บทที่ 5

บทสรุป
(Conclusion)

5.1 หลักเกณฑ์ในการพิจารณาเลือกใช้วิธีการ P.R. ในแบบการสำรวจที่จำนวนรอบแตกต่างกัน

ปัจจัย——————สำคัญที่ต้องนำมาประกอบการพิจารณาตามความสำคัญก่อนหลัง คือ :

ก) Estimator ต้องเป็น Unbiasedness

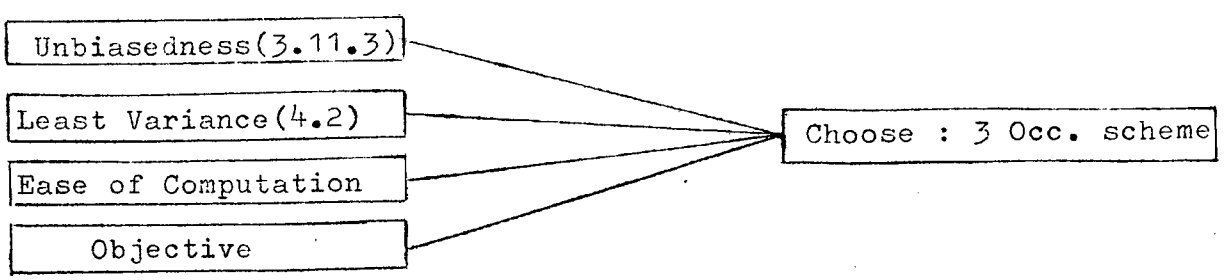
ข) ต้องให้ Gain in Efficiency สูงสุด หรือ Least Variance

ค) สดวกแก่การคำนวน

ง) ขึ้นอยู่กับความมุ่งหมายของวัตถุประสงค์——————ที่วางไว้ก่อนที่จะตัดสินใด ๆ เกี่ยวกับการเลือกใช้แบบแผนใดแบบแผนหนึ่งนั้น ควรที่จะคำนึงถึงข้อเปรียบเทียบแบบแผนที่มีอยู่ว่า แบบแผนใดมีคุณสมบัติเป็นไปตามปัจจัย——————ที่กำหนดไว้ครบถ้วนประการใดเสียก่อน จึงจะตัดสินใจได้ถูกต้องตามหลักวิชาการและเป็นไปตามความมุ่งหมายของวัตถุประสงค์ที่วางไว้

5.2 การพิจารณาเลือกใช้ระหว่างแบบแผนการสำรวจ 2 กับ 3 รอบติดต่อกัน

จากการเปรียบเทียบประสิทธิภาพซึ่งวัดด้วยค่าของเทอม Variance ตามหัวข้อ 4.2 (ก) สมการ (4.2.1) และ PGE ในหัวข้อ 4.2 (ข) ประกอบกับหลักเกณฑ์ที่กำหนดไว้ในหัวข้อ 5.1 นั้น พอจะนำไปสู่ผลการสรุปได้ดังนี้ คือ :



เหตุผลรายละเอียดเกี่ยวกับการเห็นสมควรเลือกใช้แบบ 3 รอบติดต่อกัน มากกว่าที่จะใช้แบบ 2 รอบติดต่อกันนั้น คือ :

ก) Unbiasedness ทั้ง 2 แบบต่างก็เป็น Unbiasedness ซึ่งได้แสดงผลไว้แล้วในบทที่ 3 สมการ (3.5.4) สมการ (3.11.3) สิ่งที่ต้องคำนึงต่อไปคือ Variance ของแบบแผนใดจะเล็กกว่ากัน ถ้าหากกรณีเป็น Biased ทั้ง 2 แบบแผน ก็ควรจะใช้แบบแผนที่มี Variance เล็กกว่า เป็นต้น

