณะของแผ่นชีท (Sheet) และเย็บติดกัน เมื่อนำแบบสอบถามทั้ง 2 แบบ ไปถามตัวอย่างชุดเดียวกัน แต่ต่างเวลากัน ปรากฏว่าแบบที่หนึ่งได้รับคำตอบคืนมา 95.8 เปอร์เซ็นต์ สำหรับแบบที่สองจะได้รับคำตอบคืนมาเพียง 93.1 เปอร์เซ็นต์ และพบว่า แตกต่างกันอย่างมีนัยสำคัญ (Significant) และสก๊อตยังได้กล่าวอีกว่า "สำหรับการวิจัยของ กัลลาฮอร์น (Gullahorn) พบว่ากระดาษที่ใช้ในแบบสอบถามแตกต่างกันนั้น จะไม่ทำให้ได้อัตราการตอบแบบสอบถามต่างกัน"

2.2.2.6 เทคนิคในการให้รางวัล (Reward) ให้กับผู้ที่ตอบแบบสอบถามคืนมา

เทคนิคนี้เป็นเทคนิคที่จะกระตุ้นให้ผู้ตอบอยากตอบคำถามได้เหมือนกัน โดยที่ผู้ตอบอยากจะได้รางวัลต่าง ๆ หรือได้ของสมนาคุณต่าง ๆ แต่ว่าเทคนิคนี้จะใช้ได้ดีนั้นจะต้องมีทุนพอที่จะทำกรณีเช่นนี้ได้โดยหวังผลที่จะได้รับแบบสอบถามที่ตอบแล้วคืนมาให้มาก ๆ และดูเหมือนว่า ถ้าใช้วิธีนี้แล้วจะมีแต่พวกจน ๆ เท่านั้นที่จะส่งแบบสอบถามคืนมา ซึ่งตามความเป็นจริงแล้ว คนที่มีฐานะดีก็จะตอบคำถามคืนมาเหมือนกัน (Leslie และ Conrad (1975 : 447) และยังได้มีนักวิจัยชื่อ วอทรูบา (Watruba 1966 : 398 - 400) ได้ศึกษาถึงอิทธิพลการกำหนดรางวัลหรือของแถมต่าง ๆ เมื่อมีผู้ตอบแบบสอบถามส่งคืนมาว่าจะแตกต่างจากการที่ผู้วิจัยไม่ได้กำหนดรางวัลหรือไม่ ปรากฏว่าได้อัตราการตอบแบบสอบถามคืนมา แตกต่างกันอย่างมีนัยสำคัญ กล่าวคือเมื่อไม่ได้กำหนดรางวัลจะได้รับคำตอบคืนมา 18 เปอร์เซ็นต์ แต่ถ้ากำหนดให้รางวัล จะได้รับคำตอบคืนมาถึง 40 เปอร์เซ็นต์

2.2.2.7 เทคนิคในการกำหนดวันส่งแบบสอบถามคืนมา (Deadline Dates)

สก๊อต (Scoot 1961 : 175) กล่าวว่า เมื่อใช้สติ๊กเกอร์ (Sticker) ที่มีข้อความว่า "เมื่อได้รับแบบสอบถามแล้วขอความกรุณาให้ตอบแบบสอบถามทุกข้อแล้วส่งแบบสอบถามที่ตอบแล้วคืนมาภายในวันและเวลาที่กำหนด" โดยติดไว้ที่แบบสอบถามผลจากการกระทำดังกล่าวพบว่าจะมีผลขั้นต้นในการที่จะกระตุ้นให้ผู้ตอบส่งแบบสอบถามที่ตอบแล้วส่งคืนมา

จะเห็นว่านักวิจัยต่าง ๆ ที่กล่าวมาเบื้องต้นนั้นได้วิจัยถึงการที่จะให้ได้อัตราการตอบ (Response Rate) แบบสอบถามคืนมาให้มากที่สุด แต่ยังไม่มีผู้ใดกล่าวถึงวิธีการปรับความเอนเอียง (Bias) ของตัวประมาณค่าของพารามิเตอร์ในกรณีที่เกิดการไม่ตอบ

(Nonresponse) จากการสำรวจตัวอย่างโดยส่งแบบสอบถามทางไปรษณีย์เลย ซึ่งต่อมาก็ได้มีนักวิจัยหลายคนได้ศึกษาถึงวิธีการปรับความเอนเอียงของตัวประมาณค่าของพารามิเตอร์ในกรณีที่เกิดการไม่ตอบดังกล่าว เช่น ฟิลเลี่ยน (Filian 1975 : 482 - 492) ได้ใช้สมการถดถอยเชิงเส้น (Linear Regression) ในการประมาณพารามิเตอร์ของประชากร (Population Paramiters) ในกรณีที่เกิดความเอนเอียงอันเนื่องมาจากการไม่ตอบแบบสอบถามทางไปรษณีย์ (Nonresponse Bias in Mail Surveys) และ เฮนดริกค์ (Hendrick 1949 :: 52 - 56) ได้ใช้ตัวแบบสมการถดถอยเชิงเส้น (Linear Regression Model) ในการประมาณค่าพารามิเตอร์ของประชากรในกรณีที่เกิดการไม่ตอบเมื่อส่งแบบสอบถามทางไปรษณีย์ นอกจากนี้ ยังมีนักวิจัยอีก 2 ท่าน ที่ศึกษาถึงการปรับความเอนเอียงอันเนื่องมาจากการไม่ตอบในกรณีที่ส่งแบบสอบถามทางไปรษณีย์ โดยใช้ตัวแบบการสุ่มตัวอย่างย่อย (A Subsampling Model) คือ แฮนเซ็น-เฮอร์วิทย์ (HANSEN & HURWITZ) และ เอล-บาดรี้ (M.A EL-BADRY) ซึ่งวิทยานิพนธ์นี้จะศึกษาทั้ง 2 วิธีนี้ดังได้กล่าวมาแล้วเบื้องต้น