ข) Gain in Efficiency สูงสุด หรือ Least Variance สำหรับ Variance ของทั้ง 2 แบบแผนได้ทำการเปรียบเทียบแสดงไว้ในหัวข้อ 4.2 สมการ (4.2.1) และตารางที่ 6 ปรากฏว่า Variance ของการสำรวจ 3 รอบติดต่อกันให้ค่า Variance ที่เล็กกว่าอยู่ภายใต้เงื่อนไข $0 < \mu < 1$; $0 < \rho \leq 1$ ส่วน PGE ก็เช่นเดียวกัน คือ จากตารางที่ 8 จะเห็นได้ชัดเจนว่าแบบ 3 รอบให้เปอร์เซนต์สูงกว่าในแบบ 2 รอบอยู่ประมาณ 6-20% เมื่อ $\mu = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ และ $0.6 \leq \rho \leq 0.8$ และสูงกว่าประมาณ 30-40 เปอร์เซนต์ เมื่อ $0.9 \leq \rho \leq 0.95$ ซึ่ง ρ ในอันตรภาคนี้มักจะเป็นจริงในทางปฏิบัติสำหรับความสัมพันธ์ระหว่างข้อมูลชุดเดียวกัน ซึ่งนับว่าเป็นเปอร์เซนต์ที่สูงมากพอสมควร ยิ่งเมื่อพิจารณาดูเฉพาะแบบ 3 รอบ กับแบบ No.R.U. และ C.R. แล้วจะยิ่งสูงกว่า 20-80 เปอร์เซนต์ ดังนั้นจึงคุ้มที่จะใช้แบบ 3 รอบมากกว่าแบบ 2 รอบ แม้ว่าการได้รับความร่วมมือจากผู้ตอบสัมภาษณ์จะสู้แบบ 2 รอบไม่ได้ บ้างก็ตาม เพราะเนื่องจากการเข้าสอบถามข้อมูลจากผู้ถูกสัมภาษณ์ที่ตกเป็นหน่วยตัวอย่างถึง 3 ครั้ง ย่อมต้องเกิดการไม่ให้ความร่วมมือ แน่เป็นธรรมดา แต่ก็ไม่แน่เสมอกันที่ว่าผู้ที่ตกเป็นหน่วยตัวอย่างถึง 2 ครั้งจะไม่เกิดการไม่ให้ความร่วมมือ ฉะนั้น การต่างกันจึงกล่าวได้ว่าเพียงระดับความไม่ให้ความร่วมมือมากน้อยกว่ากัน ซึ่งยากที่จะวัดออกมาเป็นค่าได้โดยตรง นอกจากจะประเมินผลได้จากความถูกต้องของคำตอบ หรือจากปฏิกริยาที่ได้รับในขณะเข้าสัมภาษณ์ เป็นต้น

ค) ความสดวกในการคำนวน นับว่าเป็นส่วนสำคัญขั้นหนึ่งต่อการพิจารณาในการตัดสินใจเช่นกัน เนื่องจากเมื่อแบบแผนทั้ง 2 ให้ผลเหมือนกันตามข้อ ก. และ ข. แต่มีขั้นตอนการคำนวนที่ยุ่งยากสลับซับซ้อนในเชิงคณิตศาสตร์ (Mathematical Difficulties) ของแบบแผนหนึ่งที่ง่ายกว่า เราก็อาจใช้แบบที่ง่ายเพื่อความสดวกเป็นต้น

สำหรับแบบ 3 รอบ ปัญหาในเชิงคณิตศาสตร์มิได้ยุ่งยากสลับซับซ้อนจนเป็นอุปสรรคต่อการนำมาใช้เลย เพราะสามารถจะทำได้เช่นเดียวกับแบบ 2 รอบ เพียงแต่มีวิธีการคำนวนที่ยาวกว่ากันเท่านั้นเอง ซึ่งในทางปฏิบัติจริงแล้วสิ่งเหล่านี้มิได้เป็นปัญหาสำคัญเนื่องจากเราสามารถอาศัยเครื่องคำนวนสมองกล (Electronic Computer) ในการคำนวนได้ทุกอย่างมีความรวดเร็ว ซึ่งสามารถประมวลผลออกมาในช่วงระยะเวลาอันสั้นและมีความถูกต้องแน่นอน 100 เปอร์เซนต์ แต่ทั้งนี้ต้องหมายความว่าสิ่งที่เราป้อนให้กับเครื่องนั้นถูกต้องด้วย อาทิเช่นโปรแกรมคำนวนการ์ดข้อมูล ฯลฯ ดังนั้นจึงกล่าวได้ว่าปัญหาทางคำนวนจึงเป็นอันหมดไปสำหรับการเลือกใช้แบบ 3 รอบ

จากเหตุผลที่กล่าวมาแล้วทั้ง 3 ข้อ จึงพอสรุปผลได้เป็นข้อเสนอแนะ คือ เห็นสมควรใช้แบบ 3 รอบ สำหรับการทำสำรวจซ้ำที่ทำมากกว่า 1 ครั้งขึ้นไปในแบบ P.R. ตามที่วิทยานิพนธ์ฉบับนี้ได้สร้างมาตราการคำนวน, การเปรียบเทียบประสิทธิภาพ, การแก้ปัญหาและเหตุผลต่าง ๆ ประกอบ แต่อย่างไรก็ตามมิได้หมายความว่าในทางปฏิบัติจำเป็นจะต้องใช้แบบแผนนี้เสมอก็หาไม่ หากแต่ขึ้นอยู่กับภาวะการณ์ตามความจำเป็นหรือตามความมุ่งหมายของวัตถุประสงค์ในขณะใดขณะหนึ่ง ซึ่งแตกต่างกันออกไป จึงได้กำหนดหัวข้อเพิ่มเติมอีกหัวข้อหนึ่งขึ้นเกี่ยวกับการเลือกใช้แบบแผน คือ :

ง)----------ความมุ่งหมายของวัตถุประสงค์ที่ตั้งไว้ ซึ่งโดยทั่ว ๆ ไปแล้วจุดสำคัญหรือหัวใจเกี่ยวกับการสำรวจจากวิธีการสุ่มตัวอย่าง เรามักจะคำนึงถึงประสิทธิภาพของ Variance ก่อนอื่น คือ ต้องมีค่าเล็กและเป็น Unbiased Estimate ด้วย นอกจากนั้นยังคำนึงถึงปัจจัยอื่น เช่น งบประมาณ ระยะเวลามาเกี่ยวข้องด้วย เพราะฉนั้นดังที่ได้กล่าวแล้วว่ายอมขึ้นอยู่กับสภาวะการณ์ตามความจำเป็นและความเหมาะสมในขณะใดขณะหนึ่งซึ่งไม่แน่นอน หากวัตถุประสงค์ เน้นหนักในด้านงบประมาณที่จำกัดและระยะเวลาสั้น ก็อาจจำเป็นต้องหันมาใช้การสำรวจ 2 รอบติดต่อกัน โดยคำนึงถึง Gain ที่จะได้รับ——ไม่เน้นหนัก——นัก หรือหากวัตถุประสงค์เน้นหนักในแง่ Gain ก็อาจต้องหันมาใช้แบบการสำรวจ 3 รอบติดต่อกันเป็นต้น

อนึ่ง สำหรับหลักเกณฑ์การพิจารณาเลือกใช้ค่าอัตราส่วน $(1-\mu)$ ควรเป็นเท่าใดในกรณีการสำรวจ 3 รอบติดต่อกันนั้น ก็คงเป็นไปตามที่ได้กล่าวแล้วในหัวข้อ 4.3

ดังนั้นจาก "วิธีการประมาณในการสำรวจจากตัวอย่างที่มีหน่วยตัวอย่างซ้ำเพียงบางส่วน" ที่ได้กำหนดไว้ในวัตถุประสงค์จึงสามารถแก้ปัญหาต่าง ๆ คือ :

1. <u>ปัญหาทางด้านทฤษฎี</u> คือ สามารถสร้าง Estimator ที่เป็น Unbiased และมีประสิทธิภาพที่ให้ค่า Variance ต่ำกว่าวิธีการอื่นที่นำมาเปรียบเทียบในวิทยานิพนธ์นี้

2. <u>ปัญหาทางด้านปฏิบัติ</u> คือสามารถลดการไม่ให้ความร่วมมือในการตอบข้อสอบถาม การหลีกเลี่ยงการได้รับคำตอบที่ผิด ๆ การลดงบประมาณค่าใช้จ่ายลงในด้านเกี่ยวกับการทำแผนที่แสดงที่ตั้งอาณาเขตของหน่วยตัวอย่าง เนื่องจากการใช้ตัวอย่างใหม่เพียงบางส่วนเท่านั้น และสามารถใช้หน่วยตัวอย่างทดแทนในกรณีที่เกิดการเปลี่ยนแปลงที่ตั้งอาณาเขตของหน่วยตัวอย่าง

นอกจากนี้แล้วยังเป็นประโยชน์ในด้านการเผยแพร่และส่งเสริมระเบียบวิธีสถิติว่าด้วยการสำรวจจากวิธีการตามหัวข้อที่เสนอดังกล่าว ซึ่งนับได้ว่าเป็นวิธีการสมัยใหม่ ———— ———— ทั้งในงานส่วนราชการและวงการธุรกิจที่มีความจำเป็นต้องเก็บรวบรวมข้อมูลตามคาบเวลาต่าง ๆ เพื่อวางแผนในการบริหารกิจการ เช่น การทำสำรวจวิจัยตลาด เพื่อประมาณปริมาณความต้องการสำหรับสินค้าบางประเภทจึงจะสามารถผลิตให้ได้ส่วนกับความต้องการ ซึ่งเป็นการป้องกันการผลิตเกิน ———— และการผลิตขาด ———— เป็นต้น

6.3 <u>ปัญหาเกี่ยวกับประสิทธิภาพของการประมาณเมื่อค่าของ Variance เปลี่ยนไปตามเวลา</u>

ภายใต้ข้อสมมุติฐาน (Assumption) เกี่ยวกับค่าของ Variance ที่

$$S_1^2 = S_2^2 = S_3^2$$

ตามที่ได้ใช้ข้อสมมุติดังกล่าวในวิทยานิพนธ์บทที่ 3 นั้น ก็เพื่อความสดวกและเป็นประโยชน์ต่อการสร้างมาตราการการคำนวนสูตร Variance of Estimated Mean for the Last Occasion ในการเปรียบเทียบประสิทธิภาพของ Variance ดังกล่าวสำหรับแบบแผนที่ต่างกัน แต่โดยทั่ว ๆ ไปแล้วไม่จำเป็นเสมอไปที่ $S_1^2 = S_2^2 = S_3^2$ อาจพิจารณาเป็นในรูป

Non Stationary Assumption $S_2^2 = \lambda_1 S_1^2$

$S_3^2 = \lambda_2 S_2^2 = \lambda_1 \lambda_2 S_1^2$

λ_1 และ λ_2 เป็น arbitrary-parameters

if $\lambda_1 = \lambda_2 = 1$ ก็จะเป็นไปตามข้อสมมุติฐานที่เราใช้ คือ $S_1^2 = S_2^2 = S_3^2$ นั่นเอง

ดังนั้นในการคำนวนหา Variance of Estimated Mean for the Last Occasion ในกรณี Variance เปลี่ยนแปลงไปตามเวลา นั่นคือ $S_1^2 \neq S_2^2 \neq S_3^2$ จึงต้องอาศัย Parameters λ_1 และ λ_2 เข้ามาประกอบการคำนวน ซึ่งขั้นตอนการคำนวนก็คงเป็นไปตามวิธีเดียวกับที่ได้แสดงไว้แล้วในบทที่ 3 และ 4 ทุกประการ โดยที่รายละเอียดการคำนวนและผลไม่ได้รวมอยู่ในขอบข่ายของวิทยานิพนธ์ฉบับนี้ แต่ผู้สนใจศึกษาเพิ่มเติมอาจนำไปค้นคว้าเพื่อหาผลออกมาได้ โดยจะมีประสิทธิภาพเช่นเดียวกันกับกรณี $S_1^2 = S_2^2 = S_3^2$ หรือไม่ในหลักการ

ภาคผนวกเชิงคณิตศาสตร์

Mathematical Appendix

1. การพิสูจน์ $E(\bar{X}_{1,1}) = E(\bar{X}_{1,2}) = \theta_1$ และ $E(\bar{X}_{2,1}) = E(\bar{X}_{2,2}) = \theta_2$

จาก $\bar{X}_{1,1} = \frac{1}{Mn} \sum_{i=1}^{Mn} {}_{1,1}X_i$

$$E(\bar{X}_{1,1}) = E\left[\frac{1}{Mn} \sum_{i=1}^{Mn} {}_{1,1}X_i\right]$$

$$= \frac{1}{Mn} \sum_{i=1}^{Mn} E({}_{1,1}X_i)$$

$$= \frac{1}{Mn} Mn \sum_{i=1}^{N} \frac{1}{N} {}_{1,1}X_i$$

$$= \frac{1}{N} \sum_{i=1}^{N} {}_{1}X_i = \theta_1 \quad \text{--------- (1)}$$

จาก $\bar{X}_{1,2} = \frac{1}{(1-M)n} \sum_{i=1}^{(1-M)n} {}_{1,2}X_i$

$$= E\left[\frac{1}{(1-M)n} \sum_{i=1}^{(1-M)n} {}_{1,2}X_i\right]$$

$$= \frac{1}{(1-M)n} \sum_{i=1}^{(1-M)n} E({}_{1,2}X_i)$$

$$= \frac{1}{(1-M)n} (1-M)n \sum_{i=1}^{N} \frac{1}{N} {}_{1}X_i$$

$$= \frac{1}{N} \sum_{i=1}^{N} {}_{1}X_i = \theta_1 \quad \text{---------(2)}$$

(1) = (2) ดังนั้น $E(\bar{X}_{1,1}) = E(\bar{X}_{1,2}) = \theta_1$

และในทำนองเดียวกันอาจพิสูจน์ได้ว่า

$$E(\bar{X}_{2,1}) = E(\bar{X}_{2,2}) = \theta_2$$

และ $$E(\bar{X}_{3,1}) = E(\bar{X}_{3,2}) = \theta_3$$

2. การพิสูจน์ว่า $V(\bar{X}_{1,1}-\bar{X}_{1,2}) = \frac{1}{\mu n}S_1^2 - \frac{1}{(1-\mu)n}S_1^2$

จาก $V(\bar{X}_{1,1}-\bar{X}_{1,2}) = V(\bar{X}_{1,1}) + V(\bar{X}_{1,2}) - 2\,Cov(\bar{X}_{1,1},\bar{X}_{1,2})$

$$V(\bar{X}_{1,1}-\bar{X}_{1,2}) = V\left(\frac{1}{\mu n}\sum_{i=1}^{\mu n} {}_{1,1}X_i\right) + V\left(\frac{1}{(1-\mu)n}\sum_{i=1}^{(1-\mu)n} {}_{1,2}X_i\right) - \frac{2}{\mu n(1-\mu)n} Cov\left[\sum_{i=1}^{\mu n} {}_{1,1}X_i, \sum_{i=\mu n+1}^{n} {}_{1,2}X_i\right]$$

$$= \frac{1}{(\mu n)^2} V\left(\sum_{i=1}^{\mu n} {}_{1,1}X_i\right) + \frac{1}{(1-\mu)^2 n^2} V\left(\sum_{i=1}^{(1-\mu)n} {}_{1,2}X_i\right) - \frac{2}{\mu n(1-\mu)n} Cov\left[\sum_{i=1}^{\mu n} X_i, \sum_{i=\mu n+1}^{n} {}_{1,2}X_i\right]$$

ในที่นี้จะไม่คำนึงถึงเทอม (f.p.c)

$$= \frac{1}{(\mu n)^2}\cdot \mu n\, V_{1,1}(X_i) + \frac{1}{(1-\mu)^2 n^2}(1-\mu)n\, V({}_{1,2}X_i) - 0$$

เนื่องจาก μn และ $\mu n+1$ independent กัน

$$= \frac{1}{\mu n}S_1^2 + \frac{1}{(1-\mu)n}S_1^2$$ QED.

3. การพิสูจน์ว่า $Cov[(\bar{X}_{1,1}-\bar{X}_{1,2}),\bar{X}_{2,1}] = \frac{-\sigma_{1,2}}{(1-\mu)n}$

จาก $Cov[(\bar{X}_{1,1}-\bar{X}_{1,2}),\bar{X}_{2,1}] = Cov(\bar{X}_{1,1},\bar{X}_{2,1}) - Cov(\bar{X}_{1,2},\bar{X}_{2,1})$

$$= \frac{1}{\mu n(1-\mu)n} Cov\left[\sum_{i=1}^{\mu n} {}_{1,1}X_i, \sum_{i=\mu n+1}^{n} {}_{2,1}X_i\right] - \frac{1}{(1-\mu)^2 n^2} Cov\left[\sum_{i=1}^{(1-\mu)n} {}_{1,2}X_i, \sum_{i=1}^{(1-\mu)n} {}_{2,1}X_i\right]$$

$$= 0 - \frac{1}{(1-\mu)^2 n^2}\cdot(1-\mu)n\, Cov[{}_{1,2}X_i, {}_{2,1}X_i]$$

$$= -\frac{1}{(1-\mu)n}\sigma_{1,2}$$ QED.

ในทำนองเดียวกันอาจพิสูจน์ได้ว่า :

$$Cov[(\bar{X}_{1,1}-\bar{X}_{1,2}),(\bar{X}_{2,1}-\bar{X}_{2,2})] = \frac{-\sigma_{1,2}}{(1-\mu)n}$$

$$Cov[(\bar{X}_{1,1}-\bar{X}_{1,2}),(\bar{X}_{3,1}-\bar{X}_{3,2})] = -\frac{\sigma_{1,3}}{(1-\mu)n}$$

$$Cov[(\bar{X}_{2,1}-\bar{X}_{2,2}),(\bar{X}_{3,1}-\bar{X}_{3,2}) = \frac{-\sigma_{2,3}}{(1-\mu)n}$$